Proceedings of the 8th European Workshop on Advanced Control and Diagnosis 2010

ACD 2010

Editors: Silvio Simani, Marcello Bonfè
Paolo Castaldi & Nicola Mimmo

18 – 19 November 2010

Department of Engineering
The University of Ferrara
Via Saragat, 1
44122 Ferrara (FE)
ITALY

Preface

This series of International Workshops is jointly organised on a rotational basis among the Member Organisations of the former European Institute for Applied Research (IAR), and of the Intelligent Control and Diagnosis (ICD). This year, for the first time, the Advance Control and Diagnosis (ACD) Workshop becomes a European event, and for the third time, it is being been held outside France and Germany.

Since 2003, the ICD Working group has organized an annual Workshop on Advanced Control and Diagnosis, which has brought together academics and engineers in automated systems. The aim is manifold, *i.e.*, to present recent developments in control and diagnosis techniques, to present practical applications or open problems, to provide an opportunity for industrial partners to express their needs and priorities, and to review the current activities in the field of Advanced Control and Diagnosis and their implication in Maintenance.

The 7th Workshop was held in Zielona Gora (Poland) on November 2009 with around 80 participants (http://www.issi.uz.zgora.pl/ACD_2009). It was organised in the continuation of the previous workshops, which took place respectively in:

- 1st Workshop in Duisburg (Germany) - 2003;
- 2nd Workshop in Karlsruhe (Germany) - 2004;
- 3rd Workshop in Mulhouse (France) - 2005;
- 4th Workshop in Nancy (France) - 2006;
- 5th Workshop in Grenoble (France) - 2007;
- 6th Workshop in Coventry (UK) - 2008;

The Department of Engineering at the University of Ferrara (Italy) is delighted to be the host of the 8th such Workshops, and the first in Italy. The emphasis of ACD Workshop is focused on general advanced control and diagnosis issues, and in particular for this year, it comprises about seventy papers, including four plenary invited papers, and the remaining as regular papers for oral presentation. As well as special sessions on Advanced Informatics and Control, Fault Diagnosis, Fault Tolerant Control and Network Control Systems, there are sessions on System Identification, Signal and Image Processing, Industrial Processes, Automotive Control Systems, Control Applications, Control Systems, Modelling for Control and Intelligent Methods for System Identification and Control.

The papers published in these proceedings indicate the growing interest in these important discipline areas in terms of both theoretical development, as well as diversity of application. The ACD2010 Workshop is organised by the Department of Engineering at the University of Ferrara (Italy), by the Automatics and Robotics Group within the Department of Engineering, Ferrara University.

On behalf of the Organisers for the ACD2010 Workshop, we would like to express our gratitude to our colleagues and members of the ACD Directorate, members of the International Programme Committee, the Session Chairs and Co-Chairs and all the Presenters and Authors of the papers attending from twenty-five countries. We also acknowledge the support of the Department of Engineering at the University of Ferrara, and the Consorzio Ferrara Ricerche, at the University of Ferrara.

In anticipation of another valuable and enjoyable meeting, it is our pleasure to welcome all delegates, both old friends and new, to this International and European Workshop here at Ferrara University. We wish you all a pleasant stay in Ferrara, and trust that you will find the Workshop to be of value and leave us having made new friends.

Dr. Silvio Simani & Dr. Marcello Bonfè
(On behalf of the Organisers for ACD2010, Ferrara University, Italy)

Acknowledgements

International Programme Committee

Jan Åslund (Sweden)
Christophe Aubrun (France)
Andrzej Bartoszewicz (Poland)
Sergio Beghelli (Italy)
Gianni Bertoni (Italy)
Sergio Bittanti (Italy)
Mogens Blanke (Denmark)
Jozsef Bokor (Hungary)
Keith J. Burnham (United Kingdom)
Marco Campi (Italy)
Alessandro Casavola (Italy)
Paolo Castaldi (Italy)
Vincent Cocquempot (France)
Maria Letizia Corradini (Italy)
Claudio De Persis (Italy)
Steven X. Ding (Germany)
Roberto Diversi (Italy)
Andras Edelmayer (Hungary)
Chris Edwards (United Kingdom)
Miroslav Fikar (Slovakia)
Giuseppe Franzè (Italy)
Erik Frisk (Sweden)
Sylviane Gentil (France)
Michael J. Grimble (United Kingdom)
David Henry (France)
Marina Indri (Italy)
Sirrka L. Jamsa-Jounela (Finland)
Andrzej Kasprzak (Poland)
Paul King (United Kingdom)
Michel Kinnaert (Belgium)
Józef Korbicz (Poland)
Leszek Koszalka (Poland)
Jan Kościelny (Poland)
Suzanne Lesecq (France)
Antoni Ligęza (Poland)
Jan Lunze (Germany)
Didier Maquin (France)
Elena Mainardi (Italy)
Massimiliano Mattei (Italy)
Nicola Mimmo (Italy)
Hans Henrik Niemann (Denmark)
Thomas Parisini (Italy)
Krzysztof Patan (Poland)
Ron J. Patton (United Kingdom)
Andrzej Pieczyński (Poland)

Marios M. Polycarpou (Greece)
Vicenç Puig (Spain)
Joseba Quevedo (Spain)
José Ragot (France)
José Sá da Costa (Portugal)
Dominique Sauter (France)
Piotr Skrzypczynski (Poland)
Miroslav Simandl (Czech Republic)
Dirk Soeffker (Germany)
Marcel Staroswiecki (France)
Ralf Stetter (Germany)
Jacob Stoustrup (Denmark)
Michele Taragna (Italy)
Piotr Tatjewski (Poland)
Didier Theilliol (France)
Andrea Tilli (Italy)
Dariusz Uciński (Poland)
Maria Elena Valcher (Italy)
Andreas Varga (Germany)
Antonio Visioli (Italy)
Holger Voos (Germany)
Marcin Witczak (Poland)

Local Organising Committee

Silvio Simani	Chairman
Marcello Bonfè	Vice-chairman
Sergio Beghelli	Honorary chairman
Elena Mainardi	Organization Chairman
Paolo Castaldi	Program Chairman
Nicola Mimmo	Student Paper Chairman
Mauro Mazza	Local Arrangement Chairman

Co-sponsoring Organisations

Department of Engineering, University of Ferrara (ENDIF-UNIFE)
Consorzio Ferrara Ricerche, University of Ferrara
VM Motors (Cento, Ferrara)

Supports

The Intelligent Control and Diagnosis (ICD, http://www.icd.cran.uhp-nancy.fr/) working group founded in 1998, leads to new developments and applications in the field of automatic control and fault diagnosis. The aim of the ICD working group is to explore research opportunities in the

direction of Fault Diagnosis and Fault-tolerant Control for technical systems. ICD Research activities can be summarized as follows:

- Development of advanced methods with applications to automatic control and fault detection and isolation (FDI);
- Design of FTC strategy providing an optimal performance of the reconfigured system according to the reliability measure in order to ensure the dependability of the system and the human safety;
- Investigation of typical application areas and technology transfer to industrial areas of special interest for control and diagnosis of technical systems. The domains of application concern different types of systems such as embedded systems, distributed systems, networked systems.

Within this working group, the members co-operate in different ways, one important one are joint European projects. The aim for the future is to initiate more of such projects, especially in co-operation with industry and to tackle with advanced methods in Fault Tolerant Control (FTC) framework in order to improve the human safety and dependability of the system.

The chairs are the Prof. C. Aubrun and the Prof. D. Theilliol. The members and partners are:

- Gerhard-Mercator-Universitaet Duisburg, Germany
- Centre de Recherche en Automatique de Nancy, France
- GIPSA-Lab Grenoble, France
- University of Karlsruhe, Germany
- Control Theory and Applications Centre, Coventry University, United Kingdom
- Institute of Control and Computation Engineering, University of Zielona Gora, Poland
- Department of Engineering, University of Ferrara, Italy
- Automatic Control Department, Universidad Politecnica de Cataluna, Spain
- Engineering Department, The University of Hull, United Kingdom

Members and Partners

University of Ferrara

Contents

Preface…….. iii
Acknowledgements…….. iv
International Programme Committee…… iv
Local Organising Committee…… v
Co-sponsoring Organisations…… v
Supports…… v
Members and Partners……... vii

Contents……... viii
Plenary Papers……..……..……..…… 2
A norm-based point of view for fault diagnosis: Application to aerospace missions
David Henry…….. 4
Design and Evaluation of Reconfiguration-based Fault Tolerance using the Lattice of System Configurations
Marcel Staroswiecki…….. 17
New Perspectives for Research in Fault Tolerant Control
Ron J Patton…….. 36
Developments in bilinear systems modelling and control with industrial applications
Keith Burnham…….. 37

Regular Papers……... 38
Design of Robust Fault Detection Filters for Plants with Quantized Information
Maria Letizia Corradini, Andrea Cristofaro, Roberto Giambò, and Silvia Pettinari…….. 40
Aircraft Sensor Fault Detection and Accommodation by Some Conventional Controllers
Emre Kiyak and Fikret Caliskan……... 46
Performance Comparison of Different Types of Controllers for the Control of the Pitch Angle of an Aircraft
Gulay Iyibakanlar and Emre Kiyak…….. 52
Fault Detection and Estimation in Networked Control Systems
Ignacio Peñarrocha and Roberto Sanchis…….. 58
Optimization of a Water for Injection Control System for a Pharmaceutical Plant
Antonio Visioli, Massimiliano Ammannito, Michele Caselli and Marco Incardona…… 64
Diagnosis for the Reliability Improvement of Embedded Systems
Ouadie Bennouna, Houcine Chafouk and Jean-Philippe Roux…… 68
Smith Predictor Based Control of Continuous-Review Perishable Inventory Systems with a Single Supply Source
Przemyslaw Ignaciuk and Andrzej Bartoszewicz…….. 73
Smoothing in Multiple Model Change Detection for Stochastic Systems
Ivo Puncochar, Jindrich Dunik and Miroslav Simandl……. 79
Predictive Fault-Tolerant Control of Takagi-Sugeno Fuzzy Systems
Lukasz Dziekan and Marcin Witczak…… 85
Communication Gains Design in a Consensus Based Distributed Change Detection Algorithm
Nemanja Ilic and Srdjan Stankovic……... 91
Control of Independent Mobile Robots by Means of Advanced Monitoring
Lothar Seybold, Jaroslaw Krokowicz, Krzysztof Patan, Ralf Stetter and Anderas Paczynski……... 95

Modelling of Positive Displacement Pumps for Monitoring, Planning, Control and Diagnosis
Stefan Kleinmann, Muhammad Fairusz Abdul Jalal and Ralf Stetter…….. 101
Concept of an Advanced Monitoring, Planning, Control and Diagnosis System for Autonomous Vehicles
Lothar Seybold, Andrzej Pieczyński, Andreas Paczynski and Ralf Stetter…… 107
Reliability Assessment of Technical Devices Based on Degradation Data and Stochastic Equations
Ryszard Kopka……... 113
Intelligent Techniques for Faults Diagnosis and Prognosis of CHP Plant with Gas Turbine Engine
Luigi Miozza, Andrea Monteriù, Alessandro Freddi and Sauro Longhi…….. 119
Periodic Linear Time-Varying System Norm Estimation Using Running Finite Time Horizon Transfer Operators
Przemyslaw Orlowski…… 125
An Application of Model Based Fault Detection in Power Plants
Goran Kvascev, Predrag Tadic and Zeljko Djurovic……... 130
Validation of a New Time Delay Estimation Method for Control Performance Monitoring
Markus Stockmann, Robert Haber and Ulrich Schmitz…… 135
Estimation and Prediction of Global Radiation by Meteosat Image Processing
Ali Zaher, Thiery Frédérik, Yao N'Goran, and Adama Traore ……..... 141
Advanced and Predictive Diagnosis on the Example of Pump Systems
Stefan Kleinmann, Anna Dabrowska, Domenico Leonardo, Ralf Stetter and Agathe Koller-Hodac……. 146
Evaluation Scheme of Task Allocation in Mesh Connected Processors with Metaheuristic Algorithms
Wojciech Kmiecik, Leszek Koszalka, Iwona Pozniak-Koszalka, and Andrzej Kasprzak….. 152
Bus Route Optimization: an Experimentation System and Evaluation of Algorithms
Krzysztof Golonka, Leszek Koszalka and Andrzej Kasprzak…… 158
Routing in Mobile Ad-hoc Networks: an Experimentation System and Evaluation of Algorithms
Maciej Foszczynski, Marek Adamczyk, Kamil Musial, Leszek Koszalka, Iwona Pozniak-Koszalka, and Andrzej Kasprzak……... 164
Testing SQL Queries: an Experimentation System and Efficiency Evaluation
Michal Hans, Pawel Kmiecik, Iwona Pozniak-Koszalka, and Andrzej Kasprzak……..... 170
Properties of NCGPC Applied to Nonlinear SISO Systems with a Relative Degree One or Two
Marcelin Dabo, Nicolas Langlois and Houcine Chafouk……. 174
Improvement of the Decoupling Feature of Decentralized Predictive Functional Control
Khaled Zabet and Haber Robert……... 180
Equality Constraints in Sensor Faults Reconfigurable Control Design
Dusan Krokavec and Anna Filasova……. 184
Set-Point Reconfiguration in Case of Severe Actuator Fault
Boumedyen Boussaid, Christophe Aubrun and Naceur Abdelkrim…….. 190
Connections of Functional States for Automaton Identification: Application in a Steam Generator Monitoring
Javier F. Botia, Henry O. Sarmiento and Claudia Isaza…….. 196
A GMDH Toolbox For Neural Network-Based Modelling
Marcel Luzar and Marcin Witczak…… 202

Decoupling Model Predictive Control in a Non-Minimal State Space Representation
Ulrich Hitzemann and Keith J. Burnham…….. 207
Design of Unknown Input Reconstruction Algorithm in Presence of Measurement Noise
Malgorzata Sumislawska, Tomasz M. Larkowski and Keith J. Burnham……. 213
Fault Tolerant Control Schemes for Nonlinear Models of Aircraft and Spacecraft: Preliminary Results
Paolo Castaldi, Nicola Mimmo and Silvio Simani…… 217
Robust Model Matching for Geometric Fault Detection Filters: A Commercial Aircraft Example
Jozsef Bokor, Peter Seiler, Balint Vanek, Gary J. Balas……..…… 223
System Programmable Logic Controller Computer Aided Development Procedure
Sergio Chiesa, Sabrina Corpino and Giovanni Medici…… 229
Task-Oriented Modelling of Rugged Terrain from Sparse Range Data
Dominik Belter, Przemyslaw Labecki and Piotr Skrzypczynski…… 235
Flight Path Optimisation Using Primitive Manoeuvres: A Particle Swarm Approach
Luciano Blasi, Simeone Barbato and Massimiliano Mattei…….. 241
A Fault Detection Filter Design Method for Hybrid Switched Linear Parameter Varying Systems
Gianfranco Gagliardi, Alessandro Casavola, Domenico Famularo and Giuseppe Franzè. 247
Improvement of the Sensitivity of T² Quality Control Charts by Grouping of Variables
Thomas Friebel and Robert Haber……... 253
A Constrained Strategy to Control Plasma Shape in ITER
C. V. Labate, M. Mattei, D. Famularo, F. Koechl, and V. Parail……..….. 257
Fault Detection and Isolation of Wind Turbines: Application to a Real Case Study
Pep Lluis Negre, Vicenç Puig and Isaac Pineda…….. 263
Second-Order Sliding Modes and Soft Computing Techniques for Fault Detection
Milan Rapaic, Zoran Jelicic, Alessandro Pisano, and Elio Usai……. 271
Unknown-Input Observation Techniques in Open Channel Hydraulic Systems
Siro Pillosu, Alessandro Pisano, and Elio Usai…..……. 278
Unknown Input Observer with Sliding Mode Disturbance Estimator for the Diffusion PDE
Alessandro Pisano, Stefano Scodina, and Elio Usai……. 284
An Efficient Algorithm For Fault Tolerant Sensor Network Design
Firas Rouissi, Ghaleb Hoblos and Nicolas Langlois…… 290
Multi-Scale PCA-Based Fault Diagnosis for Rotating Electrical Machines
Francesco Ferracuti, Andrea Giantomassi, Gianluca Ippoliti, and Sauro Longhi ……. 296
Reconfiguration of Over-Actuated Consecutive-k-out-of-n: F Systems Based on Bayesian Network Reliability Model
Philippe Weber, Christophe Simon and Didier Theilliol…….. 302
Fault Detection in Flat Systems by Constraint Satisfaction and Input Monitoring
Ramatou Seydou, Tarek Raissi, Ali Zolghadri and David Henry……. 308
Communication Sequence Design in Networked Control Systems With Communication Constraints: A Graphic Approach
Sinuhe Martinez-Martinez, Hossein Hashemi-Nejad and Dominique Sauter……... 314
Comparison on Control Allocation Methods For The High Altitude Performance Demonstrator
V. Scordamaglia, M. Mattei, C. Calabrò, A. Sollazzo, F. Corraro…….….. 320
Temporal Reliability Analysis of Embedded Systems
Afifa Ghenai and Mohamed Benmohammed…… 326
Data–Driven and Model–Based Fault Diagnosis of Wind Turbine Sensors
Silvio Simani, Paolo Castaldi and Marcello Bonfè……... 332

Central sensor cluster simulation for anti-lock-braking system validation using hardware-in-the loop
Pawel Kret, Keith, J. Burnham, Leszek Koszalka and Alexandros Mouzakitis…… 339
Extended Kalman Filter Approach for Road Condition Estimation: a preliminary study
Mariusz Ruta and Keith Burnham……. 344
Diagnostics of distributed faults in ball bearings by means of vibration cyclostationary indicators
Gianluca D'Elia, Simone Delvecchio, Marco Cocconcelli and Giorgio Dalpiaz…… 350
Robust Fault Detection of Nonlinear Systems using Local Linear Neuro-Fuzzy
Hasan Abbasi Nozari, Mahdi Aliyari Shooredeli and Silvio Simani…… 356
Fault Detection and Isolation of Tennessee Eastman Process Using Improved RBF Network by Genetic Algorithm
Somayeh Hekmati Vahed, Mohammad Mokhtare, Hassan Abbasi Nozari, Mahdi Aliyari Shoorehdeli and Silvio Simani…….. 362
HVAC system energy consumption dependency on control set-point selection
Ivan Zajic, Tomasz Larkowski, Dean Hill and Keith Burnham…… 368
Fuel moisture content analysis as a basis for process monitoring of a BioGrate boiler
Alexandre Boriouchkine, Alexey Zakharov and Sirkka-Liisa Jämsä-Jounela…….. 374
Fault Detection and Accommodation of the Boiler Unit Using State Space Neural Networks
Andrzej Czajkowski and Krzysztof Patan…….. 380

Index of Authors…….. 387

Plenary Papers

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

A Norm-based Point of View for Fault Diagnosis: Application to Aerospace Missions

David HENRY

IMS Lab / ARIA team – Bordeaux 1 University
Bordeaux, FRANCE, (e-mail: david.henry@ims-bordeaux.fr)

Abstract:

This paper deals with norm-based FDI (Fault Detection and Isolation) techniques for both LTI and LPV systems. The investigated techniques can be seen as a nice and practically relevant framework in which various design goals and trades-off are formulated and managed. It is shown that the design problem can be formulated as an optimization problem that can be solved by numerically powerful LMI-based techniques. The output of the design is a filter for Fault Detection, or a bank of filters for Fault Detection and Isolation. The approach has been developed by the author at IMS/LAPS, Bordeaux, see reference section. The developed techniques have been successfully applied to a number of aerospace applications, e.g. satellite, atmospheric re-entry and rendezvous missions.

1. INTRODUCTION

During the last decade, certain basic results concerning robustness of FDI filters using H_∞-based optimisation techniques have appeared. The majority of the presented studies involve the use of a slightly modified H_∞ fault estimator, i.e. the design objective is to minimize the effect of the fault signal, the disturbances and the modelling errors on the fault estimation error, in a H_∞-norm sense.
Residual generation is different from fault estimation because it does not only require the disturbances and model perturbations attenuation. The residual has to remain sensitive to faults while guaranteeing robustness against unknown inputs. This motivates (Ding, 1996; Niemann, 1999; Chen, 1999; Ding, 2000; Zhong, 2003, Liu, 2005; Jaimoukha, 2006) to introduce the H_∞/H- paradigm, i.e. robustness objectives are considered using the H_∞ norm while the fault sensitivity specifications are expressed using the H- norm formulation. ARE-based solutions, eigenstructure assignment, genetic algorithms and Linear Matrix Inequality (LMI) based solutions were developed by the authors to derive the optimal selection of the residual generator.

The majority of methods discussed above involve the use of an open-loop model of the system in spite of that the resulting FDI unit is supposed to supervise the system under closed-loop feedback configuration. In such situations, faults may be covered by control actions and the early detection of process faults (low frequency faults) is clearly more difficult. The control signal directly influences the FDI output when there is modelling uncertainty present. The feedback controller can then have the effect of desensitising the residual signals and deteriorate the FDI unit capability of detecting incipient faults.
A solution may then consist in the so-called integrated design of control and diagnosis systems where a robust controller and a fault detector are designed together by optimizing a set of mixed control and FDI objectives (Stoustrup, 1997; Khosrowjerdi, 2004). However, because in many systems, the already in place controller is certified, this solution cannot be applied.
Furthermore, from a practical point of view, it is convenient to take advantage of hardware redundancy. Thus the problem of optimally (in some sense) merging available information should also be addressed as an integral part of the design of a model-based FDI scheme.

These problems motivate (Henry 2002, 2003, 2005, 2008, 2009) to develop a method to design FDI schemes within the H_∞/H- paradigm, that *i)* takes into account directly the controller actions within the design procedure; *ii)* merges optimally and systematically available information coming from sensors and control signals.
The H_∞/H- based FDI techniques are generally reputed to give robust but conservative solutions. The problem comes from the fact that, once the diagnostic filter is designed, no systematic analysis procedure is proposed to refine and manage the design trade-offs. It is clear that if the design method is associated with a suitable post-analysis process, an iterative refinement process can be established to get a good balance between different design trade-offs, and to get "as close as possible" to the required robustness/performance specifications, there is not any reason for the final result to be conservative
Similarly to the H_∞ design / μ-analysis cycle used in the robust control community, the method proposed in (Henry 2002, 2003, 2005, 2008) provides a solution to the aforementioned problems by providing a complete design/analysis cycle:

- With regards to the design task, the procedure aims to generate a structured residual vector r in the following general form

$$r(s) = z(s) - \hat{z}(s), \quad z(s) = M_y y(s) + M_u u(s)$$

$$\hat{z}(s) = L(s)\begin{pmatrix} y(s) \\ u(s) \end{pmatrix}, \quad u(s) = K(s)y(s) \qquad (1)$$

where K denotes the controller. $\hat{z}$ is an estimation of $z = M_y y + M_u u$, a subset of available measurements y and inputs u. M_y, M_u are two residuals structuring (constant) matrices and $L(s)$ is a (stable) dynamical filter. The proposed method consists in jointly designing M_y, M_u and $L(s)$ such that the effects that faults have on the residuals r are maximized in the H_- norm sense whilst minimizing the influence of unknown inputs and model uncertainties, in the H_∞ norm sense. Furthermore, it is shown how robust poles assignment and H_{2g} specifications can be considered. H_{2g} specifications and regional filter poles assignment are convenient to tune the transient response and to enforce some minimum decay rate of the residuals. This feature becomes very important from a decision making point of view, as the residual is post-processed by a hypothesis-based test to make a final decision about the fault.

- With regards to the post-design analysis procedure, a test is proposed to check if all FDI objectives are achieved in the face of specified structured and/or unstructured model perturbations. The problem is formulated using an appropriate performance index, defined with respect to the effects of underlying faults on the residual signal. As outlined above, the robust residual generation problem is not equivalent to the optimal fault estimation problem which is a counterpart of robust control. Testing the performances of residual generators results in a min-max optimization problem which cannot be formulated and solved using the classical "μ-analysis" framework. The method proposed by (Henry 2002, 2003, 2005) provides a remarkably powerful solution to the problem by a FDI-oriented generalized μ-analysis procedure, denoted by the authors the μg-analysis procedure.

This method can be seen as a nice and practically "advanced" framework in which various design goals and trades-off are formulated and managed. It corresponds to a complete design/analysis cycle and has the following advantages:

i) Systematic formulation of different design trade-offs.
- H_∞ specifications are convenient to enforce robustness to model uncertainty (e.g. external disturbances, parametric uncertainties and neglected dynamics) and to take into account frequency-domain specifications.
- H- specifications are useful for fault sensitivity requirements over specified frequency ranges.
- H_2 objectives allow us to take into account the stochastic nature of disturbances and measurement noises
- H_{2g} specifications and regional filter poles assignment are convenient to tune the transient response and to enforce some minimum decay rate of the residuals. This feature becomes very important from a decision making point of view, as the residual is post-processed by a hypothesis-based test to make a final decision about the fault.

ii) The residuals structuration matrices are jointly optimised with the dynamical part of the FDI filter. Their role is to merge optimally the available on-board measurement and control signals to build the fault indicating signal.

iii) The control system can be included explicitly in the design.

iv) The μg tool is used as FDD-oriented performance measure: similarly to the μ-analysis procedure that allows for checking the robust performance of any LTI control law, the μg tool can be used as a general FDD-oriented performance measure for LTI model-based fault diagnosis scheme.

NOTATIONS:

In dealing with vectors, the Euclidean norm is always used and is written without a subscript; for example $\|x\|$. Similarly in the matrix case, the induced vector norm is used $\|A\| = \bar{\sigma}(A)$ where $\bar{\sigma}(A)$ denotes the maximum singular value of A. Signals, for example $w(t)$ or w, are assumed to be of bounded energy, and their norm is denoted by $\|w\|_2$, i.e. $\|w\|_2 = \left(\int_{-\infty}^{+\infty} \|w(t)\|^2 dt\right)^{1/2} < \infty$. Linear models, for example, $P(s)$ or simply P, are assumed to be in RH_∞, real rational functions with $\|P\|_\infty = \sup_\omega \bar{\sigma}(P(j\omega)) < \infty$. In accordance with the induced norm, the smallest gain of a transfer matrix P is defined according to $\|P\|_- = \inf_{\omega\in\Omega} \underline{\sigma}(P(j\omega))\ \Omega = [\omega_1, \omega_2]$, where $\underline{\sigma}(P(j\omega))$ denotes the minimum non-zero singular value of the complex valued matrix $P(j\omega)$ and $\Omega = [\omega_1, \omega_2]$, the evaluated frequency range in which $\underline{\sigma}(P(j\omega)) \neq 0$.

Referring to LPV systems, the worst-case RMS gain from input signal u to output signal $y = P(\theta)u$, i.e. the H_∞-norm for LPV systems, is defined according to $\|P(\theta)\|_\infty = \sup_{\substack{\forall\theta \\ \|u\|_2 \neq 0}} \frac{\|y\|_2}{\|u\|_2}$. As explained in the previous sections, for LTI systems, $\|P\|_\infty$ is accompanied by the non-zero smallest gain of P, that is the H_--index, which is the restriction of $\inf_\omega \underline{\sigma}(P(j\omega))$ to a finite frequency domain Ω. This motivated (Grenaille *et.al.*, 2008) to introduce the evaluation criteria $\|P(\theta)\|_{sens} = \inf_{\substack{\forall\theta \\ \|u\|_e \neq 0}} \frac{\|y\|_e}{\|u\|_e}$ where $\|w\|_e = \left(\frac{1}{2\pi}\int_\Omega \|w(j\omega)\|^2 d\omega\right)^{1/2}$ is the restriction of $\|w\|_2$ to Ω. In the special case of dim(u)=1 and/or dim(y)=1, it has been shown in (Henry *et al*, 2009) that $\|P(\theta)\|_{sens}$ is the generalization of $\|P\|_-$ to LPV case and it follows the definition of the the H- norm for LPV systems

$\|P(\theta)\|_- = \inf_{\substack{\forall\theta \\ \|u\|_e \neq 0}} \frac{\|y\|_e}{\|u\|_e}$. The interested reader can refer to (Grenaille *et.al.*, 2008) and (Henry *et al.*, 2009) for more details.

Block diagrams are intensively used to represent interconnections of systems. For example, the structure shown in Figure 1 represents the equations $v = \Delta w$, $w = P_{11}v + P_{12}u$ and $y = P_{21}v + P_{22}u$.

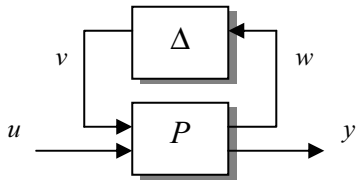


Fig. 1: The interconnection structure of systems

In terms of input *u* and output *y*, this can be expressed as a linear fractional representation (LFR) $y(s) = (\Delta(s) * P(s))u(s)$, where $(\Delta(s) * P(s))$ is the star product of matrices Δ and *P* defined according to : $(\Delta(s) * P(s)) = P_{21}(s)\Delta(s)(I - P_{11}(s)\Delta(s))^{-1}P_{12}(s) + P_{22}(s)$.

2. MODELS AND FAULT ASSUMPTIONS

2.1. Modelling the faulty system

For a realistic representation of the faulty system, it is necessary to model the effects of various faults. According to (Isermann, 1997), faults can be classified as “additive faults” and “multiplicative faults”. In the proposed theoretical developments, it is necessary to have an additive representation of faults. An approximation of the fault model is used to “multiplicative fault” cases. This approximation makes sense as long as the (controlled) system keeps stability in faulty situations. The interested reader can refer to (Isermann, 1997) and (Frank, 2001) for a discussion of such an approximation.

2.2. Modelling the faulty system

The H∞/H- method proposed in (Henry 2002, 2003, 2005, 2008) is classified in model-based approaches. From this last facet, it is necessary to have a LTI or LPV representation of the monitored system. Some addition features about the type of faults, the measurement noises, …etc… can be used to find the best FDI filter for a given application. As mentioned in the introduction, the monitored system is controlled by a validated and certified control law. The control signal provided by this in place controller directly influences the FDI output. To avoid a deterioration of the FDI unit capability of detecting faults due to the feedback controller, a LTI or LPV representation of the regulator can be necessary. From applicability point of view, this last point allows to consider, e.g. gain-scheduling LTI controllers, nonlinear dynamic inversion based control laws, see for instance (Papageorgiou & Glover, 2005).

3. THEORETICAL FOUNDATIONS

3.1. The H∞/H- design procedure

3.1.1. Problem setting

Consider the following model in the LFR form placed in a feedback control loop (see Figure 2)

$$y(s) = (\Delta(s) * P(s)) \begin{pmatrix} d(s) \\ f(s) \\ u(s) \end{pmatrix}, u(s) = K(s)y(s) \quad (2)$$

The system model consists in a nominal LTI (Linear Time Invariant) model *P* and a perturbation block $\Delta \in \underline{\Delta} : \|\Delta\|_\infty \leq 1$ acting on the nominal model. $\underline{\Delta}$ describes the set of all perturbations of a prescribed structure, i.e.

$$\underline{\Delta} = \left\{ blockdiag\left(\delta_i^r I_{k_i}, \delta_j^c I_{k_{ji}}, \Delta_l\right) \right\} \quad (3)$$

where $\delta_i^r I_{k_i}, i = 1..m_r$, $\delta_j^r I_{k_j}, j = 1..m_c$ and $\Delta_l, l = 1..m_C$ are known respectively as the "repeated real scalar" blocks, the "repeated complex scalar" blocks and the "full complex" blocks. It is assumed that all model perturbations are represented by Δ. Exogenous disturbances (including measurement noises) are denoted *d* and *f* is used to represent faults affecting the plant. The signals *v* and *w* are internal to the model. *K* denotes any LTI controller.

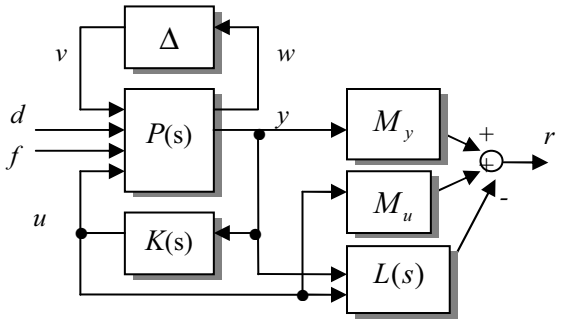


Fig. 2: The FDI filter design problem

Let *f* entering in $((\Delta * P) * K)$ be detectable faults and the residual vector *r* be defined according to eq. (1). The goal is to derive simultaneously M_y, M_u and the state space matrices of the dynamical filter *L* such that the residual vector *r* meets the following specifications:

(S.1)- $\|T_{rd}\|_\infty < \gamma_1$ for all perturbations model $\Delta \in \underline{\Delta} : \|\Delta\|_\infty \leq 1$, where T_{rd} denotes the closed-loop transfer between *r* and *d*.

(S.2)- $\|T_{rf}\|_- > \gamma_2$ over a specified frequency range Ω for all $\Delta \in \underline{\Delta} : \|\Delta\|_\infty \leq 1$. T_{rf} denotes the closed-loop transfer between *r* and *f*, and Ω is the frequency range where the energy of the faults is likely to be concentrated.

The specification (S.1) represents the worst-case robustness of the residual to disturbances d for all specified model perturbations, in the H_∞ norm sense. Under plant perturbation, the effect that the exogenous disturbances acting on the system have on the residual, can greatly increase. The fault detection performance may then be considerably degraded. A robust fault sensitivity specification is then needed to maintain a detection performance level of the FDI unit. Here the smallest gain of T_{rf} is used to guarantee the worst-case sensitivity of the residual to faults (see specification (S.2)). It is clear that the smaller γ_1 and the bigger γ_2 are, the better the fault detection performances will be.

3.1.2. The quasi-standard setup

Generally speaking, to achieve high FDI performances, model-based FDI schemes use disturbance, measurement noise and fault models into the design procedure. Here, such models are expressed in terms of shaping filters, i.e. of desired gain responses for the appropriate closed-loop transfers. The objectives are then turned into uniform bounds by means of the shaping filters. To proceed, let W_d and W_f be the (dynamical) shaping filters associated to the robustness and fault sensitivity objectives defined such that

$$\|W_d\|_\infty \le \gamma_1, \|W_f\|_- \ge \gamma_2 \tag{4}$$

Assume that W_d and W_f are invertible (this can be done without loss of generality because it is always possible to add zeros in $W_d(s)$ and $W_f(s)$ to make them invertible). Thus, it is obvious that if the condition

$$\left\|T_{rd}W_d^{-1}\right\|_\infty < 1 \ \forall \Delta \in \underline{\Delta} : \|\Delta\|_\infty \le 1 \tag{5}$$

is satisfied, then the robustness design specification (S.1) yields.
Now, we need the following proposition to transform the fault sensitivity specification (S.2) into a H_∞ requirement.

Lemma (Henry2005): Consider the shaping filter W_f defined above. Let W_F be a right invertible transfer matrix so that $\|W_f\|_- = \frac{\gamma_2}{\lambda}\|W_F\|_-$ and $\|W_F\|_- > \lambda$ where $\lambda = 1+\gamma_2$. Define the signal $\tilde{r}$ such that $\tilde{r}(s) = r(s) - W_F(s)f(s)$, see figure 3 for easy reference.

Then a sufficient condition for the fault sensitivity specification (S.2) to hold, is

$$\left\|T_{\tilde{r}f}\right\|_\infty < 1 \ \forall \Delta \in \underline{\Delta} : \|\Delta\|_\infty \le 1 \tag{6}$$

where $T_{\tilde{r}f}$ denotes the transfer between $\tilde{r}$ and f ■

Using the above lemma, the H_∞/H- filter design problem can be re-casted in a fictitious H_∞ framework: Using linear fractional algebra and including $\gamma_1, \lambda, W_F, W_d$ and K into the model P, one can derive from eq. (2) a new model $\tilde{P}(M_y, M_u)$ depending of the residual structuration matrices M_y, M_u so that (see figure 3)

$$\begin{pmatrix} r(s) \\ \tilde{r}(s) \end{pmatrix} = \left(\left(\Delta(s) * \tilde{P}(M_y, M_u, s)\right) * L(s)\right)\begin{pmatrix} \tilde{d}(s) \\ f(s) \end{pmatrix} \tag{7}$$

Then, by combining both the H_∞ requirements eq. (5) and eq. (6) into a mixed single H_∞ constraint, it follows that a sufficient condition for specifications (S.1) and (S.2) to hold is:

$$\left\| \begin{matrix} T_{rd}W_d^{-1} \\ T_{\tilde{r}f} \end{matrix} \right\|_\infty < 1 \ \forall \Delta \in \underline{\Delta} : \|\Delta\|_\infty \le 1 \tag{8}$$

With eq. (7) and by virtue of the small gain theorem, it follows that a sufficient condition is:

$$\left\|\tilde{P}(M_y, M_u) * L\right\|_\infty < 1 \tag{9}$$

This equation seems to be similar to a standard H_∞ equation. In fact, this is not the case since the transfer $\tilde{P}(M_y, M_u)$ depends of M_y, M_u that are a part of the solution we are seeking. A solution may then consist in chosen heuristically them. However, there is no guarantee to the optimal solution.

To solve this problem, a SDP (Semi Definite Programming) formulation is derived in (Henry, 2003, 2005) by means on the bounded real lemma (Boyd,94) and the projection lemma (Apkarian,94).

3.1.3. The SDP formulation

Let $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$, the state space matrices of $\tilde{P}(M_y, M_u)$, be partitioned in accordance with

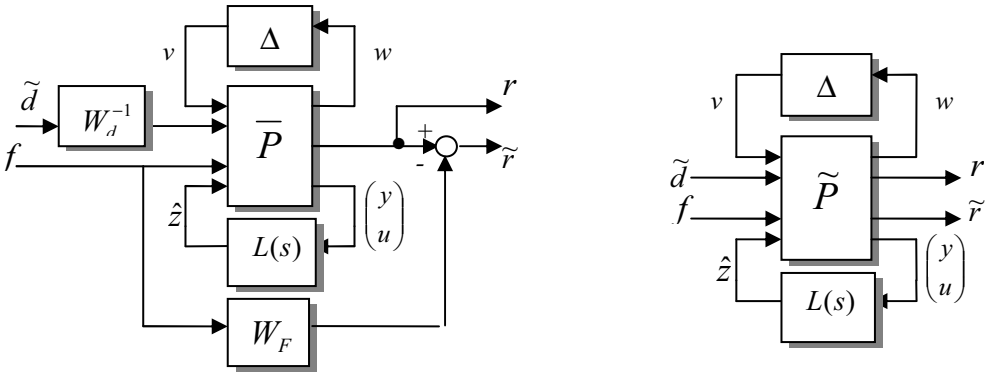


Fig 3 : The quasi standard setup

$$\widetilde{B} = \begin{pmatrix} \widetilde{B}_1 & \widetilde{B}_2 \end{pmatrix},\ \widetilde{C} = \begin{pmatrix} \widetilde{C}_1 \\ \widetilde{C}_2 \end{pmatrix}, \widetilde{D} = \begin{pmatrix} \widetilde{D}_{11} & \widetilde{D}_{12} \\ \widetilde{D}_{21} & \widetilde{D}_{22} \end{pmatrix} \tag{10}$$

It can be verified that, by definition, $\widetilde{B}_2 = 0$ and $\widetilde{D}_{22} = 0$ that outlines the fact that the filter operates in open loop with respect to the system.

The following proposition solves the problem. A complete proof of this proposition can be found in (Henry & Zolghadri, 2005):

Proposition (Henry & Zolghadri, 2005):
Let $W = \begin{pmatrix} \widetilde{C}_2 & \widetilde{D}_{21} \end{pmatrix}^{\perp}$. *There exists a solution of (9) if and only if there exist* $\gamma < 1$, M_y, M_u *and two symmetric matrices R, S solving the following SDP problem*

min γ s.t.

$$\begin{pmatrix} \widetilde{A}R + R\widetilde{A}^T & R\hat{C}_1^T & \widetilde{B}_1 \\ \hat{C}_1 R & -\gamma I & \hat{D}_{11} \\ \widetilde{B}_1^T & \hat{D}_{11}^T & -\gamma I \end{pmatrix} < 0$$

$$\begin{pmatrix} W & 0 \\ 0 & I \end{pmatrix}^T \begin{pmatrix} S\widetilde{A} + \widetilde{A}^T S & S\widetilde{B}_1 & \widetilde{C}_1^T \\ \widetilde{B}_1^T S & -\gamma I & \widetilde{D}_{11}^T \\ \widetilde{C}_1 & \widetilde{D}_{11} & -\gamma I \end{pmatrix} \begin{pmatrix} W & 0 \\ 0 & I \end{pmatrix} < 0 \tag{11}$$

$$\begin{pmatrix} R & I \\ I & S \end{pmatrix} > 0$$

Here $\hat{C}_1$ *and* $\hat{D}_{11}$ *denote the "* q_v *" first rows of* $\widetilde{C}_1$ *and* $\widetilde{D}_{11}$ *respectively. The filter L(s) is then computed from the optimal solution* $\gamma < 1$, M_y, M_u *and* (*R,S*)*, see (Henry & Zolghadri, 2005) for the computational procedure.* ■

3.2. Robust fault sensitivity performance

We shall motivate this section by asking the following question: Because the conditions stated by eq (6) and eq. (8) (and therefore eq (9)) are only sufficient conditions, what is the degree of conservatism of the obtained solution (M_y, M_u,,$L(s)$) ? The FDI filter design method described in the previous section does not account for the structure of the model perturbation block Δ. This means that the solution (M_y,M_u,,$L(s)$) can be conservative in some cases. Furthermore, $\gamma > 1$ (see the inequality eq (11)) does not imply with certainty that the FDI filter does not meet the desired H_∞ / H- specifications.

To check if the required performances are achieved, the robust test based on the generalized structured singular value (denoted μg) proposed in (Henry *et.al*., 2001), (Henry *et.al*., 2002), (Henry *et.al*., 2003) can be used.

3.2.1. Definition of the μg function

Robust stability, i.e. stability of all models in the model set $(\Delta(s) * P(s))$, is analyzed with the μ-function. The real-valued function μ is the inverse of the size of the smallest destabilizing perturbation Δ (Doyle et.al., 1982). Consequently, μ-analysis guarantees stability for perturbations up to 1/μ. In a μg-problem, the perturbation structure Δ is divided into two parts, say Δ_J and Δ_K, so that Δ_J satisfies a maximum norm constraint and Δ_K a minimum gain constraint (Henry *et.al*., 2002). The analogous stability result is that the system is stable for $\|\Delta_J\|_\infty < 1/\mu g$ and for $\|\Delta_K\|_\infty > \mu g$.

To formalize, consider a block structure $\underline{\Delta} = \{diag(\Delta_J, \Delta_K)\}$ and a complex valued matrix $N = \begin{pmatrix} N_{JJ} & N_{JK} \\ N_{KJ} & N_{KK} \end{pmatrix}$ partitioned in accordance with $\underline{\Delta} = \{diag(\Delta_J, \Delta_K)\}$ that satisfies the closed-loop equations

$$z = Nv,\ v = \Delta z,\ z = \begin{pmatrix} z_J \\ z_K \end{pmatrix},\ v = \begin{pmatrix} v_J \\ v_K \end{pmatrix} \tag{12}$$

The μg-function is a positive real-valued function of the matrix N and the specified block structure $\underline{\Delta}$ defined according to:

$$\mu g_{\underline{\Delta}}(N) = \max_{\|v\|=1} \left\{ \gamma : \begin{matrix} \|v_j\|\gamma \le \|z_j\|, \forall j \in J \\ \|v_k\| \ge \|z_j\|\gamma, \forall k \in K \end{matrix} \right\} \tag{13}$$

The μg function is defined in a domain dom(μg) given by:

$$N \in dom(\mu g) \quad iff\ N_{KK} v_K 0 \Rightarrow v_K = 0 \tag{14}$$

which is equivalent to a nontrivial solution, i.e. the maximization part in the μg problem is finite.

3.2.2. Robust fault sensitivity performance analysis

Consider the block diagram depicted in figure 2 and the shaping filters W_d and W_f given by eq. (4). Including K, M_y,M_u,,L and the shaping filters W_d and W_f into the model P leads to the set up described by the block diagram shown in figure 4. $\widetilde{d}$ is defined as in figure 3 and $\widetilde{f}$ is a fictitious signal which is defined according to:

$$\widetilde{f}(s) = W_f(s) f(s) \tag{15}$$

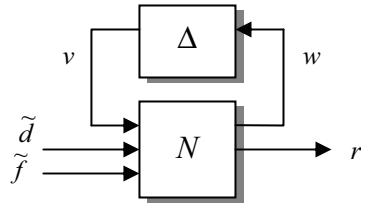


Fig 4 : The generic structure of robust detection performance analysis problem.

The filter performances analysis problem over the plant perturbations $\Delta \in \underline{\Delta}$ is then a min-max gain problem over the specified frequency grid Ω. This problem can be formulated as:

$$\begin{matrix} \sup\limits_{\omega} \overline{\sigma}\left(T_{r\tilde{d}}(j\omega)\right) < 1 \\ \inf\limits_{\omega\in\Omega} \underline{\sigma}\left(T_{r\tilde{f}}(j\omega)\right) > 1 \end{matrix} \quad \forall \Delta \in \underline{\Delta} : \|\Delta\|_{\infty} \le 1 \tag{16}$$

where $T_{r\tilde{d}}$ and $T_{r\tilde{f}}$ denote respectively the closed loop transfer between r and $\tilde{d}$ and r and $\tilde{f}$ .The following theorem gives the solution of the robust fault sensitivity analysis problem.

Theroem (Henry,2005): Consider the model structure depicted in figure 4 and partition N according to $N = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}$, *where* N_{22} *denotes the transfer between the signals r* and $\tilde{f}$. *Let* $\sup\limits_{\omega} \mu_{\overline{\underline{\Delta}}}\left(N_{11}(j\omega)\right) < 1$ *where* $\overline{\underline{\Delta}} = \{diag(\Delta, \Delta_d)\}$ *where* $\Delta_d \in C^{\dim(\tilde{d})\times\dim(r)}$ *is a fictitious plant perturbation block introduced to closed the loop between r* and $\tilde{d}$, *and let* $N \in dom(\mu g)$. *Then a necessary and sufficient condition for eq. (16) to hold is:*

$$\sup_{\omega\in\Omega} \mu g_{\tilde{\underline{\Delta}}}\left(N(j\omega)\right) < 1 \tag{17}$$

The block structure $\tilde{\underline{\Delta}}$ *is defined according to* $\tilde{\underline{\Delta}} = \left\{diag\left(\overline{\Delta}, \Delta_f\right)\right\}$ *where* $\Delta_f \in C^{\dim(\tilde{f})\times\dim(r)}$ *is a fictitious uncertainty block introduced to close the loop between r and* $\tilde{f}$. *The condition* $\sup\limits_{\omega} \mu_{\overline{\underline{\Delta}}}\left(N_{11}(j\omega)\right) < 1$ *is equivalent to the maximum norm constraint in eq. (16) is satisfied over the block structure* $\underline{\Delta}$ *which is equivalent to the robustness performance specification (S.1) yields* $\forall \Delta \in \underline{\Delta} : \|\Delta\|_{\infty} \le 1$ ■

Because this theorem involves a necessary and sufficient condition which takes into account the structure of the model perturbations Δ, the robust sensitivity performance (i.e. the specification (S.2)) can be tested by calculating the μg function of N over the block structure $\tilde{\underline{\Delta}}$. Unfortunately, the exact calculation of μg is not currently available. As a result, it is necessary to develop computable bounds. Computationally inexpensive upper and lower bounds have been developed in (Morris, 1996). If the bounds are equal, then an exact value of μg has been found. An upper bound of μg can be formulated as a convex optimization problem, which results in checking a LMI feasibility. In (Newlin & Smith, 1998) it is shown that for three or fewer complex blocks in the $\tilde{\underline{\Delta}}$ structure, the proposed LMI-based algorithm gives the exact solution. For more than three blocks, the authors state that it might be reasonably expected that the LMI upper bound is as accurate as the LMI upper bound for μ. In the standard μ case with more than three complex blocks, the gap between the upper bound and μ is typically only a few percent and is very rarely more than ten percent. A lower bound algorithm from the "Power Algorithm" family is also proposed in (Morris, 1996), which seeks to optimize Δ_J and Δ_K explicitly. The algorithm is developed in a similar fashion to the power algorithm for μ. The authors have studied the convergence property of the algorithm and it appears that the lower bound generally converges when the $\tilde{\underline{\Delta}}$ structure is restricted to complex full blocks. For structures involving more than two real blocks, the lower bound algorithm fails to converge.

An important point regarding the robust fault sensitivity test given by eq. (17) is that the convergence of the upper bound is much more critical than the lower, as we are checking if μg (or any upper bound) is below 1 or not.

3.3. Discussion on fault isolation

After a judgment "fault", fault isolation is required if one desires to gain deeper insight into the faulty situation. The aim is to provide the system operator with the fault location (i.e. sensors and/or actuators and/or components faults). One approach to fulfil the fault isolation task is to design a set of structured residuals. Each residual is designed to be sensitive to a subset of faults, whilst remaining robust to the remaining faults (which are treated like the disturbances *d*). The residual set which has required sensitivity to specific faults and insensitivity to other faults is known as the *structured residual set* (Chen & Patton, 1999). The design procedure consists of two steps. The first step is to specify the sensitivity and insensitivity relationships between residuals and faults according to the assigned isolation task, and the second is to design a set of FDI filters according to the desired sensitivity and insensitivity relationships. The advantage of this approach is that the diagnostic analysis is simplified to determining which of the residuals are affected by the faults (e.g. non-zero). The decision test may then be performed separately for each residual, yielding a Boolean decision table, and the isolation task can be fulfilled using this table. This is called as a *dedicated residual set* which is inspired by the DOS (Dedicated Observer Scheme) approach (Chen & Patton, 1999). From a practical point of view, even when this structured residual set can be designed, there is no design freedom left to achieve other desirable performances. This motivates the so called *generalized residual set* which is inspired by the GOS (Generalized Observer Scheme) approach. The method consists in designing the residual set to make each residual sensitive to all but one fault. The isolation task can again be performed using a Boolean decision table. Finally, an alternative way of achieving the isolability of faults is to design a *directional residual vector* which lies in a fixed and fault-specified direction (or subspace) in the residual space, in response to a particular fault. In this case, the fault isolation problem is one of determining which of the known fault signature directions the generated residual vector lies the closest to. Of course, to isolate faults reliably, each fault signature has to be uniquely related to one fault.

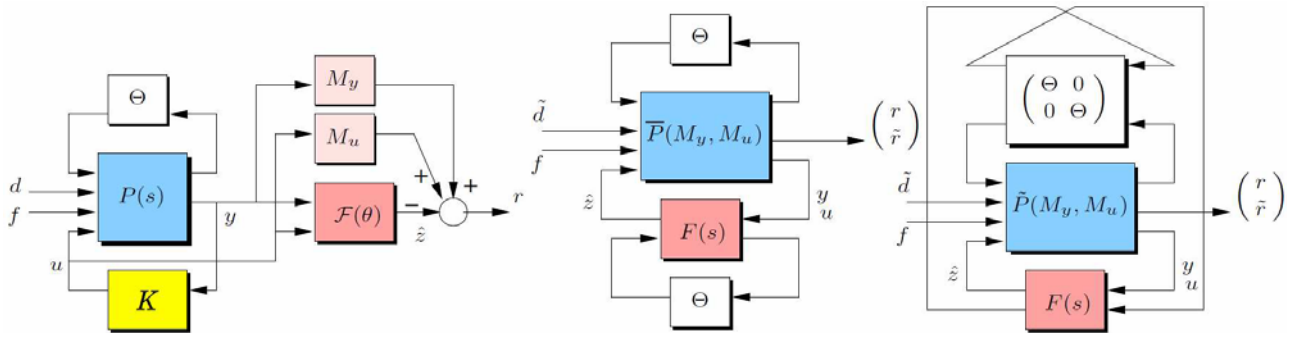


Fig. 5: (a) FDI filter design problem, (b) LPV filter structure, (c) "Quasi-standard" structure

3.4. *Extension to LPV cases*

In recent years, linear parameter varying (LPV) methods have gained a great deal of interest, both in the robust control and diagnostic communities. LPV theory is advantageous because

- it offers an efficient paradigm to model nonlinear systems with on-line measurable state depending parameters,
- it provides stability and performance guarantee over wide range of changing parameters.

The two commonly used approaches for designing fault detection and isolation schemes for LPV systems are *i*) the fault estimation approach and *ii*) the residual generation approach. Similarly to the LTI case, the fault estimation technique is convenient since the design problem can be formulated as a pure LPV H_∞ filtering problem where the error between the fault estimates and the faults is minimized for all varying parameters. Residual generation is fundamentally different from fault estimation because it can not be formulated in a minimization problem. In a residual generation problem the residuals have to be simultaneously robust to unknown inputs and sensitive to faults.

The purpose of the following is to describe a method for designing a structured residual generator for fault detection and isolation for LPV systems modelled in a LFR manner.

3.4.1. *Problem formulation*

Consider the following LPV model in the LFR form, placed in a feedback control loop (see figure 5.a for easy reference)

$$y = F_u(P,\Theta)\begin{pmatrix} d \\ f \\ u \end{pmatrix}, u = Ky \tag{18}$$

d denotes the exogenous disturbances (including measurement noise) and f models the faults to be detected. K is a LPV or LTI controller that is assumed to be known. P denotes a known LTI model and Θ is a block diagonal time-varying operator specifying how θ enters P, so that

$$\Theta = blockdiag(\theta_1 I_{k_1}, \cdots, \theta_q I_{k_q}) \tag{19}$$

where $k_i > 1$ whenever the parameter θ_i is repeated. It is assumed that all parameters $\theta_i(t)$ are measured in real time and bounded so that, without loss of generality, $|\theta_i(t)| \le 1, \forall t \Rightarrow \|\Theta\|_\infty \le 1$.

The FDI design problem we are interested in is formulated as follows:

Let f entering in eq. (20) be detectable faults (the interesting reader can refer to (Saberi et.al., 2000) for a discussion on fault detectability) and consider the residual vector r defined by figure 5.a. The goal is to derive simultaneously M_y , M_u and the state space matrices of the LTI filter $F: \mathcal{F}(\Theta)=F_l(F,\Theta)$ defined according to

$$F(s) = \begin{pmatrix} C_{F1} \\ C_{F\theta} \end{pmatrix}(sI - A_F)^{-1}\begin{pmatrix} B_{F1} & B_{F\theta} \end{pmatrix} + \begin{pmatrix} D_{F11} & D_{F1\theta} \\ D_{F\theta 1} & D_{F\theta\theta} \end{pmatrix} \tag{20}$$

where $\mathcal{F}(\Theta)$ is internally stable for all parameter trajectories $\theta(t)$, that solve the following optimisation problem (see figure5.b for an illustration):

$$\min_{M_y, M_y, F} \gamma_1 \quad s.t. \|T_{d\to r}(\theta)\|_\infty < \gamma_1 \qquad \text{and} \qquad \max_{M_y, M_y, F} \gamma_2 \quad s.t. \|T_{f\to r}(\theta)\|_- > \gamma_2 \tag{21}$$

$T_{d\to r}(\theta)$ and $T_{f\to r}(\theta)$ denote the LPV transfers between d and r, and f and r respectively. γ_1 and γ_2 are two positive constants. θ playing the role of a scheduling variable, $\mathcal{F}(\Theta)=F_l(F,\Theta)$ gives the rule for updating the FDI filter state-space matrices (20) based on the measurements of θ.

In this formulation, $\|T_{d\to r}(\theta)\|_\infty$ and $\|T_{f\to r}(\theta)\|_-$ denote respectively the H_∞ and the H- norms for LPV systems, see the notation section. These norms are used to specify the robustness objectives and the fault sensitivity requirements, respectively. The performance index γ_1 guarantees a minimum nuisances attenuation H_∞ gain, whereas the performance index γ_2 guarantees a maximum faults amplification H- gain.
This problem could also be interpreted as a multiobjective optimisation problem whereby the choice of γ_1 and γ_2 is guided by the Pareto optimal points. However, in practice,

γ_1 and γ_2 are better considered as parameters to be selected by the designer since finding "optimal" values is highly related to the system under consideration.

Remark : It should be outlined that a great advantage of the considered formulation is that the controller K can either be a LTI controller, or a LPV controller. This allows to consider, e.g. gain-scheduling LTI controllers, nonlinear dynamic inversion based control laws, see for instance (Papageorgiou & Glover, 2005). However, in some cases, it may be difficult to assess to a such model. To overcome this problem, the solution consists in considering eq. (18) in open loop, i.e. the controller K is removed and the controlled signal u is considered in the same manner than the exogenous signals d. Note that in this case, $F_u(P,\Theta)$ must be internally stable for all parameter trajectories $\theta(t)$.

Remark : In the above problem definition, it is assumed that the residual structuring matrices M_y, M_u do not depend on θ. This assumption will be justified later, see section 3.4.3.

3.4.2. Solution to the problem

■ **The standard setup**

To solve the problem, we proceed very similarly (thanks to the LFR paradigm and the definition of the H- norm) to the LTI case. The design objectives are formulated in terms of desired gain responses for the appropriate closed-loop transfers.

To proceed let $W_d : \|W_d\|_\infty \le \gamma_1$ and $W_f : \|W_f\|_- \le \gamma_2$ be the shaping filters associated to $T_{d\to r}(\theta)$ and $T_{f\to r}(\theta)$ respectively. It is assumed that W_d is invertible (this can be done without loss of generality since it is always possible to add zeros in W_d to make it invertible). Then, there exists a solution to the H_∞ specification in eq. (21) iff:

$$\exists M_y, M_u, F : \left\| T_{\tilde{d}\to r}(\theta) \right\|_\infty < 1 \qquad (22)$$

where $\tilde{d}$ is a fictitious signal generating d through W_d^{-1}.

The following lemma allows the H- constraint in eq. (21) to be formulated in terms of a fictitious H_∞ one.

Lemma (Henry et al, 2009): Let W_f be an invertible LTI transfer matrix defined such that $\|W_f\|_- \le \gamma_2 / \lambda \|W_F\|_-$ \$|| and $\|W_F\|_- > \lambda$ where $\lambda = 1 + \gamma_2$. Define the (fictitious) signal $\tilde{r}$ such that $\tilde{r} = r - W_F f$. Then a sufficient condition for the H- constraint in eq. (21) to hold is

$$\exists M_y, M_u, F : \left\| T_{f\to\tilde{r}}(\theta) \right\|_\infty < 1 \qquad (23)$$

■

Following the above developments, the design problem can be re-casted in a framework which looks like a standard H_∞ problem for LPV systems, by combining both requirements eq. (22) and eq. (23) into a single constraint: A sufficient condition for M_y,M_u and F to solve the problem is

$$\left\| T_{(\tilde{d}^T \quad f^T)^T \to (r^T \quad \tilde{r}^T)^T}(\theta) \right\|_\infty < 1 \qquad (24)$$

where $T_{(\tilde{d}^T \quad f^T)^T \to (r^T \quad \tilde{r}^T)^T}(\theta)$ denotes the transfer between $(\tilde{d}^T \quad f^T)^T$ and $(r^T \quad \tilde{r}^T)^T$.

This equation shows that the original LPV problem can be viewed as a gain-scheduling H_∞ performance problem, where the time varying parameters θ enter both the plant and the filter F, see figure 5.b.

Now, let us introduce the following augmented plant

$$\tilde{P}(M_y, M_u) = \begin{pmatrix} 0 & 0 & I \\ 0 & \bar{P}(M_y, M_u) & 0 \\ I & 0 & 0 \end{pmatrix} \qquad (25)$$

where $\bar{P}(M_y, M_u)$ is a LTI transfer that is deduced from P, K, W_d, W_F, M_y and M_u using some linear algebra manipulations (the mathematical details about these manipulations are omitted here for clarity). Then, it can be verified that the closed-loop mapping from exogenous inputs $(\tilde{d}^T \quad f^T)^T$ to output signals $(r^T \quad \tilde{r}^T)^T$ can be expressed according to:

$$\begin{pmatrix} r \\ \tilde{r} \end{pmatrix} = F_u\left(F_l\left(\tilde{P}(M_y, M_u), F\right), \begin{pmatrix} \Theta & 0 \\ 0 & \Theta \end{pmatrix} \right) \begin{pmatrix} \tilde{d} \\ f \end{pmatrix} \qquad (26)$$

This expression amounts to redrawing the diagram illustrated in figure 5.b as in figure 5.c. It follows with eq. (24) that a sufficient condition for M_y, M_u and F to solve problem 1 is:

$$\left\| F_u\left(F_l\left(\tilde{P}(M_y, M_u), F\right), \begin{pmatrix} \Theta & 0 \\ 0 & \Theta \end{pmatrix} \right) \right\|_\infty < 1 \qquad (27)$$

Thus, the FDI filter design problem can be viewed as a standard LPV H_∞ performance problem for the LTI plant $\tilde{P}(M_y, M_u)$ in the face of the norm-bounded block-repeated uncertainty $\begin{pmatrix} \Theta & 0 \\ 0 & \Theta \end{pmatrix}$. In fact, this is not the case since the transfer $\tilde{P}(M_y, M_u)$ depends on M_y and M_u that are unknown. However, sufficient conditions for solvability can be provided by means of the small gain theory using adequate commutable scaling matrices. This will be done in the next paragraph.

■ **The SDP formulation**

Let the state-space realization of $\bar{P}(M_y, M_u)$ be given according to

$$\overline{P}(M_y, M_u) = \begin{pmatrix} C_\theta \\ C_1 \\ C_2 \end{pmatrix} (sI - A)^{-1} (B_\theta \quad B_1 \quad B_2) + \begin{pmatrix} D_{\theta\theta} & D_{\theta 1} & D_{\theta 2} \\ D_{1\theta} & D_{11} & D_{12} \\ D_{2\theta} & D_{21} & D_{22} \end{pmatrix} \tag{28}$$

where the partitioning is conformable to the setup illustrated in figure 5.b. The problem dimensions are as follows

$$A \in \Re^{n\times n}, D_{\theta\theta} \in \Re^{q\times q}, D_{11} \in \Re^{p_1\times m_1}, D_{22} \in \Re^{p_2\times m_2}, A_F \in \Re^{n_F\times n_F} \tag{29}$$

It can be verified that, by construction:

- B_2, $D_{\theta 2}$, D_{22} are null matrices. This shows that the FDI filter does not affect neither the state, nor the measurements of the system, i.e. the FDI filter operates (obviously) in open loop with regards to the system;
- C_1, $D_{1\theta}$, D_{11} depend on M_y, M_u that are the "static" part of the solution we are looking for;
- $m_1 = \dim(\tilde{d}) + \dim(f)$, $p_1 = 2\dim(r)$, $m_2 = \dim(r)$ and $p_2 = \dim(y) + \dim(u)$.

Furthermore, we assume that D_{11} is a square matrix. This can be always fulfilled by augmenting the problem with columns/rows of zeros.

Now, let Δ denote the structure set associated with Θ defined by eq. (19) and L_Δ be the set of commutable scaling matrices defined so that

$$L_\Delta = \{L > 0 : L\Theta = \Theta L, \forall \Theta \in \Delta\} \subset \Re^{q\times q} \tag{30}$$

The following theorem allows to formulate the design problem in terms of a SDP optimisation one. The proof can be found in (Henry et.al., 2009):

Theorem (Henry et.al., 2009): Consider Δ, L_Δ and the state-space realization of $\overline{P}(M_y, M_u)$ defined above. Let $W = (C_2 \quad D_{2\theta} \quad D_{21})^\perp$ and consider any matrix $X \in \Re^{m_2\times m_2}$. The H_∞ requirement eq. (27) is satisfied and $\mathcal{F}(\theta)$ is of full-order and internally stable for all parameter trajectories $\theta(t)$, if there exist $\gamma < 1$ $M = (M_y, M_u) \in \Re^{m_2\times p_2}$ and pairs of symmetric positive definite matrices $(R, S) \in \Re^{n\times n}$ and $(L_3, J_3) \in L_\Delta$ solving the following SDP problem:

min γ *subject to:*

$$\begin{pmatrix} AR + RA^T & RC_\theta^T & RC_1^T H^T & B_\theta J_3 & B_1 \\ C_\theta R & -J_3 & 0 & D_{\theta\theta} J_3 & D_{\theta 1} \\ HC_1 R & 0 & -\gamma \mathrm{diag}(2X^T X, I_j) & HD_{1\theta} J_3 & HD_{11} \\ J_3 B_\theta^T & J_3 D_{\theta\theta}^T & J_3 D_{1\theta}^T H^T & -J_3 & 0 \\ B_1^T & D_{\theta 1}^T & D_{11}^T H^T & 0 & -\gamma I \end{pmatrix} < 0 \tag{31}$$

$$\left(\begin{array}{c|c} W & 0 \\ \hline 0 & I \end{array}\right)^T \left(\begin{array}{ccc|cc} A^T S + SA & SB_\theta & SB_1 & C_\theta^T L_3 & C_1^T \\ B_\theta^T S & -L_3 & 0 & D_{\theta\theta}^T L_3 & D_{1\theta}^T \\ B_1^T S & 0 & -\gamma I & D_{\theta 1}^T L_3 & D_{11}^T \\ \hline L_3 C_\theta & L_3 D_{\theta\theta} & L_3 D_{\theta 1} & -L_3 & 0 \\ C_1 & D_{1\theta} & D_{11} & 0 & -\gamma I \end{array}\right) \left(\begin{array}{c|c} W & 0 \\ \hline 0 & I \end{array}\right) < 0 \tag{32}$$

$$\begin{pmatrix} R & I \\ I & S \end{pmatrix} \geq 0, \quad \begin{pmatrix} L_3 & I \\ I & J_3 \end{pmatrix} \geq 0 \tag{33}$$

$$H = diag\left(\begin{bmatrix} X^T & -X^T \end{bmatrix}, I_j\right), \quad n_F = n \tag{34}$$

where "j" denotes the number of added rows to make D_{11} square. ■

Remark: Coming back to the statement of the above theorem, it can be noted that X is a user-defined matrix. It is shown in (Henry et al, 2009) that the choice of X is only guided by numerical aspects.

3.4.3. The case of parameter-dependent residual structuring matrices

Consider now the case of parameter-dependent residual structuring matrices defined so that:

$$\begin{aligned} r(s) &= M_y(\theta) y(s) + M_u(\theta) u(s) - z(s) \\ \hat{z}(s) &= F(s, \theta) \begin{pmatrix} y(s) \\ u(s) \end{pmatrix} \end{aligned} \tag{35}$$

Our aim is to derive simultaneously $M_y(\theta)$, $M_u(\theta)$ and the state space matrices of the LTI filter F defined by (20), that solve the optimisation problem defined by equations (21). A particularity of this formulation is that the time varying parameters enter now both the structuring matrices and the filter $\mathcal{F}(\Theta)=F_l(F,\Theta)$.

To apprehend this problem with the small gain theory, we must first gather all parameter-dependent components into a single uncertainty block. To proceed, let $\mathbf{M}(\theta)=(M_y(\theta)\ M_u(\theta))$ be put into a LFR-form so that:

$$\mathbf{M}(\theta)=F_l(M,\Theta): M = \begin{pmatrix} M_{11} & M_{1\theta} \\ M_{\theta 1} & M_{\theta\theta} \end{pmatrix} \tag{36}$$

θ playing the role of a scheduling variable, this equation gives the rule for updating $M_y(\theta)$ and $M_u(\theta)$ based on the measurements of θ.

Following the same developments than those presented in section 3.4.2, it can be verified that a sufficient condition for M and F to solve the FDI filter design problem is

$$\left\| F_u\left(F_l\left(\tilde{P}(M), F\right), \begin{pmatrix} \Theta & & \\ & \Theta & \\ & & \Theta \end{pmatrix} \right) \right\|_\infty < 1 \tag{37}$$

This equation shows that the original problem can be viewed as a gain-scheduling H_∞ performance problem in the face of the norm-bounded block-repeated uncertainty $\begin{pmatrix} \Theta & & \\ & \Theta & \\ & & \Theta \end{pmatrix}$. Clearly, the dimension of this new set is bigger than those defined for constant residual structuring matrices and, by virtue of the small gain theory, this may drive to more conservative solutions than those presented in section 3.4.2. Thus, it is preferred constant residual structuring matrices at this stage.

4. EXAMPLES OF APPLICATIONS

4.1.1. The Microscope satellite (LTI context)

Microscope is a spine satellite due to be launched on a circular, quasi-polar, sun-synchronous orbit at an altitude of *700*km with ascending and descending nodes at 6:00 and 18:00, respectively, see figure 6 for an illustration performed by the CNES-France.

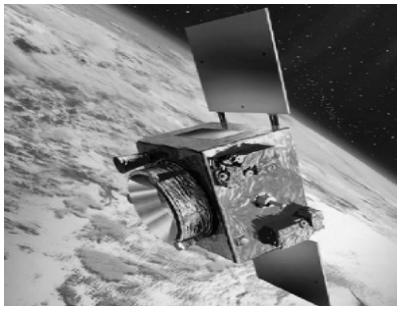


Figure 6: The Microscope satellite

To control its trajectory, Microscope uses the coupling of six ultra-sensitive accelerometer sensors, a stellar sensor and a very precise electric propulsion system composed by twelve Field Emission Electric Propulsion (FEEP) thrusters. The mission can be in danger if a FEEP thrusters fault occurs, since the satellite may not compensate for non-gravitational disturbances which are indispensable prior conditions for its mission: testing the Equivalence Principle.

To overcome this problem, a FDI scheme that consists of a bank of 12 H_∞/H- residual generators is proposed in (Henry, 2008). The design is done so that the sensitivity level of the *ith* residual with respect to the *ith* FEEP thruster fault f_i is maximised in the H- norm sense, whilst guaranteeing robustness against measurement noises and spatial disturbances in the H_∞ norm sense.
Figure 7 illustrates the behaviour of the residuals $r_i(t)$, i=1…12, the behaviour of the decision test and the isolation criteria, for some faulty situations. The (nonlinear) simulations were done using the Microscope simulator provided by the CNES-France. As can be seen in the figures, after a small transient behaviour, all faults are successfully detected and isolated by the FDI unit.

4.1.2 The HL-20 Re-entry Launched Vehicle (LTI context)

The HL-20 Re-entry Launched Vehicle (RLV) (see figure 8) was defined as a component of the Personnel Launch System (PLS) mission. This has initially been designed to support several manned-space missions including the orbital rescue of astronauts, the International Space Station (ISS) crew exchange and some satellite repair missions.

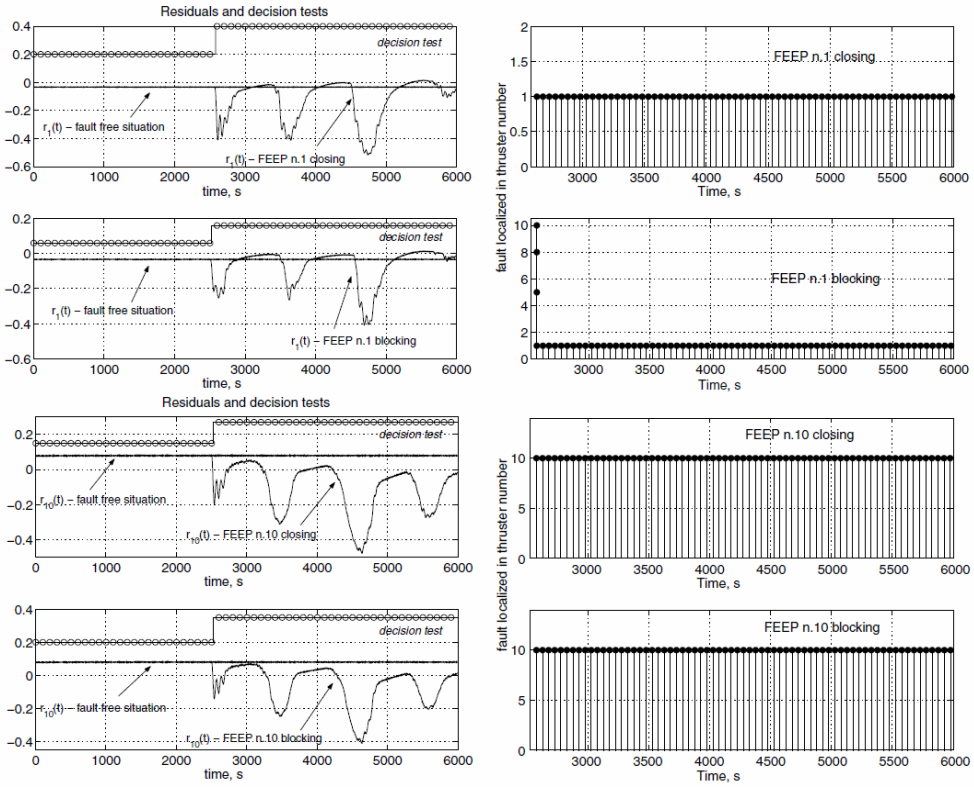


Figure 7: Fault-free and faulty residuals with the decision test (left) and the isolation criteria (right)

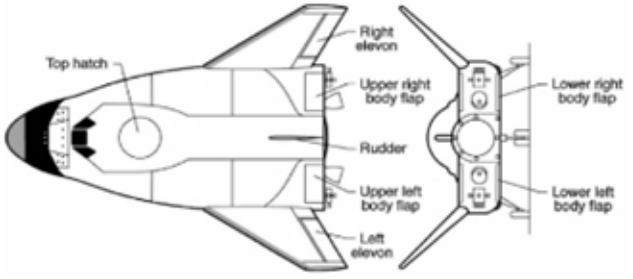


Figure 8: The HL-20 vehicle

A typical atmospheric re-entry for a medium or high L/D vehicle consists of performing three successive flight phases, namely the Hypersonic phase from about 120 km high down to TAEM (Terminal Area Energy Management) handover, the TAEM phase from Mach 2 gate down to Mach 0.5 gate and the auto-landing phase from Mach 0.5 gate down to the wheel stop on the runway. After having achieved the hypersonic path, the vehicle initiates the TAEM phase characterized by an entry point called TEP (Terminal Exit Point), typically defined when crossing Mach 2 gate, and an exit point called NEP (Nominal Exit Point) which is defined in terms of altitude, velocity and distance to the runway. Finally, the landing path is defined in terms of desired altitude from the runaway threshold and is composed of three successive sections, i.e. a steep outer glideslope, a parabolic pullup manoeuvre and a shallow inner glideslope.

The work presented in (Falcoz et.al., 2007, 2008) focuses on any type of faults in the wing flap actuators during the landing phase. The strategy proposed by the authors consists of a bank of two H_∞/H- fault detection filters that are designed so that a given filter is made robust against measurement noise, winds turbulence, guidance reference signals and faults in a given wing flap actuator, whilst remaining sensitive to all faults in the other wing flap actuator. For the purpose of estimating the position of the faulty control surfaces, the nonlinear EKF method presented in (Falcoz et.al., 2007, 2008) is used.
Figure 9 illustrates the results for some nonlinear simulations coming from a medium fidelity Matlab simulator in the presence of wind and atmospheric turbulences. A monte carlo campaign reveals that the faults are successfully detected, isolated and estimated by the FDI unit.

4.1.3 Academic example (LPV context)

To illustrate the potential of the proposed LPV approach, an illustrative example of academic nature is considered. Consider the following system

$$G(\theta):\begin{cases}\dot{x} = A_o(\delta)x + B_o u + K_1 f \\ y = C_o x + n\end{cases} \tag{38}$$

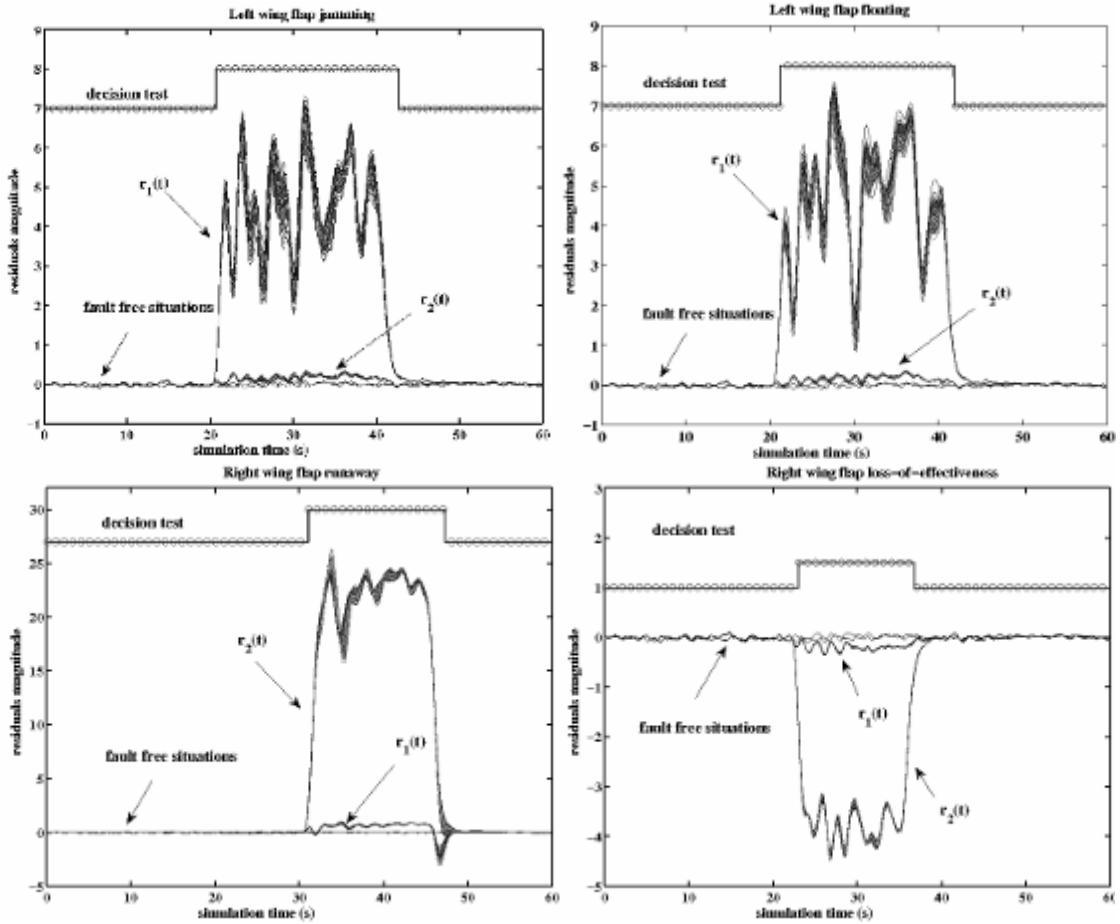


Figure 9: A monte carlo campaign for fault detection and isolation in left and right wing flaps of the HL20 RLV

$$A_o(\delta)=\begin{pmatrix}0 & \delta_2\\ -0.1\delta_1 & -\delta_3\end{pmatrix},\ B_o=\begin{pmatrix}0\\ 0.1\end{pmatrix},\ K_1=\begin{pmatrix}1\\ 1\end{pmatrix},\ C_o=I_2 \quad (39)$$

$\delta_i(t), i=1,2,3$ are assumed to vary in the following bounds

$$5\le\delta_1(t)\le 8,\quad -2\le\delta_2(t)\le -1,\quad 2\le\delta_3(t)\le 4$$

with arbitrary time variations. It is assumed that the system operates in a feedback control loop so that the closed loop is internally stable for all parameter trajectories $\theta(t)$ (a LQ-based controller was computed for this purpose).

Following the developments presented in sections 3.4, the system is put into a LFR form according to the setup illustrated in figure 5.a. This boils down to the block diagonal time-varying operator Θ defined according to $\Theta=diag(\theta_1,\theta_2,\theta_3)$ which has been normalized so that $|\theta_i|\le 1, i=1,2,3$. The method presented in sections 3.4 is then used to derive M_y, M_u, $\mathcal{F}(\theta)$ and the residual $r(t)$ is computed. The shaping filters W_d and W_f that allow to specify the robustness and the sensitivity objectives, have been fixed according to:

$$W_d=10\frac{1+10^{-2}s}{1+10s}I_2,\quad W_f=0.1\frac{1}{1+2s} \quad (40)$$

By this choice, it is required an attenuation factor of, at least, 40*dB* of $n(t)$ on the residual in the frequency range [100,+∞[*rd/s*, and an amplification factor of, at least, -20*dB* of $f(t)$ on $r(t)$ in low frequencies.

To analyse the computed solution, the principal gains $\bar{\sigma}(T_{d\to r}(j\omega))$ and $\underline{\sigma}(T_{f\to r}(j\omega))$ are plotted versus the required objectives $|W_d(j\omega)|$ and $|W_f(j\omega)|$ for some arbitrary fixed values of $\theta_i(t), i=1,2,3$, see figure 10. Despite these plots only offers necessary conditions since the time-varying aspect of $\theta_i(t), i=1,2,3$ is not considered, it can be argued that the required objectives are met since $\bar{\sigma}(T_{d\to r}(j\omega))$ < $|W_d(j\omega)|$ and $\underline{\sigma}(T_{f\to r}(j\omega))$ > $|W_f(j\omega)|, \forall\omega\in\Omega$ for all fixed $\theta_i(t), i=1,2,3$. Note that, by virtue of the theorem in section 3.4.2, we know that it still yields for all $\theta_i(t), i=1,2,3$, since the optimal value of γ is found to be ≈ 0.98.

The behaviour of $r(t)$ is illustrated in figure 11. Chirp signals 0-1KHz have been considered for the simulation of $\delta_i(t), i=1,2,3$. For the purpose of the fault simulation, a step of magnitude "1" is considered between $50s\le t\le 100s$. As it can be seen, the design objectives are met, leading to the detection of the fault despite the variations of $\delta_i(t), i=1,2,3$.

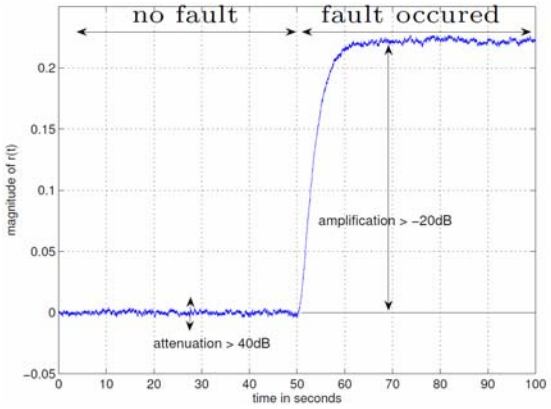


Figure 11. Behaviour of the residuals

5. CONCLUDING REMARKS

To summarize, the technique investigated in this paper can be seen as a nice and practically relevant framework in which various design goals and trades-off are formulated and managed. The optimization problem is then solved by numerically powerful LMI-based techniques. The output of the design is a filter for Fault Detection, or a bank of filters for Fault Detection and Isolation. The approach has been developed by the authors at IMS-LAPS, Bordeaux. The developed techniques have been successfully applied to a number of applications, see http://extranet.ims-bordeaux.fr/aria.

A last section is devoted to FDD based on LPV models in order to take into account wider and more rapid parameters variations. LPV models can be used efficiently to represent some nonlinear systems. Robustness against exogenous disturbances and sensitivity against faults are considered in a framework similar to the H∞/H− setting for LTI systems

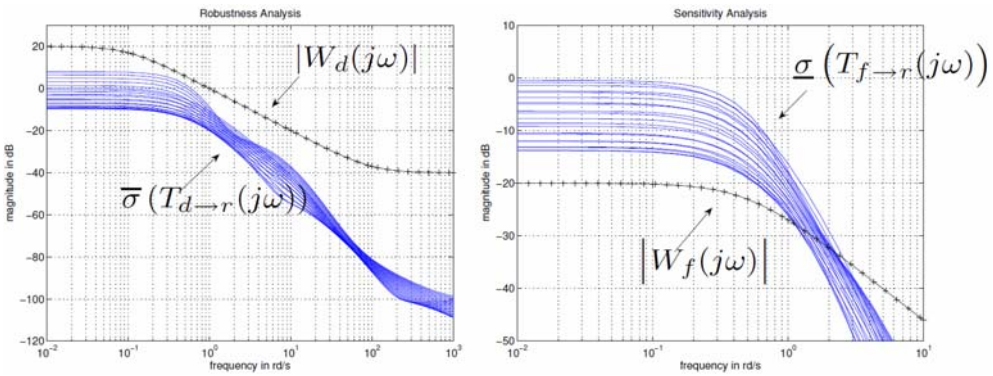


Figure 10. Performance vs objectives / LPV FDI scheme

as described previously. The main difference between this problem and the standard H∞ problem for LPV systems is that it involves the residual structuring matrices that are a priori unknown.

REFERENCES

Apkarian, P., P. Gahinet and G. Becker 1995. "Self-scheduled h1 control in linear parameter-varying systems: a design example". Automatica 31(9), 1251–1261.

Boyd, S., El.Ghaoui, L., Feron, E. and Balakrishnan, V. 1994. "Linear Matrix Inequalities in System and Control Theory". Studies in Applied Mathematics.

Chen, J. and Patton, R.J. 1999. "Robust model-based fault diagnosis for dynamic systems". Kluwer Academic Publishers.

Ding, X. and Guo, L. 1996. "Observer based optimal fault detector". In: Proceedings of the 13th IFAC World Congress. IFAC, San Francisco - USA.

Ding X. & Guo L., 1996. "Observer based optimal fault detector" 13th IFAC World Congress, San Francisco, 187-192

Ding, S., Jeinsch, T., Frank, P. and Ding, E. 2000. „A unified approach to the optimization of fault detection systems. International Journal of Adaptive Control and Signal Processing". Vol 14, pp.725-745.

Falcoz, A., D. Henry and A. Zolghadri 2007. "Development of a robust model-based fault diagnosis technique for re-entry launch vehicles: A case study". In: Progress report.

Falcoz, A., Henry, D., Zolghadri, A., Bornschleg, E. and Ganet, M. 2008. "On-board model-based robust FDIR strategy for reusable launch vehicles (RLV)". In: 7th International ESA Conference on Guidance, Navigation and Control Systems, County Kerry, Ireland.

Falcoz, A., Henry, D., Zolghadri, A. 2008. "A nonlinear fault identification scheme for reusable launch vehicles control surfaces". International Review of Aerospace Engineering", vol. October.

Frank P.M., Alcorta-Garcia E. et Köppen-Seliger 2001, "Modelling for fault detection and isolation versus modelling for control". *Mathematical and Computer Modelling of Dynamical Systems*, vol. 7, no. 1, pp. 1-46.

Gahinet, P. and Apkarian, P. 1994. "A linear matrix inequality approach to Hinf control". Int. Journal Robust Nonlinear Control 4, pp. 421-428.

Grenaille, S., D. Henry and A. Zolghadri 2008. "A method for designing fault diagnosis filters for lpv polytopic systems". Journal of Control Science and Engineering.

Henry, D., Zolghadri, A., Castang, F. and Monsion, M. "A new multi-objective filter design for guaranteed robust FDI performance". In: Proceedings of the 40th Conference on Decision and Control. IEEE, Orlando, USA. pp.173-178.

Henry, D., Zolghadri, A., Castang, F. and Monsion, M. 2002. "A multi-objective filtering approach for fault diagnosis with guaranteed sensitivity performance". In: Proceedings of the 15th IFAC World Congress. IFAC, Barcelona, Spain.

Henry, D. and Zolgahdri, A. 2003. "Hinf/H- filters for fault diagnosis in systems under feedback control". In: Proceedings of SAFEPROCESS'2003,Washington DC, USA, pp.87-92.

Henry, D. and Zolghadri, A. 2005. "Design and analysis of robust residual generators for systems under feedback control". Automatica, Vol.41, pp.251-264.

Henry, D. and Zolghadri, A. 2005. "Design of Fault Diagnosis Filters: A Multi-Objective Approach". Journal of Franklin Institute. Vol.342, No.~4, pp.421-446.

Henry, D. and Zolghadri, A. 2006. "Norm-based design of robust FDI schemes for uncertain systems under feedback control: Comparison of two approaches". Control Engineering Practice, vol.14, no.9, pp.1081-1097.

Henry, D. 2008. "Fault diagnosis of the Microscope satellite actuators using Hinf/H- filters". AIAA Journal of Guidance, Control, and Dynamics. vol.31, no.3, pp.699-711.

Henry D., Falcoz A., Zolghadri A. 2009. "Structured H∞/H− LPV filters for faultdiagnosis: Some new results". Preprints of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, Barcelona, Spain, June 30 - July 3, 2009, pp. 420-425.

Isermann R., 1997, "Supervision, fault detection and fault diagnosis methods – an introduction". Control. Eng. Practice 5(5), 639-652.

Jaimoukha, I., Li, Z., and Papakos V. 2006. "A matrix factorization solution to the H-/H_inf fault detection problem". Automatica. Vol.42, pp.~1907-1912.

Khosrowjerdi, M., Nikoukhah, R., and Safari-Shad, N. 2004. "A mixed H2/Hinf approach to simultaneous fault detection and control". Automatica, Vol.40, No.2, pp.261-267.

Liu, J., Wang, J., and Yang, G. 2005. "An LMI approach to minimum sensitivity analysis with application to fault detection". Automatica, Vol.41, pp.1995-2004.

Morris, J., 1996. "Experimental control and model validation: A helicopter case study". Ph.D. thesis, California Institute of Technologie.

Newlin, M. and Smith, R. 1998. "A generalization of the structured singular value and its application to model validation". IEEE Transactions on Automatic Control. Vol. 43, pp. 901-907.

Niemann H. & Stoustrup J., 1999. "Gain scheduling using the Youla parametrization". Conference on Decision and Control, New York.

Packard, A., Fan M. et Doyle J. (1988), A power method for the structured singular value, *IEEE CDC,* pp. 2132-2137, December, 1988

Papageorgiou, C. and K. Glover 2005. "Robustness analysis of nonlinear flight controllers". AIAA Journal of Guidance, Control and Dynamics 28(4), 639–648.

Rank, M., and Niemann, H., 1999. "Norm based design of fault detectors." International Journal of Control 72 (9), pp 773-783.

Saberi, A., A.A. Stoorvogel, P. Sannuti and H. Niemann 2000. "Fundamental problems in fault detection and identification". International Journal of Robust and Nonlinear Control 10(14), 1209–1236.

Stoustrup, J., Grimble, M., and Niemann, H. 1997. "Design of integrated systems for the control and detection of actuator/sensor faults". Sensor Review, Vol.17, No.2, pp.138-149.

Zhong, M., Ding, S., Lam, J. and Wang, H. 2003. "An lmi approach to design robust fault detection filter for uncertain lti systems". Automatica 39 (2), pp. 543-550.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Design and Evaluation of Reconfiguration-based Fault Tolerance using the Lattice of System Configurations.*

M. Staroswiecki
SATIE, ENS Cachan, USTL, CNRS, UniverSud
61 avenue du Président Wilson
94235 Cachan Cedex, France

Abstract

This paper addresses reconfiguration-based fault tolerance under actuator faults. For this problem, the set of possible system configurations is a lattice. Based on the concept of bottom-up monotonous property, extensive controls are defined, and their design is characterized. The combination of a small number of extensive controls is shown to define a fault tolerance scheme that mixes the passive and active approaches (the PACT scheme) and recovers from all recoverable faults, while minimizing the reliability overcost.

Keywords : Fault tolerant control, actuator outages, LQ problem

1 Introduction

Fault tolerant control (FTC) aims at guaranteeing stability and performance (at least in a degraded sense) under system component faults. The faults considered in this paper are actuator outages, or any actuator fault under the reconfiguration strategy, i.e. when faulty actuators are switched-off (which is assumed to be possible). Therefore, the set of all system configurations to be considered is a lattice. Whatever the objective to be achieved (stability [6], [8], [21], tracking [7], [22], optimal control [14], [15], [23], [24], robustness and disturbance attenuation [26], [27]), the control law associated with each configuration can be developed off-line and implemented in a bank of control laws. Switching from an impaired control law to the correct one only needs the fault to be detected and isolated, avoiding the on-line fault estimation and control re-design steps [9], [12], [14].

*This work has been carried out in the SIRASAS project (Strategies Innovantes et Robustes pour l'Autonomie des Systemes Aeronautiques et Spatiaux) founded by the FRAE (Fonds pour la Recherche dans l'Aeronautique et l'Espace).

Designing the FT control involves passive or active schemes. In passive fault tolerance (PFT) a single controller that guarantees stability and performance whatever the fault is designed. This is a simple approach, whose feasibility can be proven in specific cases [21], [23], [24], but most often, only sufficient existence conditions are available, e.g. in terms of a matrix inequality solvability problem [6], [7]. Moreover, PFT induces a reliability overcost, hence the interest of reducing its conservativeness [20]. On the contrary, in active fault tolerance (AFT), a control law dedicated to each fault is designed. Stability and performance are therefore guaranteed for each fault, provided it is recoverable, a concept that has motivated relatively few works [16], [25].

Combining the advantages of PFT and AFT is an appealing idea. In a recent work, a sequential combination was used to avoid instability during the Detection/Recovery transient [26]. In the present paper a parallel combination, namely the PACT (PAssive / ACTive) approach, where several controllers (AFT), each of them dedicated to a subset of recoverable faults (PFT), is designed. For the PACT scheme, FT stability and FT quadratic performances were studied in [17] and [18], based on the reliable control (RC) idea from [23] and [24]. This paper extends the results in [17], [18] and proposes a design algorithm to minimize the reliability overcost.

Evaluating the performance of the obtained scheme and of each system component from the FT point of view is important in applications. However, few works deal with this problem: different FT evaluations, based on *ad hoc* concepts, were addressed in [25], [3], [16] while in [2] sensors were classified into useless, useful and essential. In this paper, the lattice of system configurations, enriched with reliability data, is shown to provide a sound basis for the evaluation of FT performance.

It is organized as follows : the system is described in Section 2, and the lattice of system configurations is introduced as the result of considering the reconfiguration strategy. Section 3 presents the design of a PACT bank and its associated decision procedure, while Section 4 addresses the design trade-off associated with the reliability overcost. Fault Tolerance evaluation is developped in Section 5, and an example illustrates the approach in Section 6. Some concluding remarks are finally given.

2 System description

2.1 Nominal and faulty systems

Consider a LTI system equipped with a set of m actuators, $s_0 \triangleq \{\alpha_1, \alpha_2, ..\alpha_m\}$. Its nominal operation is described by the state equations

$$\dot{x} = Ax + B_0 u \tag{1}$$

where $x \in R^n$ is the state, $u \in R^m$ is the control, and A, B_0 are constant matrices of appropriate dimensions.

Remark 1 *Disturbances are not considered in (1) for the sake of simplicity. Introducing disturbances and disturbance related specifications is straightforward, using e.g. the disturbance characterization as in [11].*

We are interested in actuator faults and system reconfiguration. System reconfiguration (SR) is based on a mechanism that switches-off the faulty actuators, whatever the fault [1], therefore, it does not require any fault model to be identified on-line. Let $s_i \subset s_0$, $i = 1, ...2^m - 1$ be a proper subset of s_0, the situation in which actuators in $s_0 \backslash s_i$ are faulty is abbreviated as fault i (remark that single as well as multiple faults are considered). In that case, the system model is

$$\dot{x} = Ax + B_i u \tag{2}$$

where $B_i = B_0 \Sigma_i$, $\Sigma_i = diag\{\sigma_i(k), k = 1, ...m\}$ and $\sigma_i(k)$ is a set membership function, $\sigma_i(k) = 1$ if actuator α_k belongs to s_i, and $\sigma_i(k) = 0$ otherwise.

Remark 2 *Partition the control signals into $\hat{u}(s_i)$ - those associated with the actuators in s_i - and $\check{u}(s_i)$ - those associated with actuators in $s_0 \backslash s_i$. Note that $\check{u}(s_i)$ has no effect on the solutions of (2) since the associated columns of matrix B_i are zero.*

2.2 Performance requirement

Define the set of system configurations $S = \{s_i : s_i \subseteq s_0, i \in I\}$ where $I = \{0, 1, ..2^{m-1}\}$. s_0 is the nominal configuration, and $s_i \subset s_0$ is the one that is used under fault i. Let S_{stab} be the subset of configurations such that (A, B_i) is stabilizable (we obviously assume that S_{stab} contains at least the nominal configuration s_0).

Let $Q = C^T C \geq 0$ and $R_0 = diag\{r_k, k = 1, ..m\} > 0$ be given, such that (C, A) is detectable. For a configuration $s_i \subseteq s_0, s_i \in S_{stab}$, let R_i be obtained from R_0 by zeroing all rows and columns associated with the faulty actuators in $s_0 \backslash s_i$. Let $u = Kx$ be a state feedback that stabilizes the closed-loop matrix $F_i \triangleq A + B_i K$. The performance of configuration s_i under the state feedback K is evaluated by the cost

$$J_i(x_0, K) \triangleq \int_0^\infty \left[x^T Q x + u^T R_i u\right] dt \tag{3}$$

which is well known to be $J_i(x_0, K) = x_0^T W_i x_0$ where $W_i = W_i^T > 0$ satisfies the Lyapunov equation

$$Q + K^T R_i K + W_i F_i + F_i^T W_i = 0 \tag{4}$$

Definition 3 *A state feedback K is **admissible** for configuration $s_i \in S_{stab}$ if*

$$\forall x_0 \in R^n : J_0(x_0, K) \leq x_0^T N x_0 \tag{5}$$

where $N = N^T > 0$ is a given specification.

Remark 4 *From the partition of u into $\hat{u}(s_i)$ and $\check{u}(s_i)$ one gets a partition of the rows of K into $\hat{K}(s_i)$ and $\check{K}(s_i)$. Since $\check{u}(s_i)$ has no influence on the trajectories of the system (A, B_i) it follows that W_i does not depend on the values in $\check{K}(s_i)$.*

Assessing whether a given state feedback K is admissible for a configuration s_i can be done thanks to the following lemma.

Lemma 5 *The state feedback K is admissible for configuration $s_i \subseteq s_0$ if and only if there exists a matrix $W = W^T > 0$ such that*

$$\begin{aligned} W &\leq N \\ Q + K^T R_i K + W^T F_i + F_i^T W &\leq 0 \end{aligned} \tag{6}$$

Proof. *Necessity is evident, because F_i must be Hurwitz and the cost matrix W associated with the stabilizing feedback K satisfies (4) and (5). Sufficiency follows from the fact that under (6) $x^T W x$ is a Lyapunov function for the system $\dot{x} = F_i x$, and the cost satisfies*

$$J_i(x_0, K) \leq x_0^T W x_0 \leq x_0^T N x_0$$

■

Remark 6 *Considering state-feedbacks assumes that the state is measured, or reconstructed from sensors. When reconstructed, a convergence delay obviously exists after the occurrence of a fault (the observer uses the pre-fault model as long as the fault is not detected and isolated, while the system behaves according to the post-fault model). The effect of such mismatches are considered small enough to be neglected (all the more as only fault detection and isolation is necessary in the SR strategy, excluding the need for fault estimation and control accommodation).*

2.3 The set of recoverable configurations

Note that some configurations may be such that whatever the state feedback $K \in R^{m \times n}$ the feasibility problem (6) has no solution, hence the following definition.

Definition 7 *Configuration $s_i \subseteq s_0$ is **recoverable** by a state feedback control (SF-recoverable) if there exists a pair (K, W) such that (6) holds.*

Remark 8 *SF-recoverability is a structural property, i.e. it only depends on the fault (i.e. on the configuration s_i that is analyzed), not of the control law that is used. Let S_{recov} be the subset of recoverable configurations, one obviously has $S_{recov} \subseteq S_{stab}$. A configuration $s_i \in S_{recov}$ may be recovered by some feedback law, and not recovered by another one. On the contrary, there is no control law that can recover a configuration $s_i \notin S_{recov}$. When a non recoverable fault occurs, objective reconfiguration must be considered [1].*

Determining the set S_{recov} is a basic problem, whose solution is given by the following Lemma.

Lemma 9 *Given the specification N, a configuration $s_i \subseteq s_0$ is SF-recoverable if and only if $s_i \in S_{stab}$ and $W_i^* \leq N$ where W_i^* is the unique stabilizing solution of the Riccati equation $A^T W_i^* + W_i^* A - W_i^* \beta_i W_i^* + Q = 0$ and $\beta_i = B_i R_0^{-1} B_i^T$.*
Proof. *The stabilizability of s_i is obviously necessary. The system (A, B_i) controlled by $u = Kx$ has exactly the same response as the system $\left(A, \hat{B}_i\right)$ controlled by $\hat{u}(s_i) = \hat{K}(s_i)\, x$. Let $\hat{R}_i$ be obtained from R_0 by deleting all rows and columns associated with faulty actuators (note that $\hat{R}_i \neq R_i$: in R_i the rows and columns associated with faulty actuators were zeroed). Then $\hat{R}_i^{-1}$ exists, and one has $\beta_i = \hat{B}_i \hat{R}_i^{-1} \hat{B}_i^T$. It follows that W_i^* is the minimal cost associated with the optimal control of configuration s_i hence the result.* ■

2.4 The lattice of system configurations

Because set inclusion is a partial order relation, the set of configurations is a lattice, noted $L(S, \subseteq)$, meaning that every pair $(s_1, s_2) \in S^2$ has a minimal element $(s_1 \cap s_2)$ and a maximal element $(s_1 \cup s_2)$ [10].

Definition 10 *The predecessors $\mathbb{P}(s)$ and the successors $\mathbb{S}(s)$ of a configuration $s \in S$ are defined by*

$$\mathbb{P}(s) = \{s_i, i \in I : s_i \supseteq s\}$$
$$\mathbb{S}(s) = \{s_i, i \in \mathcal{I} : s \subseteq s_i\}$$

Note that a configuration s belongs both to $\mathbb{P}(s)$ and $\mathbb{S}(s)$.

Definition 11 *A property is bottom-up monotonous (bum) on a lattice $L(S, \subseteq)$ if the fact that it is satisfied for one configuration implies that it is satisfied for all its predecessors.*

Definition 12 *Let $S_{\mathcal{P}}$ be the set of configurations that satisfy some property $\mathcal{P}$. A minimal configuration for $\mathcal{P}$ is a configuration $s_m \in S_{\mathcal{P}}$ such that $\mathbb{S}(s_m) \cap S_{\mathcal{P}} = \{s_m\}$.*

Lemma 13 *Let $m(S_{\mathcal{P}})$ be the set of minimal configurations for $\mathcal{P}$. If the property $\mathcal{P}$ is bum on $L(S, \subseteq)$ then it holds that*

$$S_{\mathcal{P}} = \bigcup_{s_m \in m(S_{\mathcal{P}})} \mathbb{P}(s_m) \tag{7}$$

Proof. *From the fact that $\mathcal{P}$ is bum, it follows that $\forall s_m \in m(S_{\mathcal{P}})$, $\mathbb{P}(s_m) \subseteq S_{\mathcal{P}}$. Conversely, let $s \in S_{\mathcal{P}}$, then either $s \in m(S_{\mathcal{P}})$ or $\exists s_m \in \mathbb{S}(s) : s_m \in m(S_{\mathcal{P}})$.* ■

Remark 14 *Stabilizability is bum on $L(S,\subseteq)$. Indeed, one obviously has*

$$\forall s_1 \in S_{stab}, \forall s_2 \in \mathbb{P}(s_1) : s_2 \in S_{stab}$$

Lemma 15 *SF-recoverability is bum on $L(S,\subseteq)$.*
Proof. *Let $s_i \in S_{recov}$. Any configuration $s_j \in \mathbb{P}(s_i)$ is such that $W_j^* \leq W_i^*$ because the set of non zero columns in B_j includes the set of non-zero columns in B_i.* ■

From the characterization of bum properties (7), it follows that S_{stab} and S_{recov} are completely defined by the knowledge of $m(S_{stab})$ and $m(S_{recov})$ the sets of minimal stabilisable and minimal recoverable configurations (mSC and mRC).

3 Fault tolerance design

3.1 Definitions

(1) The span of an admissible state feedback $u = Kx$ is the subset $S_{recov}(K) \subseteq S_{recov}$ of configurations that it recovers (i.e. for which it is admissible).

(2) Let K_1 and K_2 be two admissible state feedbacks. K_1 dominates K_2 ($K_1 \succeq K_2$) if $S_{recov}(K_2) \subseteq S_{recov}(K_1)$.

(3) Let $\mathcal{K} = \{K_i, i = 1,..q\}$ be a bank of admissible state feedbacks, and let $S_{specif} \subseteq S_{recov}$ be given. The bank is complete over a given set of configurations S_{specif} (abbreviated into "complete") if

$$S_{specif} \subseteq \bigcup_{i=1,..q} S_{recov}(K_i) \tag{8}$$

meaning that any configuration in the specified subset S_{specif} is recovered by at least one state feedback in $\mathcal{K}$.

(4) A complete bank $\mathcal{K}$ is minimal if no proper subset of $\mathcal{K}$ is complete (i.e. it contains no dominated state feedback).

Remark 16 *In passive fault tolerance (PFT), the same control law K_0 is used in the nominal and in the faulty cases, hence $\mathcal{K} = \{K_0\}$ and $S_{specif} \subseteq S_{recov}(K_0)$. Obviously, there may be no solution K_0 for some S_{specif}. In active fault tolerance (AFT), a solution K_l exists for each recoverable configuration, hence $\mathcal{K} = \{K_i, i = 1,..|S_{specif}|\}$. The passive / active (PACT) intermediate case is associated with $\mathcal{K} = \{K_i, i = 1,..q\}$ where $q < |S_{specif}|$ and (8) holds true.*

3.2 Problem setting

Given the system (A, B_0), an admissible performance specification N, and a set of configurations $S_{specif} \subseteq S_{recover}$ to be recovered, the Linear Quadratic Reconfiguration-based Fault Tolerance design problem is to find a bank of admissible control laws that is complete over S_{specif}. Since there may exist several

control laws that recover a given configuration, the problem also includes the design of a decision procedure to select the one to be applied when configuration s_i occurs.

3.3 The design of a complete bank

The simplest algorithm for the design of a complete bank is as follows :

Algorithm 1.

(1) For every configuration $s_i \in S_{specif}$

(2) Design a state feedback K_i that recovers configuration s_i.

Note that in (2), any state feedback that results in an admissible performance can be chosen. Obviously, selecting an optimal state feedback K_i^* associated with each configuration s_i results in the optimal performance when this configuration occurs. Since the nominal configuration and each faulty configuration are recovered by a specific control law, the obtained complete bank implements AFT, at the cost of having $|\mathcal{K}| = |S_{specif}|$ (unless there exists several configurations that have identical optimal solutions). Note that a complete bank with less control laws is obtained by discarding the dominated state feedbacks K_i^* (if such state feedbacks exist).

In the sequel, we use the fact that SF-recoverability is bum to design a complete bank that contains at most $|m(S_{specif})|$ control laws.

3.4 Bottom-up monotonicity and complete banks

Definition 17 *A state feedback K is bum-extensive over a configuration $s_i \in S_{recov}$ if its span includes $\mathbb{P}(s_i)$.*

Remark 18 *K being bum-extensive over s_i means that it is a reliable control in the sense defined by [23], [24], namely it recovers any configuration that includes s_i (i.e. such that the subset of faulty actuators is included in $s_0 \backslash s_i$).*

The interest of bum-extensivity is shown by the following lemma.

Lemma 19 *Let $\mathcal{K}$ be a bank such that, for any configuration in $m(S_{specif})$, there is a state feedback that is bum-extensive over it. Then it is complete.*

Proof. *For any configuration in $m(S_{specif})$, there is a state feedback in $\mathcal{K}$ that recovers from the faults associated with all its predecessors. $\mathcal{K}$ is complete because SF-recoverability is a bum property and therefore S_{specif} is the union of the predecessors of its minimal elements.* ∎

3.5 Design of a PACT bank

From the previous lemma, a complete bank can be obtained by designing a bum-extensive control K_m associated with each configuration $s_m \in m(S_{specif})$. The design algorithm becomes:

Algorithm 2.

(1) For every minimal configuration $s_m \in m(S_{specif})$

(2) Design a state feedback K_m that is bum-extensive over s_m.

From (6), this boils down to finding a feedback K_m such that $\forall s_i \in \mathbb{P}(s_m)$ there exists a matrix $W_i = W_i^T > 0$ that satisfies

$$W_i \leq N \tag{9}$$
$$Q + K_m^T R_i K_m + W_i^T F_{im} + F_{im}^T W_i \leq 0$$

where $F_{im} \triangleq A + B_i K_m$.

The following proposition gives a solution that is based on the reliable control idea introduced by [24].

Proposition 20 *Let B_m be the actuation matrix associated with $s_m \in m(S_{specif})$, and let W_m^* be the unique stabilizing solution of the algebraic Riccati equation*

$$A^T W_m^* + W_m^* A - W_m^* \beta_m W_m^* + Q = 0 \tag{10}$$

Then, $K_m = -R_0^{-1} B_0^T W_m^*$ and $\forall s_i \in \mathbb{P}(s_m) : W_i = W_m^*$ satisfy (9).

Proof. *With each configuration $s_m \in m(S_{specif})$ associate the control law $u_m^* = -R_0^{-1} B_0^T W_m^* x$ where W_m^* is the unique stabilizing solution of (10). Configuration s_m being recoverable because $m(S_{specif}) \subseteq S_{recov}$, it is stabilized by u_m^* at an admissible cost, i.e. $W_m^* \leq N$. From [24] u_m^* is a reliable control, that stabilizes any configuration $s_i \in \mathbb{P}(s_m)$ at a cost less than $x_0^T W_m^* x_0$, therefore it is bum-extensive over s_m.* ∎

3.6 Decision procedure

Let $\mathcal{K} = \{K_l, l = 1, .., q\}$ be a complete bank. For each configuration $s_i \in S_{specif}$, let $\mathcal{K}_i \subseteq \mathcal{K}$ be the subset of state feedbacks by which it can be recovered. Selecting the one to be applied when s_i occurs is the aim of the decision procedure.

Assume that s_i has been detected and isolated at time t_i, and let $K_j \in \mathcal{K}_i$ and $x(t_i)$ be respectively the feedback gain that is switched-on and the system state at time t_i. Let T_i be the length of the time window until the next configuration is switched-on. The cost paid during the time interval $[t_i, t_i + T_i[$ is

$$\tilde{J}_i(x(t_i), K_j, T_i) \triangleq x^T(t_i) W_{ij} x(t_i) - x^T(t_i + T_i) W_{ij} x(t_i + T_i)$$

where W_{ij} is the solution of the Lyapunov equation

$$Q + K_j^T R_i K_j + W_{ij}^T F_{ij} + F_{ij}^T W_{ij} = 0$$

$F_{ij} = A + B_i K_j$ is the closed-loop matrix that results from configuration s_i being controlled by the state feedback K_j and $x(t_i + T_i) = [\exp F_{ij} T_i] x(t_i)$. The minimal cost decision is to select the state feedback $\tilde{K}_i = \arg\min_{K_j \in \mathcal{K}_i} \tilde{J}_i(x(t_i), K_j, T_i)$, which is seen to depend both on the initial state $x(t_i)$ and on the time window T_i. Since T_i is obviously unknown, an estimation has to be used, the simplest one being $T_i = \infty$ (it is estimated that no further fault will occur after the loss of the actuators in $s_0 \backslash s_i$). Then one has

$$\lim_{T_i \to \infty} \tilde{J}_i(x(t_i), K_j, T_i) = J_i(x(t_i), K_j) = x^T(t_i) W_{ij} x(t_i)$$

Note that the decision procedure can be made independent on the initial state by selecting

$$\tilde{K}_i = \arg \min_{K_j \in \mathcal{K}_i} \max_{\|x\|=1} x^T W_{ij} x \tag{11}$$

which boils down to select the state feedback for which the largest eigenvalue of the associated cost matrix W_{ij} is minimal, $\tilde{K}_i = \arg\min_{K_j \in \mathcal{K}_i} \lambda_{\max}(W_{ij})$.

4 Design trade-off

4.1 Reliability overcost

In Algorithm 1, selecting the optimal state feedback K_i^* associated with each configuration $s_i \in S_{specif}$ results in the minimal cost $J_i(x, K_i^*) = x^T W_i^* x$ when this configuration occurs, and in the minimal performance degradation with respect to the nominal case, namely

$$J_i(x, K_i^*) - J_0(x, K_0^*) = x^T \left(W_i^* - W_0^*\right) x$$

Consider now the design that results from Algorithm 2 and Proposition 1. For each configuration $s_i \in S_{specif}$ let $s_{m(i)}$ be the minimal configuration in $m(S_{specif})$ whose state feedback $K_{m(i)}^*$ is selected when fault i occurs. Then the performance is $J_i(x, K_{m(i)}^*) = x^T W_{m(i)}^* x$ and the performance degradation with respect to the nominal case is

$$J_i(x, K_{m(i)}^*) - J_0(x, K_{m(0)}^*) = x^T \left(W_{m(i)}^* - W_{m(0)}^*\right) x$$

In both designs, the performance is admissible for each configuration $s_i \in S_{specif}$ (and the performance degradation is obviously zero in the nominal case). However, controlling a configuration $s_i \in S_{specif}$ by the reliable state feedback $K_{m(i)}^*$ instead of the optimal state feedback K_i^* induces a reliability overcost $\rho(x, s_i) = x^T \left(W_{m(i)}^* - W_i^*\right) x$, which is the performance degradation with respect to the best one that could have been obtained when configuration i occurs.

The reason to accept this overcost is that the design based on Algorithm 1 results in a bank of $|S_{specif}|$ control laws while the design of Algorithm 2 produces only $|m(S_{specif})|$ control laws. It is therefore of interest to consider the design trade-off between the reliability overcost and the cost of the bank of control laws to be implemented.

4.2 Minimizing the reliability overcost

The trade-off problem can be set in many different ways. A sensible one is to find a bank of bum-extensive state feedbacks over $m(S_{specif})$ such that the reliability overcost associated with the nominal configuration is minimized. Looking for bum-extensive feedbacks over $m(S_{specif})$ guarantees that the bank is complete. Minimizing the reliability overcost associated with the nominal configuration

aims at reducing the conservativeness of the FT strategy when the nominal configuration occurs, a case that (hopefully) occurs most frequently.

In this problem setting, Algorithm 2 is modified as follows:

Algorithm 3.

(1) For each minimal configuration $s_m \in m\left(S_{specif}\right)$

(2) Design a state feedback K_m that is optimal for s_0 under the constraint that it is bum-extensive over s_m.

From (6), point (2) boils down to finding a feedback K_m such that

$$\begin{aligned} &\min \lambda_{\max}\left(W_{0m}\right) \\ &Q + K_m^T R_0 K_m + W_{0m}^T F_{0m} + F_{0m}^T W_{0m} = 0 \end{aligned} \tag{12}$$

and

$$\forall s_i \in \mathbb{P}\left(s_m\right), \exists W_{im} = W_{im}^T > 0 : \left\{ \begin{array}{l} W_{im} \leq N \\ Q + K_m^T R_i K_m + W_{im}^T F_{im} + F_{im}^T W_{im} \leq 0 \end{array} \right. \tag{13}$$

4.3 A Newton-Kleinman algorithm

Solving problem (12) under the constraints (13) is not an easy task. However, note that from [24], the control law $u_m^* = -R_0^{-1} B_0^T W_m^* x$ satisfies (13). Let $H_m\left(t\right)$, $t = 0, 1, 2, ...$ be a sequence of symmetric positive definite matrices, initialized as $H_m\left(0\right) = W_m^*$. Starting with $K_m^* = -R_0^{-1} B_0^T W_m^*$, the following Newton-Kleinman algorithm produces a sequence of state feedbacks $K_m\left(t\right) = -R_0^{-1} B_0^T H_m\left(t\right)$ that decrease the reliability overcost while remaining bum-extensive over s_m. The notations are as follows : $F_{0m}(t) \triangleq A - \beta_0 H_m\left(t\right)$ is the closed-loop matrix associated with the nominal configuration s_0 and the state feedback $K_m\left(t\right)$, $G_{0m}(t) \triangleq A - \beta_0 W_{0m}\left(t\right)$ is the closed-loop matrix associated with the state feedback $W_{0m}\left(t\right)$ where $W_{0m}\left(t\right)$ is the solution of the Lyapunov equation

$$Q + K_m^T\left(t\right) R_0 K_m\left(t\right) + W_{0m}^T\left(t\right) F_{0m}\left(t\right) + F_{0m}^T\left(t\right) W_{0m}\left(t\right) = 0 \tag{14}$$

P_2 is the set of pairs (p, q) such that $p, q \geq 0, p + q = 1$, ε is an arbitrary small positive number, $\|\|$ is a given matrix norm, and $\mathcal{H}_m$ is the set of symmetric positive definite matrices such that

$$H_m \in \mathcal{H}_m \implies K_m = -R_0^{-1} B_0^T H_m \text{ satisfies (13)}$$

Algorithm 4.

(1) Initialization : $H_m\left(0\right) = W_m^*, W_{0m}\left(-1\right) = \infty$

(2) While $\left\|W_{0m}\left(t\right) - W_{0m}\left(t-1\right)\right\| > \varepsilon$

(3) Solve (14) for $W_{0m}\left(t\right)$

(4) Update $H_m\left(t+1\right) = \bar{p}\left(t\right) H_m\left(t\right) + \bar{q}\left(t\right) W_{0m}\left(t\right)$ where

$$\begin{aligned} &\left(\bar{p}\left(t\right), \bar{q}\left(t\right)\right) \in P_2 \\ &\bar{q}(t) = \max\left\{q : q \in \left[0, 1\right], H_m\left(t+1\right) \in \mathcal{H}_m\right\} \end{aligned} \tag{15}$$

Proposition 21 *The sequence of state feedbacks* $K_m(t) = -R^{-1}B_0^T H_m(t)$, $t = 0, 1, 2, ...$ *stabilize configuration* s_0, *are bum-extensive over* s_m, *and such that*

$$W_0^* \leq ... \leq W_{0m}(t+1) \leq W_{0m}(t) \leq ... \leq W_{0m}(0) \leq W_m^* \leq N \qquad (16)$$

Proof. *We here only give a sketch of the proof. (A) The initial solution* $K_m(0) = -R^{-1}B_0^T W_m^*$ *stabilizes* s_0, *is bum-extensive over* s_m *and such that* $W_{0m}(0) \leq W_m^*$, *from the basic result in [24]. Moreover,* $W_m^* \leq N$ *since* $s_m \in m(S_{specif}) \subseteq S_{recov}$. *(B) As long as condition (2) is not fulfilled, all solutions stabilize* s_0 *while remaining bum-extensive over* s_m. *bum-extensivity is satisfied since* $H_m(t) \in \mathcal{H}_m$ *from the updating rule (15). The proof of stability is similar to the one in [5], by using an update law* $F_{0m}(t+1) = p(t) F_{0m}(t) + q(t) G_{0m}(t)$, *where* $(p, q) \in P_2$, *instead of Kleinman's update law* $F_{0m}(t+1) = G_{0m}(t)$. *This also allows to prove that the successive costs* $W_{0m}(t)$ *and* $W_{0m}(t+1)$ *are decreasing, which guarantees the convergence of the algorithm.* ■

5 Fault Tolerance Evaluation

5.1 Reliability data

The lattice of configurations can be enriched with temporal information. Let $r_\alpha(t_1, t_2)$ be the reliability of actuator α, i.e. the probability that it is functional at time $t_2 > t_1$ subject to the condition that it was functional at time t_1 [4]. Assuming that this function is known for all actuators of an autonomous system (meaning that it cannot be repaired in operation), that actuator failures are independent, and that s_0 is the system configuration at time 0, the probability that the system is in configuration s at time t is given by

$$\Pr(s, t) = \prod_{\alpha \in s} r_\alpha(t, 0) \prod_{\alpha \notin s} [1 - r_\alpha(t, 0)] \qquad (17)$$

5.2 The reliability of a property

Remember the notation $S_{\mathcal{P}}$ for the set of configurations that satisfy some property $\mathcal{P}$. The probability for this property to remain true during a given time window $[0, T]$ is obviously equal to the probability that, on the time window $[0, T]$, only configurations that belong to $S_{\mathcal{P}}$ occur as the result of faults. It follows that the number $REL(\mathcal{P}, T)$ defined by

$$REL(\mathcal{P}, T) \triangleq \sum_{s \in S_{\mathcal{P}}} \Pr(s, T) \qquad (18)$$

is nothing but the **reliability of property** $\mathcal{P}$ as a function of time T, while $MTTF(\mathcal{P})$ defined by

$$MTTF(\mathcal{P}) \triangleq \int_0^\infty REL(\mathcal{P}, T)\, dT$$

is the **mean-time to failure of property** $\mathcal{P}$. In particular, $REL\left(stab,T\right)$ is the probability that no fault such that the system becomes non-stabilisable occurs during the time window $[0,T]$, while $REL\left(recov,T\right)$ is the probability that no non-recoverable fault occurs on $[0,T]$, thus providing a most natural and sensible FT performance measure.

Remark 22 *Assume that, in order to reduce the cost of the implemented PACT bank, the designer accepts to recover only the faults in* $S_{specif} \subseteq S_{recov}$. *The maximal accepted decrease in the FT performance,* $REL\left(recov,T\right) - REL\left(specif,T\right)$, *is obviously the cost to pay for this design trade-off.*

5.3 The usefulness of components

Faults in different actuators (more generally in different system components) may have more or less severe consequences. Based on the previous evaluation of the FT performance, we propose a usefulness measure that extends the one in [2], and may be used to decide about the actuators required reliability, their maintenance policy, their possible duplication (hardware redundancy).

For this, $\forall s \subseteq s_0$ we extend the previous notation as:

- $S\left(s\right) = \{s_i : s_i \subseteq s\}$ the subset of configurations associated with the nominal configuration s,
- $S_{recov}\left(s\right)$ the subset of recoverable ones,
- $REL\left(recov,T,s\right)$ the probability that no non-recoverable fault occurs during $[0,T]$ when the initial configuration is s.

Note that with this extension, the previous notations S, S_{recov} and $REL\left(recov,T\right)$ become respectively $S\left(s_0\right)$, $S_{recov}\left(s_0\right)$ and $REL\left(recov,T,s_0\right)$. Let us now define the **usefulness** of a subset of actuators $s_1 \subseteq s_0$ by

$$u\left(s_1\right) \triangleq \frac{REL\left(recov,T,s_0\right) - REL\left(recov,T,s_0 \backslash s_1\right)}{REL\left(recov,T,s_0\right)}$$

namely the usefulness of the subset s_1 is the fraction of the FT performance that is lost when replacing the nominal configuration s_0 by the nominal configuration $s_0 \backslash s_1$ (assuming $REL\left(recov,T,s_0\right) \neq 0$ is obviously not restrictive). This usefulness measure enjoys the following properties :

1) Since $\forall T > 0, s_1 \subseteq s_0 \implies REL\left(recov,T,s_0 \backslash s_1\right) \leq REL\left(recov,T,s_0\right)$ one obviously has

$$\forall s_1 \subseteq s_0 : u\left(s_1\right) \in [0,1]$$

2) Actuator subsets $s_1 \subseteq s_0$ such that $u\left(s_1\right) = 0$ are **useless** for the problem under consideration, since the FT performance is the same with the smaller actuation system $s_0 \backslash s_1$.

3) It is easy to show that actuators that do not belong to any configuration in $m\left(S_{recov}\right)$ are useless.

4) Actuator subsets $s_1 \subseteq s_0$ such that $u(s_1) = 1$ are called **cut-sets**, meaning that there is no FT at all for the problem under consideration, when those actuators are not present.

5) Cut-sets are not unique. **Critical actuator subsets** are minimal cut-sets, i.e. any proper subset of a critical actuator subset is not a cut-set.

6) It is easy to show that critical subsets are minimal hitting sets of $m(S_{recov})$ (the interested reader may consult [13] for a link between diagnosis and hitting sets.)

6 Example

6.1 Nominal system

We consider the academic example of a 6^{th} order system with 4 actuators $abcd$. The pair (A, B_0), where $B_0 = (b_a, b_b, b_c, b_d)$ is :

$$A = \begin{pmatrix} 0 & 1 & 1 & 2 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \quad B_0 = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{19}$$

Let $Q = I_6$ and $R = I_4$. For configuration s_0, the optimal cost matrix W_0^* is characterized by $\lambda_{\max}(W_0^*) = 7.3554$.

For simplicity, configurations are indexed from the subset of their missing actuators, according to the following table.

Configuration	$abcd$	abc	abd	ab	acd	ac	ad	a
Index i	0	1	2	3	4	5	6	7
Configuration	bcd	bc	bd	b	cd	c	d	$\varnothing$
Index i	8	9	10	11	12	13	14	15

Table 1 : Correspondence between configurations and their index

6.2 Admissibility

Define admissibility by

$$J_i(x_0, K) \leq \tau x_0^T W_0^* x_0 \tag{20}$$

where $s_i \subseteq s_0$, $u = Kx$ is the control law used in configuration s_i, $J_i(x_0, K)$ is the resulting cost, and $\tau > 1$ is an admissible performance degradation factor.

Using $\tau = 15$ in the specification (20), results in the set of recoverable configurations $S_{recov} = \{0, 1, 2, 3, 4, 8, 9, 10\}$ with the mRCs $m(S_{recov}) = \{3, 4, 9, 10\}$. On Figure 1, vertices in S_{recov} are white and vertices in $m(S_{recov})$ have a bold contour.

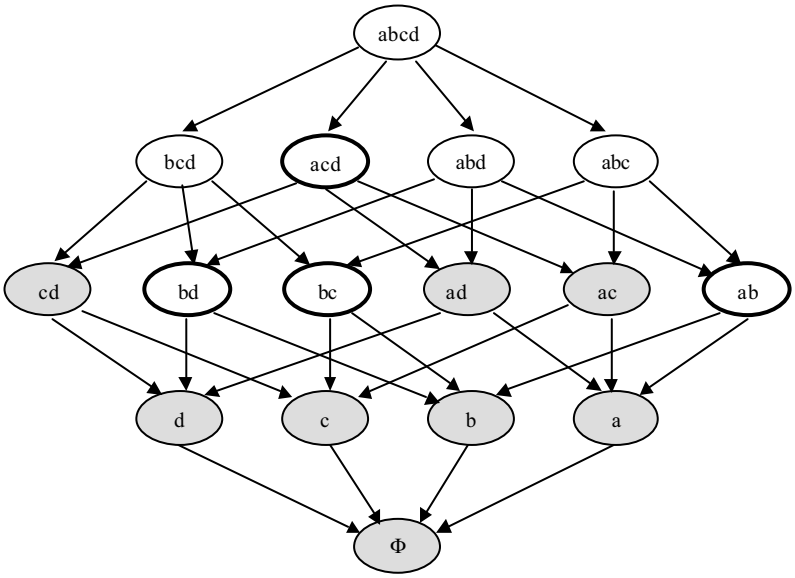


Figure 1 : The lattice of the example

6.3 PACT design

Design with algorithm 2 and Proposition 1. This design uses the matrices W_i^* associated with the mRCs $s_i \in m(S_{recov}) = \{3, 4, 9, 10\}$. The respective performance indexes are $\lambda_{\max}(W_3^*) = 17.4285$, $\lambda_{\max}(W_4^*) = 32.9450$, $\lambda_{\max}(W_9^*) = 16.5649$, $\lambda_{\max}(W_{10}^*) = 18.6938$, and the spans of the associated state feedbacks are $S(K_3^*) = \{0, 1, 2, 3, \underline{8}, \underline{9}\}$, $S(K_4^*) = \{0, \underline{1}, 4, \underline{8}\}$, $S(K_9^*) = \{0, 1, 8, 9\}$, $S(K_{10}^*) = \{0, 2, 8, 10\}$. Note that for each $s_m \in m(S_{recov})$ the span $S(K_m^*)$ indeed includes $\mathbb{P}(s_m)$ - the configurations in $S(K_m^*) \setminus \mathbb{P}(s_m)$ are underlined - and that non-mRC can be recovered by several control laws, hence the need for a decision procedure. Table 2 gives, for each recoverable configuration, the list of control laws by which it can be recovered. Among these, the one with the smallest maximal cost is underlined.

s_0	s_1	s_2	s_3	s_4	s_8	s_9	s_{10}
$K_3^*, K_4^*, \underline{K_9^*}, K_{10}^*$	$K_3^*, K_4^*, \underline{K_9^*}$	$\underline{K_3^*}, K_{10}^*$	$\underline{K_3^*}$	$\underline{K_4^*}$	$K_3^*, K_4^*, \underline{K_9^*}, K_{10}^*$	$K_3^*, \underline{K_9^*}$	$\underline{K_{10}^*}$

Table 2 : The RC-based PACT

Domination relation. Table 3 shows the domination relation deduced from the sets $S(K_m^*)$, $s_m \in m(S_{recov})$ where 1 means that the state feedback K_i^* (row i) dominates the state feedback K_j^* (column j).

$\mathcal{D}\nearrow$	K_3^*	K_4^*	K_9^*	K_{10}^*
K_3^*	1	0	1	0
K_4^*	0	1	0	0
K_9^*	0	0	1	0
K_{10}^*	0	0	0	1

Table 3 : Domination relation

It follows that a bank of only three state feedbacks, namely $K_i^*, s_i \in \{3,4,10\}$, can recover from all recoverable faults, since the set of non-dominated mRCs is $\{3,4,10\}$. The resulting PACT is

s_0	s_1	s_2	s_3	s_4	s_8	s_9	s_{10}
$\underline{K_3^*}, K_4^*, K_{10}^*$	$\underline{K_3^*}, K_4^*$	$\underline{K_3^*}, K_{10}^*$	$\underline{K_3^*}$	$\underline{K_4^*}$	$\underline{K_3^*}, K_4^*, K_{10}^*$	$\underline{K_3^*}$	$\underline{K_{10}^*}$

Table 4 : The RC-based PACT with three control laws

Some costs are increased with respect to the PACT of Table 2, indeed the state feedback K_9^* associated with $\lambda_{\max}(W_9^*) = 16.5649$ is replaced by K_3^* associated with $\lambda_{\max}(W_3^*) = 17.4285$.

Design with Algorithm 4. We illustrate this point by considering configuration 3. In the previous design, $K_3^* = -R^{-1}B_0W_3^*$ is admissible for all configurations $s_i \in \{0,1,2,3,8,9\}$. Using K_3^* for configuration 0 gives the cost matrix $W_{03}(0)$ with $\lambda_{\max}[W_{03}(0)] = 12.2667$. However, any control law $u_3 = -R_0^{-1}B_0H_3x$ where H_3 satisfies

$$H_3 = H_3^T > 0 \tag{21}$$

$$k = 0,1,2,3 : A - \beta_k H_3 \text{ Hurwitz} \tag{22}$$

$$k = 1,2,3 : \left\{ \begin{array}{c} W_{k3} \leq 15W_0^* \\ Q + H_3\beta_k H_3 + W_{k3}F_{k3} + F_{k3}^T W_{k3} = 0 \end{array} \right. \tag{23}$$

$$W_{03} < W_{03}(0) \tag{24}$$

is bum-extensive over configuration 3, i.e. it recovers the configurations in $\mathbb{P}(3) = \{0,1,2,3\}$ - from (23)(24) - and it is better than K_3^* for configuration 0, from (24).

In order to illustrate the computations, we first apply the pure Newton-Kleinman algorithm, where the control update uses $(p(t), q(t)) = (0,1)$ at every iteration t. Starting with $H_3(0) = W_3^*$ the cost matrix $W_{03}(t)$ decreases from $W_{03}(0)$ to W_0^*, providing the sequence $\lambda_{\max}(W_{03}(t))$ shown in Table 5.

Iteration t	0	1	2	3	4	5	6
$\lambda_{\max}(W_{03}(t))$	12.2667	9.0039	7.6082	7.3744	7.3564	7.3554	7.3554

Table 5 : The pure Newton-Kleinman sequence

However, as soon as the first iteration, $H_3(1)$ violates (23). The update law in the algorithm must therefore be used. The results displayed in Table 6 were obtained in two iterations. $K_3(2) = -R_0^{-1}B_0H_3(2)$ is bum-extensive, and decreases the reliability overcost by $17,18\%$, when compared with $K_3^* = -R_0^{-1}B_0W_3^*$.

Iteration t	0	1	2
$\lambda_{\max}[W_{03}(t)]$	12.2667	10.1593	10.1592
$q_{max}(t)$	0.6479	0.0001	0

Table 6 : Results for configuration s_3

Cost based decision procedure. Applying the proposed approach to each configuration in $m(S_{recov}) = \{3,4,9,10\}$ gives the results in Table 7 (the convergence is achieved in 2 iterations for $s_m \in \{3,4,10\}$ and 3 iterations for $s_m \in \{9\}$).

	Algorithm 1	Algorithm 4	Improvement
$\lambda_{\max}[W_{03}(2)]$	12.2667	10.1592	17,18%
$\lambda_{\max}[W_{04}(2)]$	19.3397	15.8924	17.83%
$\lambda_{\max}[W_{09}(3)]$	12.7427	7.8729	38.22%
$\lambda_{\max}[W_{0,10}(2)]$	10.8692	9.1817	15.53%

Table 7 : Results for the configurations in $m(S_{recov})$

Following (11), the state feedback to be chosen in configuration 0 is $K_9(3) = -R_0^{-1}B_0^TH_9(3)$.

6.4 FT Evaluation

We assume that actuator reliabilities are modeled by Poisson distribution with failure rates

$$\begin{aligned}\lambda(a) &= \lambda(b) = 4\times\ 10^{-6}\ \text{hour}^{-1}\\ \lambda(c) &= \lambda(d) = 4\times\ 10^{-7}\ \text{hour}^{-1}\end{aligned}$$

and the time horizon of interest is $T = 10^5$ hours.

Recoverability. Using these data, the probability for a non-recoverable configuration to occur on $[0,10^5]$ is $0,285$ and therefore $REL(recov,10^5,0) = 0.715$.

Actuators usefulness. From the set $m(S_{recov}) = \{3,4,9,10\}$ it is seen that there is no useless actuator, and that the critical actuator subsets are $\{3,9,10\}$.

7 Conclusion

Under the SR strategy, the set of actuator configurations to be managed by the control system has a lattice structure. Based on the notions of bum property

and bum-extensive control, this paper has proposed a design approach for the Reconfiguration-based Fault Tolerance Linear Quadratic problem. From the characterization of recoverable faults, mRCs have been defined and have been shown to allow the design of a PACT bank of control laws, that is able to recover from all recoverable faults. Such a scheme results in a large simplification of the diagnosis and the reconfiguration procedures, since FDI must only distinguish between those configurations that are associated with different control laws in the designed bank, and FTC must only switch to the appropriate one, as determined by the decision procedure. A domination relation has been shown to decrease - when possible - the number of PACT control laws, and a Newton - Kleinman algorithm has been proposed to reduce the reliability overcost. Finally, it has also been shown that the lattice of system configurations, associated with reliability data, provides a convenient and sensible framework for the evaluation of the system FT performance and of its components usefulness. Within the general frame so defined, future research is aimed at investigating more general algorithms and approaches for the design of bum-extensive controls, and addressing the trade-off between the reliability overcost and the complexity of the PACT strategy.

Acknowledgment. The author is indebted to Denis Berdjag and Ke Zhang for their help in the example computations.

References

[1] Blanke,M., Kinnaert, M., Lunze, J. and M. Staroswiecki (2003), Diagnosis and Fault Tolerant Control, Springer Verlag, Berlin, Heidelberg, New York.

[2] Commault, C., Dion, J. M., Trinh, D. H. and T. H. Do (2010). Sensor classification for the fault detection and isolation, a structural approach, *Int. J. Adapt. Control Signal Processing*, to appear.

[3] Frei, C. W., Kraus, F. J. and M. Blanke (1999). Recoverability viewed as a system property. Proceedings of European Control Conference, Karlsruhe, Germany.

[4] Kapur, K.C. and L.R Lamberson (1977). Reliability in Engineering Design, John Wiley & Sons, New York.

[5] Kleinman, D. L. (1968), On an iterative technique for Riccati equation computation. IEEE Transactions on Automatic Control, 13(1):114–115.

[6] Liang Y-W. and D-C Liaw (2006), Common stabilizers for linear control systems in the presence of actuators outage, Applied mathematics and computation, 177(2), 635-643.

[7] Liao, F., Wang, J. L. and G. H. Yang (2002), Reliable robust flight tracking control : an LMI approach, IEEE Tran. Contr. Syst. Techno., 10(1):76-89.

[8] Maki, M., Jiang, J. and K. Hagino (2004), A stability guaranteed active fault-tolerant control system against actuator failures, Int. J. of Robust and Nonlinear Control, 14(12):1061-1077.

[9] Moerder, D. D. et al. (1989), Application of pre-computed control laws in a reconfigurable aircraft flight control system. Journal of Guid. Con. & Dyn. 12(3):325-333.

[10] Priestly, H. A. and Davey, B. A. (1990). Introduction to Lattices and Order. Cambridge, England: Cambridge University Press.

[11] Pujol, G., Rodellar, J., Rossell, J. M. and F. Pozo (2007), Decentralized reliable guaranteed cost control of uncertain systems : an LMI design, IET Control Theory Appl., 1(3):779-785.

[12] Rauch, H. E. (1994), Intelligent fault diagnosis and control reconfiguration, IEEE Control System Magazine, 14(3):6-12.

[13] Reiter, R. (1987). A theory of diagnosis from first principles, *Artificial Intelligence*, 32:57-95.

[14] Staroswiecki, M., Yang, H. and B. Jiang (2007), Active Fault Tolerant Control Based on Progressive Accommodation, Automatica 43(12) : 2070–2076.

[15] Staroswiecki, M. (2003), Actuator faults and the linear quadratic control problem, Proc. of the 42d IEEE Conf. on Decision and Control, CDC'03, Hawaii, USA, 959-965.

[16] Staroswiecki M. (2002), On reconfigurability with respect to actuator failures, Proceedings of the IFAC World Congress, Barcelona, Spain, 15(1).

[17] Staroswiecki, M., Berdjag, D., Jiang B. and K. Zhang (2009), PACT : a PAssive / ACTive approach to fault tolerant stability under actuator outages, Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R. China, 7819-7824.

[18] Staroswiecki M. and D. Berdjag (2009), Passive / active fault tolerant control for LTI systems with actuator outages, European Control Conference 2009 - ECC'09, Budapest, Hungary, 2506-2511.

[19] Staroswiecki M. and D. Berdjag (2010), A general fault tolerant linear quadratic control strategy under actuator outages, to appear in Int. J. of System Science, special issue on Fault Detection and Isolation and Fault Tolerant Control.

[20] Steffen T., Michail K., Dixon R., Zolotas A. and R. Goodall (2009), Optimal Passive Fault Tolerant Control of a High Redundancy Actuator, IFAC Symposium Safeprocess 2009, Barcelona, Spain, 1234-1239.

[21] Stoustrup J. and V. D. Blondel, (2004), A simultaneous stabilization approach to (Passive) fault tolerant control, in Proc. of the ACC, Boston, Mass, 1817-1822.

[22] Tao, G., Joshi, S. M. and X. L. Ma (2001), Adaptive state feedback and tracking control of systems with actuator failures, IEEE Tran. Automat. Contr., 46 (1):78 - 95.

[23] Veillette, R. J. et al. (1992). Design of Reliable Control systems, IEEE Trans. Automat. Contr., 37(3):290-304.

[24] Veillette, R. J. (1995), Reliable Linear-quadratic State-feedback Control, Automatica, 32(1):137-143.

[25] Wu N. E., Zhou K. and G. Salomon (2000), Control reconfigurability of linear time-invariant systems, Automatica, 36(9):1767-1771.

[26] Yu, X., Li, Y., Wang, X., and K. Zhao (2006), An Autonomous Robust Fault Tolerant Control System, IEEE International Conference on Information Acquisition, Shandong, China, 1191-1196.

[27] Yang, G. H., Wang, J. L. and Y. C. Soh (2003), Reliable H_∞ controller design for linear systems, Automatica, 37(5):717 - 725.

New Perspectives for Research in Fault Tolerant Control

Extended Abstract

Ron J Patton
University of Hull, UK r.j.patton@hull.ac.uk

Conventional feedback control designs may result in unsatisfactory performance and stability in the event of component malfunctions or faults. Hence, fault tolerant control (FTC) systems are being developed which are capable of maintaining acceptable system integrity and performance subject to faults and failures with applications to a wide range of engineering systems, from vehicles and aerospace, marine systems, mechatronics, electric power (generation and distribution), offshore wind technology, chemical processes and bio-medical applications involving control.

Even after more than two decades of research FTC remains today a mainly research topic. The driving research challenge is to produce architectures and methods that are attractive for technology transfer with serious intention for practical end-user added value.

Fault-tolerance and also robustness in control can only be traded with performance and it is well known that the so-called active FTC methods (or AFTC) facilitate a mechanism, for maintaining satisfactory control performance and stability in the presence of faults, uncertain dynamics and feedback system complexity. The FTC goal is to maintain functional integrity, reliability, stability and admissible performance. Typically the required goal may be achieved through (1) control robustness, (2) fault compensation/accommodation or (3) through system redundancy and reconfiguration. But the main question is how should these concepts be combined?

This goal to achieve admissible performance in the midst of complexity and subject to a limited set of system component fault conditions is certainly an advanced multi-objective requirement that is not achievable using the simpler fixed controller passive approaches to FTC. This requirement is also compounded by the fact that the faults themselves add discrete-event complexity into the feedback system. Hence, current research is concerned with architectures and advanced control and estimation methods calling on recent developments throughout the advanced control and decision literature. The emerging challenge is to accommodate the controller such that there can be a guarantee that the closed-loop system has “admissible behaviour” subject to an expected repertoire of faults.

At or around 2000 the research focused on definitions and FTC concepts associated with reconfigurability, diagnosibility, fault accommodation, robust fault estimation, redundancy structure estimation, etc. At the same time some research in robust fault detection and isolation (FDI) or fault detection and diagnosis (FDD) continued to mature giving rise to interesting applications in reconfigurable FTC systems. Architectures are now emerging that are capable of supporting complex FTC operation for automata, hybrid systems, distributed networks and hierarchical systems. The IFAC Safeprocess Technical Committee community expanded rapidly and the main symposium event of this community is now amongst the largest of IFAC symposia in terms of successful paper contributions. The first of a new conference series dealing with systems control and fault tolerance methods, *Systol’10* bears further testament to the healthy growth in this important field.

In view of these developments this plenary paper has *two* goals. The first is to provide a review of recent research on FTC, focusing on research output that has created significant impact, particularly during the period 2000 to 2010. The second goal is to outline the emerging directions and to map out future perspectives in research that can be of value to end-user practitioners as well as the academic control and systems engineering communities.

Regular Papers

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Design of robust fault detection filters for plants with quantized information

M.L.Corradini * **A. Cristofaro** ** **R. Giambò** * **S. Pettinari** *

* *Scuola di Scienze e Tecnologie, Università di Camerino, via Madonna delle Carceri, 62032 Camerino (MC), Italy (email: {letizia.corradini, roberto.giambo, silvia.pettinari}@unicam.it)*
** *INRIA Rhône-Alpes, 655 Avenue de l'Europe, 38334 St Ismier Cedex (Grenoble), France (email: andrea.cristofaro@inrialpes.fr)*

Abstract: This paper addresses the robust (in the disturbance de-coupling sense) design problem of fault detection filters for quantized uncertain sampled data systems. In the considered set-up, the only available signal is the output variable, which undergoes a quantization process. A novel reduced-order unknown-input observer is proposed such that the estimation error is asymptotically bounded, robustly with respect to disturbances affecting the continuous-time system. Ultimate boundedness of the state variables of the continuous-time plant is also guaranteed.

Keywords: Robust Fault Detection Filters, Unknown Input Observers, Sampled data systems, Quantization.

1. INTRODUCTION

The issue of designing fault detection filters able to decouple disturbances affecting the plant with respect to eventual faults occurring in the system has been widely addresses in the continuous-time framework, see, e.g., C. Aubrun, D. Sauter and J. Yamé (2008), J. Chen, R. J. Patton (1999), P. M. Frank (1990), J. J. Gertler (1991), J. J. Gertler (1998), A. S. Willsky (1976). On the contrary, this issue has received only a limited attention as far as quantized sampled-data (SD) systems are concerned, at least as far as authors are aware, albeit SD systems have been intensively studied during the last decade and are largely used in practice (T. Chen, B. Francis , 1995), (E.N. Rosenwasser, B.P. Lampe , 2000). SD systems are particular digital control systems consisting of continuous-time plant to be controlled, discrete-time controllers controlling them, and ideal continuous-to-discrete and discrete-to-continuous trasformers. Measurements to be used for feedback are transmitted by a digital communication channel, in other words data are quantized before transmission. Therefore the given system evolves in continuous time, but output variables available to measurement are quantized.

In general, the continuous-time plant under consideration may be affected by faults and/or uncertain terms (such as unknown disturbances or model uncertainties). In the case of sampled data systems, what happens during discretization is that disturbances satisfying the matching condition in the original continous-time plant do loose such property in the corresponding discrete-time description (P. Zhang, S. X. Ding , 2008). This effect could make even harder the design of robust tools to be used for detecting the eventual occurrence of faults.

As well known, faults are unespected changes in the systems, due to components malfunction and variations in operating conditions, whose effect is some degradation of the overall system performance (R. Isermann , 1997). In order to perform the control reconfiguration needed to account for the eventual occurrence of faults, it is first necessary to robustly detect them, i.e. to determine if a fault has occured in the system regardless the presence of uncertain term. The most common robust model-based fault detection and isolation (FDI) approach makes use of unknown input observers (UIO's) designed as to make the state estimation error decoupled from the unknown inputs (J. Chen, R. J. Patton, H. Y. Zhang, 1996), (P. M. Frank, X. Ding). In other words, it is generated a diagnosis signal, called residual, which should be indipendent with respect to the system operating state, should respond to faults in characteristic manners, and should be de-coupled from disturbances.

This paper addresses the robust (in the disturbance de-coupling sense) design problem of fault detection filters for quantized uncertain sampled data systems. In the considered set-up, the only available signal is the output variable, which undergoes a quantization process. So, the contribution provided is the design of a novel reduced-order filter coupled with a robust stabilizing controller such that the estimation error is asymptotically bounded, robustly with respect to disturbances affecting the continuous-time system, and the state variables of the continuous-time plant are bounded as well.

2. PROBLEM STATEMENT

Consider a digital feedback control system consisting of the interconnection of a SISO completely observable continuous-time plant, a digital controller and a A/D converter. The plant is affected by an additive unknown disturbance term and may also undergo possible actuator faults belonging to the classes of abrupt faults (stepwise) or incipient faults (drift-like) (R. Isermann , 1997). With

no loss of generality the continuous-time systems is given in the observability canonical form and it is described as follows

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}(u(t) + d(t)) + \mathbf{f}\phi(t) \\ y(t) = \mathbf{c}'\,\mathbf{x}(t) \end{cases} \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the state vector which is not available for measurement, $y(t) \in \mathbb{R}$ is the output, $u(t) \in \mathbb{R}$ is the known input vector, $d(t) \in \mathbb{R}$ is the unknown input (or disturbance), and $\phi(t) \in \mathbb{R}$ is an unknown actuator failures whose distribution matrix $\mathbf{f} \in \mathbb{R}^n$ is supposed known. $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}'$ are known real constant matrices with appropriate dimensions. (Note that with the bold capital letters are matrices while the bold small letters are column vectors.)

Discretize the plant equations, assuming that u is constant during each sampling interval T_c, and assume that observability is preserved by a proper choice of the sampling frequency, so with no loss of generality the discretized system can be trasformed in the observability canonical form by a suitable square invertible matrix $\mathbf{M}$ (see for example P. J. Antsaklis, A. N. Michel (2006)), obtaining:

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{q}\,u(k) + \mathbf{\Delta}(\mathbf{k}) + \mathbf{\Phi}(\mathbf{k}) \\ y(k) = \mathbf{c}'\mathbf{x}(k) = x_2(k) \end{cases} \tag{2}$$

with

$$\mathbf{G} = \mathbf{M}^{-1}\, e^{\mathbf{A}\,T_C}\, \mathbf{M}\,, \tag{3}$$

$$\mathbf{q} = \mathbf{M}^{-1} \left(\int_0^{T_C} e^{\mathbf{A}\tau}\, d\tau \right) \mathbf{b}\,, \tag{4}$$

$$\mathbf{\Delta}(k) = \mathbf{M}^{-1} \int_0^{T_C} e^{\mathbf{A}\sigma}\, \mathbf{b}\, d((k+1)T_C - \sigma)\, d\sigma \tag{5}$$

$$\mathbf{\Phi}(k) = \mathbf{M}^{-1} \int_{kT_C}^{(k+1)T_C} e^{\mathbf{A}((k+1)T_C - \sigma)} \mathbf{f}\, \phi(\sigma) d\sigma \tag{6}$$

Partitioning the state vector $\mathbf{x}(t)$ as $\mathbf{x}(t) = (\mathbf{x_1}(t),\, x_2(t))'$ with $\mathbf{x_1}(t) \in \mathbb{R}^{n-1}$ and $x_2(t) \in \mathbb{R}$, the output signal is exactly the last compotent of the state vector $y(k) = x_2(k)$. Moreover, plant matrices can be partitioned accordingly

$$\mathbf{G} = \begin{pmatrix} \mathbf{G_{11}} & \mathbf{g_{12}} \\ \mathbf{g_{21}}' & g_{22} \end{pmatrix}, \; \mathbf{q} = \begin{pmatrix} \mathbf{q_1} \\ q_2 \end{pmatrix},$$

and $\mathbf{\Delta}(k) = (\mathbf{\Delta_1}(k),\, \Delta_2(k))'$, where $\mathbf{G_{11}} \in \mathbb{R}^{(n-1)\times(n-1)}$, $\mathbf{q_1}$, $\mathbf{\Delta_1}(k) \in \mathbb{R}^{n-1}$ and the other matrices have appropriate dimensions.

Fixed-point quantization is assumed to be added to the A/D converter.

Assumption 1. The state vector is unavailable for measurement except for the output variable

$$w(k) = y(k) + p(k), \quad k \in \mathbb{N} \tag{7}$$

which is affected by the quantization error $p(k)$ bounded by a known constant ρ.

Assumption 2. The scalar q_2 is not null.

Assumption 3. The invariant zeros of $(\mathbf{G}, \mathbf{q}, \mathbf{c}')$ system are Schur. (See P. J. Antsaklis, A. N. Michel (2006).)

With a slight abuse of notation we have written $p(k)$ and $\mathbf{x}(k)$ in place of $p(\mathbf{x}(k))$ and $\mathbf{x}(k\,T_C)$ respectively. In this new approach the class of disturbance signals bounded by a known function of the output available measurements is considered:

Assumption 4. The disturbance term $\mathrm{d}(t) = \mathrm{d}(\mathbf{x}(t))$ depends on the state vector and is bounded by a known piecewise continuous function depending on the quantized output signals available for measurements, that is

$$|\,\mathrm{d}(\mathbf{x}(t))| \leq \beta\,|w(k)| \tag{8}$$

for every $t \in [k\,T, (k+1)T[$, and $\beta > 0$.

Remark 1. This particular class of disturbances have been considerated in order to achieve a better fault detection performance. At present authors are working to expand this disturbance class.

Remark 2. As well known, the matched disturbance term $d(t)$ affecting the continuous time plant produces, after discretization, an unknown input $\bar{\mathbf{\Delta}}(k)$ which does no longer fulfill the matching condition.

This paper addresses the problem of designing robust diagnosis filters for sampled-data systems in the presence of quantization errors. Starting from an extension of the full-order unknown input observer (UIO) proposed by J. Chen, R. J. Patton, H. Y. Zhang (1996), presented by the authors in M. L. Corradini, R. Giambò, S. Pettinari (2009) and shortly summarized in Section 3, a new approach is here proposed in order to overcome the strong conditions required to ensure the decoupling of the residual signals with respect to the disturbance terms. The present work is the SISO case of the previous study presented by the authors in M. L. Corradini, A. Cristofaro, R. Giambò, S. Pettinari (2010). In the following, a novel UIO is proposed guaranteeing the asymptotic boundedness of the estimation error, robustly with respect to disturbances affecting the continuous-time system, and the simultaneous boundedness of state variables of the plant.

3. PRELIMINARY RESULTS

The section summarizes previous results presented in M. L. Corradini, R. Giambò, S. Pettinari (2009). Consider the following full-order observer J. Chen, R. J. Patton, H. Y. Zhang (1996)

$$\begin{cases} \mathbf{z}(k+1) &= \mathbf{F}\mathbf{z}(k) + \mathbf{T}\mathbf{q}\,u(k) + \mathbf{k}\,w(k) \\ \hat{\mathbf{x}}(k) &= \mathbf{z}(k) + \mathbf{h}\,w(k) \end{cases} \tag{9}$$

where $\hat{\mathbf{x}}(k) \in \mathbb{R}^n$ is the estimated state vector and $\mathbf{z}(k) \in \mathbb{R}^n$ is the state of full-order observer, matrices $\mathbf{F}$, $\mathbf{T}$, $\mathbf{k}$ are defined as follows

$$\mathbf{k} = \mathbf{k_1} + \mathbf{k_2} \tag{10}$$

$$\mathbf{T} = \mathbf{I}_{n\times n} - \mathbf{h}\mathbf{c}'\,, \tag{11}$$

$$\mathbf{F} = \mathbf{T}\mathbf{G} - \mathbf{k_1}\mathbf{c}'\,, \tag{12}$$

$$\mathbf{k_2} = \mathbf{F}\mathbf{h}\,, \tag{13}$$

and $\mathbf{h}$ is such that $(\mathbf{c}', \mathbf{T}\mathbf{G})$ is a detectable pair.

Remark 3. Assumpions 2-3 are not necessary in this framework.

Definition 5. Let the residual signal be the output estimation error,

$$r(k) \stackrel{def}{=} w(k) - \mathbf{c}'\hat{\mathbf{x}}(k)\,. \tag{14}$$

In M. L. Corradini, R. Giambò, S. Pettinari (2009) it is proved that a straightforward extention of the full-order UIO proposed by J. Chen, R. J. Patton, H. Y. Zhang (1996) to detect sensor and actuator faults that may affect (1) is not possible. In fact, the following Theorem shows that (9) is unable to ensure, in general, the decoupling of the residual signals with respect to the disturbance terms, still maintaining the sensitivity of residuals with respect to faults.

Theorem 6. The observer (9) is such that the residual signal (14) is disturbance de-coupled if and only if the continuous-time system (1) verifies the condition

$$\mathbf{c}' e^{\mathbf{A}\sigma}\mathbf{b} = 0 \tag{15}$$

for any $\sigma \geq 0$.

See (M. L. Corradini, R. Giambò, S. Pettinari , 2009) for a proof. The condition (15) is structural, since depends on the continuous-time plant matrices. A sufficient condition more general than the one proposed in (M. L. Corradini, R. Giambò, S. Pettinari , 2009) is presented below.

Proposition 4. Consider $\mathcal{S} \subseteq \mathcal{C}^{\perp}$ with $\mathcal{C}^{\perp} = \{\mathbf{v} \in \mathbb{R}^n \,|\, \mathbf{c}'\mathbf{v} = 0\}$ the orthogonal space of $\mathbf{c}'$. If $\mathcal{S}$ is an invariant set under the linear map $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ associated to matrix $\mathbf{A}$ and $\mathbf{b} \in \mathcal{S}$, then the condition (15) is verified.

Proof. By definition for any $\sigma \geq 0$

$$\mathbf{c}' e^{\mathbf{A}\sigma}\mathbf{b} = \sum_{i=0}^{\infty} \frac{\sigma^i}{i!} \mathbf{c}'\mathbf{A}^i\mathbf{b}\,;$$

by hypothesis $\mathbf{c}'\mathbf{A}^i\mathbf{b} = 0$ for any $i \geq 0$, so $\mathbf{c}' e^{\mathbf{A}\sigma}\mathbf{b} = 0$.

Proposition 5. Under the hypotesis of Proposition 4, faults can be detected checking when the norm of the residual signal is larger than a constant $\xi = (1 + |\mathbf{c}'\mathbf{k_1}|)\,\rho$.

Remark 6. Note that in order to detect actuator faults $\phi(t)$ the column vector $\mathbf{f}$ has to be not a multiple of $\mathbf{b}$.

4. A NEW ROBUST FILTER FOR QUANTIZED MEASUREMENTS

In order to avoid the structural constraints arising from the filter (9), a new disturbance decoupled filter is here designed. Note that the class of disturbances here considered has been restricted as defined in Assumption 4. The idea is to design a reduced-order filter such that both the estimation error and the state variables are asymptotically bounded and a particular function based on the norm of $w(k)$ gives information about the occurrence of an actuator fault.

Remark 7. Though this paper does not explicitly address sensor faults, the extension of the presented results to these type of failures is straightforward.

Consider the reduced-order observer:

$$\begin{cases} \hat{\mathbf{x}}(k+1) = \mathbf{G_{11}}\,\hat{\mathbf{x}}(k) + \mathbf{q_1}\,u(k) + \mathbf{g_{12}}\,w(k) \\ r(k) = w(k) \end{cases} \tag{16}$$

where $\hat{\mathbf{x}}(k) \in \mathbb{R}^n$ is the estimated state vector and $r(k)$ is the residual signal generated.

Observer design will be carried out in two steps: *i*) it will be proved that under proper conditions the whole state vector is asymptotically bounded, and the same holds for the estimation error; *ii*) it will be defined a function depending on the norm of $w(k)$ as an indicator of the occurrence of actuator faults.

Before describing the filter design procedure, a crucial lemma for the set-up is proved.

Lemma 7. The invariant zeros of (2) are Schur if and only if

$$\mathbf{P} = \mathbf{G_{11}} - \frac{\mathbf{q_1}}{q_2}\mathbf{g_{21}}' \tag{17}$$

is a Schur matrix.

Proof. See Appendix A for the proof.

4.1 Fault-free case

Supposing for the moment that no actuator faults affect the continuous-time system ($\phi(t) = 0$), define the estimation error as

$$\mathbf{e}(k) \stackrel{def}{=} \mathbf{x_1}(k) - \hat{\mathbf{x}}(k)\,. \tag{18}$$

With the use of Lemma 8 the asymptotic boundedness of the output variable $w(k)$ will be proved in Theorem 8. This result will be used to state the asymptotic boundedness of the estimation error and at last an asymptotic threshold for the whole state vector will be found in Theorem 9.

Lemma 8. $\mathbf{\Delta}(k)$ is bounded by $\gamma\,|w(k)|$, where

$$\gamma = ||\mathbf{M}^{-1}||\, e^{||\mathbf{A}||}\, ||\mathbf{b}||\, \beta\, T_C \tag{19}$$

.

Proof. The proof is straightforward, indeed by (5) and Assumption 4

$$||\mathbf{\Delta}(k)|| < ||\mathbf{M}^{-1}||\, e^{||\mathbf{A}||}\, ||\mathbf{b}|| \int_{kT_C}^{(k+1)T_C} |d(\mathbf{x}(\sigma))|\, d\sigma = \\ = \gamma |w(k)|\,.$$

Theorem 8. Setting

$$u(k) = -\frac{\mathbf{g_{21}}'}{q_2}\hat{\mathbf{x}}(k) - \frac{g_{22}}{q_2}w(k)\,, \tag{20}$$

if

$$\beta < \frac{1}{||\mathbf{M}^{-1}||\, e^{||\mathbf{A}||}\, ||\mathbf{b}||\, T_C\, n}, \tag{21}$$

then the output variable $w(k)$ is asymptotically bounded.

Proof. The proof will be carried out in two steps: at first it is showed that the asyptotic bound of the estimation error depends only on the asymptotic bound of the residual signal; then the asymptotic bound of $w(k)$ is provided.

By (2), (7) and (16) the estimation error dynamics (18) are given by

$$\mathbf{e}(k+1) = \mathbf{G_{11}}\,\mathbf{e}(k) + \mathbf{\Delta_1}(k) - \mathbf{g_{12}}\,p(k)\,. \tag{22}$$

Since $(\mathbf{G_{11}})^{n-1} = \mathbf{0}_{(n-1)\times(n-1)}$ and $||\mathbf{G_{11}}|| = 1$, if $k \geq n-1$,

$$||\mathbf{e}(k)|| \leq \sum_{j=k-n+1}^{k-1} ||\mathbf{\Delta_1}(j) - \mathbf{g_{12}}\,p(j)||\,, \tag{23}$$

and by Lemma 8 it follows that for $k \geq n-1$

$$||\mathbf{e}(k)|| < \gamma \sum_{j=k-n+1}^{k-1} |w(j)| + (n-1)||\mathbf{g_{12}}||\,\rho\,. \tag{24}$$

From (7), (16) and (20) the dynamics of $w(k)$ are

$$w(k+1) = \mathbf{g_{21}}'\mathbf{e}(k) - g_{22}\,p(k) - p(k+1) + \Delta_2(k)\,. \quad (25)$$

Since by hypotesis $||\mathbf{g_{21}}|| = 1$, by Lemma 8 and (24), for every $k \geq n-1$

$$|w(k+1)| \leq ||\mathbf{e}(k)|| + (|g_{22}|+1)\rho + \gamma\,|w(k)| \leq$$
$$< \gamma \sum_{j=k-n+1}^{k} |w(j)| + \alpha\,, \quad (26)$$

where

$$\alpha = ((n-1)||\mathbf{g_{12}}|| + |g_{22}| + 1)\rho. \quad (27)$$

Define

$$\bar{w} = \max_{0\leq j\leq k} |w(j)|\,.$$

From equation (26)

$$|w(k+1)| \leq \gamma\, n\bar{w} + \alpha\,, \quad (28)$$

choosing β as in (21) one gets that $\gamma\, n < 1$ so for every instant k there is always a $k_1 < k$ such that

$$|w(k)| \leq \gamma\, n\,|w(k_1)| + \alpha\,. \quad (29)$$

By induction

$$|w(k)| \leq (\gamma\, n)^{k-k_1}\,|w(k_1)| + \frac{1-(\gamma\, n)^{k-k_1}}{1-\gamma\, n}\alpha$$

therefore

$$\limsup_{k\to\infty} |w(k)| \leq \frac{\alpha}{1-\gamma\, n}\,. \quad (30)$$

Theorem 9. Setting the input $u(k)$ as in (20), and β as in (21), the state vector $\mathbf{x}(k)$ is asymptotically bounded.

Proof. Since the bound of the estimation error found in the proof of Theorem 8 (24) depends only on the norm of the measurable term $\mathbf{w}$, the asymptotic boundedness of the estimation error (18) follows directly from Theorem 8, and in particular

$$\limsup_{k\to\infty} ||\mathbf{e}(k)|| \leq (n-1)\left(\frac{\gamma\alpha}{1-\gamma\, n} + ||\mathbf{g_{12}}||\,\rho\right) \quad (31)$$

where α and γ verify (27) and (19) respectively.

If the estimation state $\hat{\mathbf{x}}$ is asymptotically bounded, then the state vector $\mathbf{x}$ behaves accordingly due to (2). From (16), (17) and (20) the dynamics of the estimated vector are

$$\hat{\mathbf{x}}(k+1) = \mathbf{P}\hat{\mathbf{x}}(k) + \left(\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2} g_{22}\right) w(k)\,. \quad (32)$$

By Assumption 3 and Lemma 7, $\mathbf{P}$ is a Schur matrix, and due to Theorem 8 $w(k)$ is asymptotically bounded, so by induction

$$\limsup_{k\to\infty} ||\hat{\mathbf{x}}(k)|| = \left|\left|\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2} g_{22}\right|\right| \cdot$$
$$\cdot \limsup_{k\to\infty} \sum_{j=0}^{k-1} ||\mathbf{P}||^{k-1-j} |w(j)| \leq$$
$$\leq \frac{\left|\left|\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2} g_{22}\right|\right|}{1-||\mathbf{P}||} \frac{\alpha}{1-\gamma\, n}\,. \quad (33)$$

By definition $||\mathbf{x}(k)|| \leq ||\mathbf{x_1}(k)|| + |x_2(k)|$, from (7) $|x_2(k)| \leq |w(k)| + \rho$, so by Theorem 8 $x_2(k)$ is asymptotically bounded, in fact

$$\limsup_{k\to+\infty} |x_2(k)| \leq \limsup_{k\to+\infty} |w(k)| + \rho = \frac{\alpha}{1-\gamma\, n} + \rho \quad (34)$$

It is enough to show that $||\mathbf{x_1}(k)||$ is asymptotically bounded in order to complete the proof. From (**??**) and (20) the evolution of $\mathbf{x_1}(k)$ is

$$\mathbf{x_1}(k+1) = \mathbf{G_{11}}\mathbf{x_1}(k) + \left(\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2}\, g_{22}\right) w(k) +$$
$$- \mathbf{g_{12}}\, p(k) - \frac{\mathbf{q_1}}{q_2}\, \mathbf{g_{21}}'\hat{\mathbf{x}}(k) + \mathbf{\Delta_1}(k)$$

since $\mathbf{G_{11}}$ is a nilpotent matrix of degree $n-1$, by induction for every $k \geq n-1$

$$||\mathbf{x_1}(k)|| \leq \sum_{j=k-n+1}^{k-1} \left|\left|\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2}\, g_{22}\right|\right| |w(j)| +$$
$$+ \sum_{j=k-n+1}^{k-1} \frac{||\mathbf{q_1}||}{|q_2|}\, ||\hat{\mathbf{x}}(j)|| +$$
$$+ (n-1)||\mathbf{g_{12}}||\rho + \sum_{j=k-n+1}^{k-1} ||\mathbf{\Delta_1}(j)||\,.$$

By Lemma 8, (30) and (33),

$$\limsup_{k\to+\infty} ||\mathbf{x_1}(k)|| \leq \frac{(n-1)\,\alpha}{1-\gamma\, n}\left(\left|\left|\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2}\, g_{22}\right|\right| \cdot \right.$$
$$\left. \cdot\left(1 + \frac{||\mathbf{q_1}||}{|q_2|(1-||\mathbf{P}||)}\right) + \gamma\right) +$$
$$+ (n-1)||\mathbf{g_{12}}||\rho\,.$$

So the asymptotic bound of the state vector is

$$\limsup_{k\to+\infty} ||\mathbf{x}(k)|| \leq \frac{(n-1)\,\alpha}{1-\gamma\, n}\left(\left|\left|\mathbf{g_{12}} - \frac{\mathbf{q_1}}{q_2}\, g_{22}\right|\right| \cdot \right.$$
$$\left. \cdot\left(1 + \frac{||\mathbf{q_1}||}{|q_2|(1-||\mathbf{P}||)}\right) + \gamma\right) +$$
$$+ (n-1)||\mathbf{g_{12}}||\rho + \frac{\alpha}{1-\gamma\, n} - \rho\,.$$

Remark 9. As expected, all the asymptoticthresholds found depend on the known constant bound of the quantization error ρ.

4.2 Presence of actuator faults affecting the plant

Let's now suppose that an actuator fault $\phi(t)$ may affects the plant (1). By definition (18) the dynamics of the estimation error are

$$\mathbf{e}(k+1) = \mathbf{G_{11}}\mathbf{e}(k) + \mathbf{\Delta_1}(k) - \mathbf{g_{12}}\, p(k) + \mathbf{\Phi_1}(k)\,, \quad (35)$$

and by (7) the evolution of the measurable output variable is

$$w(k+1) = \mathbf{g_{21}}'\,\mathbf{e}(k) - g_{22}p(k) - p(k+1) +$$
$$+ \Delta_2(k) + \Phi_2(k)\,. \quad (36)$$

A test function r_f depending on the norm of the residual signal (16) will be defined to detect actuator faults.

Definition 10. Let the test function be

$$r_f(k) \stackrel{def}{=} \frac{|w(k)|}{\gamma\, n\,\overline{w}(k) + \alpha} \quad (37)$$

where γ and α verify (19) and (27) respectively, and

$$\overline{w}(k) = \max_{i=k-n,\ldots,k-1} |w(i)| .$$

Proposition 10. If $r(\bar{k}) > 1$ then an actuator fault has occured at a time $k \leq \bar{k}$.

Proof. The proof is straightforward. It simply consists in showing that using equations (7), (26) and (37) one has:

$$r_f(k) = \frac{|w(k)|}{\gamma\, n\, \overline{w}(k) + \alpha} \leq$$
$$\leq 1 + \frac{|\Phi_2(k-1)| + \sum_{i=k-n}^{k-2} |\mathbf{\Phi_1}(i)|}{\gamma\, n\, \overline{w}(k) + \alpha} .$$

So if no faults affect the continuous-time system (1), then $r_f < 1$.

Remark 11. The previous Proposition gives only a sufficient condition for detecting actuator faults. It may indeed happen that an actuator fault occurs but is "small" enough that the test of Proposition 10 fails. This would mean, however, that condition (26) still holds, therefore the bounded behavior of the output variable $w(k)$ has not been destroyed by the fault.

5. SIMULATION RESULTS

The theoretical development discussed in Section 4 is here supported by a simulation study with reference to an unstable (the set-up presented holds both when $\mathbf{A}$ is Hurwitz and when $\mathbf{A}$ is unstable), uncertain, continuous-time plant of the form (1) with:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1.1 \\ 1 & 0 & 0.2 \\ 0 & 1 & 0.8 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{f} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$
$$\mathbf{c} = [\, 0 \;\; 0 \;\; 1 \,]$$

The plant has been discretized with a sampling time $T_C = 10^{-3}$ s, and quantization has been assumed to produce a quantization error bounded by $\rho = 10^{-3}$. A disturbance term $d(t) = \beta\, x_2(t)\, \sin(t)$ has been supposed to perturb the continuous-time system, with $\beta = 0.1511$, and $\gamma = ||\mathbf{M}^{-1}||\, e^{||\mathbf{A}||}\, ||\mathbf{b}|| = 19.8576$. Assuming that the largest variation D which the variable $w(k)$ can undergo between two consecutive samples $w(k)$ and $w(k+1)$ with $k \in \mathbb{N}$ is equal to 2, it can be seen that the disturbance function considered is consistent with the Assumption 4, since $|d(t)| \leq \beta|w(k)| + \beta(\rho + D)$ in every interval of amplitude T_C.

Simulations have been performed with initial conditions $\mathbf{x}(0) = [0.5\ 0.5\ 0.5]'$. Results have been reported in Figures 1-6. Figures 1, 2 show the dynamics of the estimation error and of the test function (37) when no actuator faults affect the sampled-data system. The red lines reported in Fig.1 are the estimation error asymptotic bound (31), while in Fig. 2 the red threshold set at 1 is consistent with Proposition 10. Fig. 3 displays the evolution of the test function (37) when an abrupt fault $\phi_a(t) = 25$ (resp. an incipient fault $\phi_i(t) = 2t - 100$ in Fig. 4) affects the third component of the state vector $\mathbf{x}(t)$ of the sampled data system (1) at test function (37) when the abrupt fault $\phi_a(t) = 25$ and a larger abrupt fault $\phi_b(t) = 300$, respectively, affect the second component of the sampled data system (1) at time $t = 50$. These two pictures support what has been emphasized in Remark 11. Indeed, due to the sampled-data system under consideration, an abrupt fault of amplitude 25 affecting the plant (1) with a distribution vector $\mathbf{f} = [0\ 1\ 0]'$ is so "small" that the test of Proposition (10) fails.

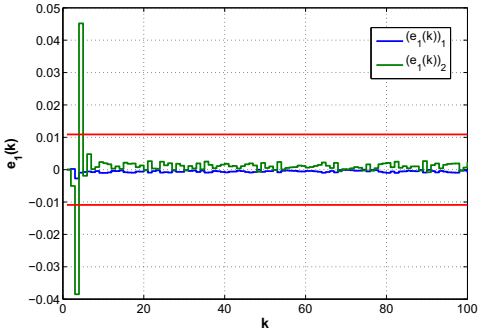


Fig.1 - Estimation error $\mathbf{e_1}(k)$ in the fault-free case.

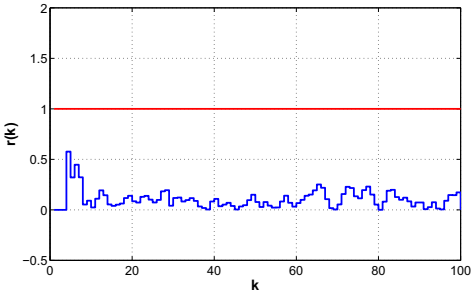


Fig.2 - Residual signal $r(k)$ when no actuator faults

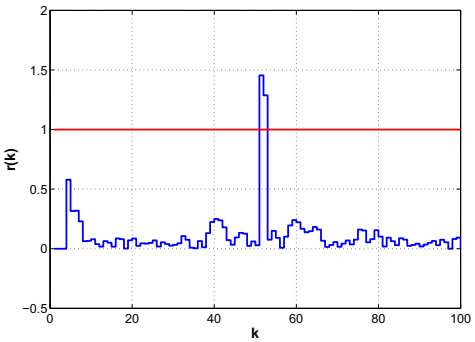


Fig.3 - Residual signal $r(k)$ when an abrupt fault $\phi_a(t) = 25$ affects the third component of the state vector $\mathbf{x}(t)$ at time $t = 50$.

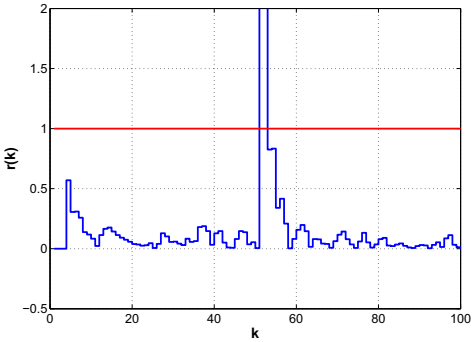


Fig.4 - Residual signal $r(k)$ when an incipient fault $\phi_i(t) = 2t - 100$ affects the third component of the state vector $\mathbf{x}(t)$ at time $t = 50$.

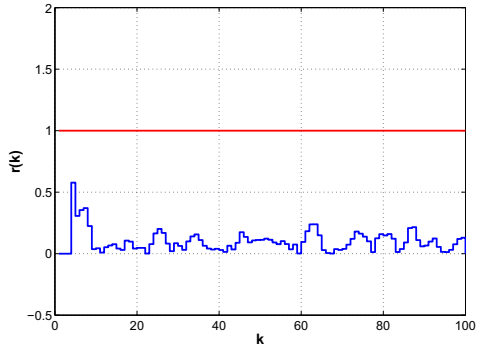


Fig.5 - Residual signal $r(k)$ when an abrupt fault $\phi_a(t) = 25$ affects the second component of the state vector $\mathbf{x}(t)$ at time $t = 50$.

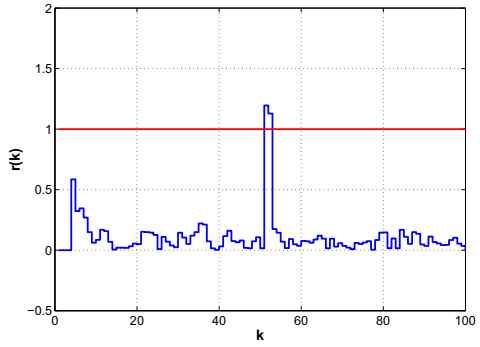


Fig.6 - Residual signal $r(k)$ when an abrupt fault $\phi_b(t) = 5000$ affects the second component of the state vector $\mathbf{x}(t)$ at time $t = 50$.

6. CONCLUSIONS

A robust (in the disturbance-decoupling sense) unknown-input observer design method for quantized uncertain sampled-data systems has been presented in this note. In the considered setup, the only signal available is the output variable which undergoes quantization. In order to overcome the strong structural conditions required to ensure the disturbance decoupling of the residual signals generated by the extension M. L. Corradini, R. Giambò, S. Pettinari (2009) of the full-order UIO proposed in J. Chen, R. J. Patton, H. Y. Zhang (1996), a new approach has been here proposed. The reduced-order filter presented, coupled with a controller, ensures the asymptotic boundedness of the estimation error, robustly with respect to disturbances affecting the continuous-time system, and the simultaneous boundedness of state variables of the plant.

REFERENCES

P. J. Antsaklis, A. N. Michel *Linear Systems.* Birkhauser, Boston, 2006.

C. Aubrun and D. Sauter and J. Yamé Fault diagnosis of networked control systems. *Int. J. Appl. Math. Comput. Sci.*, volume 18, no 4 pages 525–537, 2008.

T. Chen and B. Francis. *Optimal Sampled-Data Control Systems.* Springer, 1999.

J. Chen, R. J. Patton *Robust model-based fault diagnosis for dynamic systems.* Norwell, MS: Kluwer Academic Publishers.

J. Chen, R. J. Patton, H. Y. Zhang Design of unknown input observers and robust fault detection filters. *Int. J. Control*, volume 63, no 1 pages 85–105, 1996.

M. L. Corradini, A. Cristofaro, R. Giambò, and S. Pettinari. Robust fault detection filters for a class of MIMO uncertain sampled-data systems. In Proc. of Conference on Control and Fault-Tolerant Systems (SysTol10), Nice, France, 2010.

M. L. Corradini, R. Giambò, and S. Pettinari. Design of robust fault detection filters for plants with quantized information. In *Proc. 7th Workshop on Advanced Control and Diagnosis*, Zielona Gòra, PL, 2009.

P. M. Frank Fault diagnosis in dynamic system using analytical and knoledge based redendancy - a survey and some new results. *Automatica*, volume 26, pages 459–474, 1990.

P. M. Frank, X. Ding. Survey of robust residual generation and evaluation methods in observe-based fault detection systems. J. Process Control, pages 403424, 1997.

G. F. Franklin, J. D. Powell, A. Emami-Naeini *Feedback control of Dynamic Systems.* Prentice Hall: Englewood Cliffs, New Jersey.

J. J. Gertler Analytical redundancy mathods in fault detection and isolation. In *Proc. IFAC/IMACS Symp. SAFEPROCESS 91*, Baden-Baden, 1991.

J. J. Gertler *Fault detection and diagnosis in engineering systems.* New York: Marcel Dekker, 1998.

A. N. Michel, L.Hou, D. Liu *Stability of dynamical systems: Continuous, Discontinuous, and discete systems.* Birkhauser, Boston, 2008.

R. Isermann Supervision, fault-detection and fault diagnosis methods - an introduction. *Control Eng. Pract.*, volume 5, no 5 pages 639–657, 1997.

E.N. Rosenwasser, B.P. Lampe *Computer controlled systems - Analysis and design with process-oriented models.* Springer-Verlag, London, 2000.

A. S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, volume 12, no 6 pages 601611, 1976.

P. Zhang, S. X. Ding On fault detection in linear discrete-time, periodic, and sampled-data systems. *Journal of Control Science and Engineering*, 2008.

Appendix A. PROOF OF LEMMA 7

The proof consists of showing that the invariant zeros of (2) are exactly the eigenvalues of $\mathbf{P}$. By definition (see P. J. Antsaklis, A. N. Michel (2006)), the invariant zeros are values $z \in \mathbb{C}$ that let the Rosenbrock matrix $R(z)$ lose rank. In this case

$$R(z) = \begin{bmatrix} z\mathbf{I}_{n-1} - \mathbf{G_{11}} & -\mathbf{g_{12}} & -\mathbf{q_1} \\ -\mathbf{g_{21}} & z - g_{22} & -q_2 \\ \mathbf{0}_{1\times(n-1)} & 1 & 0 \end{bmatrix}$$

which loses rank if $det(R(z)) = 0$, that is if and only if

$$det \begin{bmatrix} z\mathbf{I}_{n-1} - \mathbf{G_{11}} & -\mathbf{q_1} \\ -\mathbf{g_{21}} & -q_2 \end{bmatrix} = 0. \qquad (A.1)$$

Due to the next matrix decomposition

$$\begin{bmatrix} z\mathbf{I}_{n-1} - \mathbf{G_{11}} & -\mathbf{q_1} \\ -\mathbf{g_{21}} & -q_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n-1} & \frac{\mathbf{q_1}}{q_2} \\ \mathbf{0} & 1 \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} z\mathbf{I}_{n-1} - \mathbf{G_{11}} + \frac{\mathbf{q_1}}{q_2}\mathbf{g_{21}} & 0 \\ 0 & -q_2 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n-1} & 0 \\ \frac{\mathbf{g_{21}}}{q_2} & 1 \end{bmatrix},$$

from (A.1) follows that $det(z\mathbf{I}_{n-1} - \mathbf{G_{11}} + \frac{\mathbf{q_1}}{g_2}\mathbf{g_{21}}) = 0$, so z is an eigenvalue of (17).

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Aircraft Sensor Fault Detection and Accommodation by Some Conventional Controllers

E. Kiyak*. F. Caliskan.**

**Anadolu University, Civil Aviation School, Eskisehir, 26470, Eskisehir, Turkey*
(Tel: +90 222 335 05 80 / 6879; e-mail: ekiyak@anadolu.edu.tr)
***Istanbul Technical University, Control Engineering, Istanbul, Turkey*
(e-mail: caliskanf@itu.edu.tr)

Abstract: The purpose of the study is to present an approach to detect and accommodate sensor faults using an unknown input observers based fault detection and isolation approach. After the detection and isolation of the faults, accommodation is very important in the systems such as an aircraft. In this study, the fault detection and isolation scheme is implemented for an aircraft model using observers and accommodation is provided by some conventional controllers. Accommodation results are compared with each other.

Keywords: Aircraft control, Controllers, Fault detection, Fault isolation, Fault tolerance

1. INTRODUCTION

In this paper, a fault detection and accommodation procedure is applied to a simple flight control model. Fault detection, isolation and accommodation techniques are widely used in fail safe control systems in which failures may cause hazardous incidents.

Hammouri et al. (1999) designed a residual generator for fault detection and isolation in nonlinear systems which are affine in the control signals and in the failure modes.

Kabore et al. (2000) proposed an approach based on decoupling techniques using differential geometry theory and observer synthesis for non-linear systems. A design procedure is provided for residual generation. The faults considered in the application are malfunctions of the feed pumps of both monomers and the initiator, and the presence of an inhibitor. A detailed construction of fault detection filters is presented, and their performances in the presence of parameter uncertainties are discussed through some simulations.

Kabore and Wang (2001) presented a set of algorithms for fault diagnosis and fault tolerant control strategy for affine nonlinear systems subjected to an unknown time-varying fault vector. The proposed algorithm is applied to a combined pH and consistency control system of a pilot paper machine, where simulations are performed to show the effectiveness of the proposed approach.

Hajiyev and Caliskan (2001) designed a Kalman filter for the effects of the sensor and actuator faults in the innovation process, and used a decision approach to isolate the sensor and actuator faults. The presented reconfigurable control algorithm is based on the Extended Kalman Filter (EKF). Reconfiguration procedure is executed by considering the identified control distribution matrix. In the simulations, the longitudinal dynamics of an aircraft control system is considered, and control reconfiguration is examined. A principal block diagram of a fault tolerant aircraft control system is proposed.

Hajiyev and Caliskan (2003) covered the combined fault diagnosis and reconfiguration in flight control systems.

Aykan et al. (2005) maintained safe flight and improved existing deicing (in-flight removal of ice) and anti-icing (prevention of ice accretion) systems under in-flight icing conditions. An offline artificial neural network is used as an identification technique. The Kalman filter is used to increase the state measurement's accuracy such that neural network training performance gets better. An aircraft linear model is simulated in time varying manner in terms of changing icing parameters in a system dynamic matrix and the obtained data are used in neural network training and testing.

Hajiyev and Caliskan (2005) addressed the flight control system's failures of sensor, actuator and control surface. The extended Kalman filter (EKF) was developed for nonlinear flight dynamic estimation of an F-16 fighter and the effects of the sensor and control surface/actuator failures in the innovation sequence of the designed EKF are investigated. A robust Kalman filter was used to isolate the control surface/actuator failures and sensor failures.

Amato et al. (2006) proposed a nonlinear Unknown Input Observer (UIO) to detect and isolate sensor faults in an aircraft. To guarantee robustness against nonlinearities and disturbances H_∞ theory is employed for the observer gain design.

Kiyak et al. (2008) addressed the flight control system's failures of sensor. They used residuals generated by an unknown input observer to detect fault and isolate the sensor failures in a VTOL dynamic model. Although there exist unknown inputs such as system non-linearities, noise and

disturbances, any single sensor fault could be detected and isolated correctly. They show this kind of observer is robust and flexible.

In this study, a fault detection and accommodation procedure is implemented in a simple roll flight control system. The fault is detected using an unknown input observer and accommodation is obtained through some controllers. Controllers are compared with each other in terms of performances.

2. FAULT DETECTION

2.1 Observers

An observer is a dynamical system whose state converges to the state of the plant (Hajiyev and Caliskan, 2003).

Consider a continuous linear time invariant steady space model of the system:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{1}$$

$x \in R^{nx1}$, $u \in R^{mx1}$, $y \in R^{nx1}$, $A \in R^{nxn}$, $B \in R^{nxm}$, and $C \in R^{nxn}$ represents state vector, input vector, sensor output, system coefficient matrix, input coefficient matrix, and output coefficient matrix, respectively.

The structure of the observer is described as:

$$\dot{z}(t) = Fz(t) + Gy(t) + Lu(t) \tag{2}$$

where $F \in R^{nxn}$ observer dynamics, $G \in R^{nxn}$ measurement distribution matrix, $L \in R^{nxm}$ control distribution matrix, and $z(t) \in R^{nx1}$ observation vector.

The error vector is given by:

$$e(t) = z(t) - Tx(t) \tag{3}$$

Using Equation (1) and (2), derivative of the error vector is obtained:

$$\dot{e}(t) = F(z(t) - Tx(t)) + (FT - TA + GC)x(t) + (L - TB)u(t) \tag{4}$$

Equations;

$$FT - TA + GC = 0 \tag{5}$$

$$L - TB = 0 \tag{6}$$

are satisfied, equation (4) can be written as:

$$\dot{e}(t) = Fe(t) \tag{7}$$

The solution of the Equation (7) is:

$$e(t) = e^{Ft} e(0) \tag{8}$$

If the matrix F is selected Hurwitz, the error approaches zero asymptotically:

$$\lim_{t \to \infty} e(t) = 0 \tag{9}$$

and it follows:

$$\lim_{t \to \infty} z(t) = \lim_{t \to \infty} Tx(t) \tag{10}$$

2.2 Unknown Input Observers

Consider a continuous linear time invariant steady space model of the system:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Ed(t) \\ y(t) &= Cx(t) \end{aligned} \tag{11}$$

d represents the unknown input vector and E represents the unknown input distribution matrix.

The structure of the unknown input observer is described as (Guan and Saif, 1991):

$$\begin{aligned} \dot{z}(t) &= Fz(t) + TBu(t) + Ky(t) \\ \hat{x}(t) &= z(t) + Hy(t) \end{aligned} \tag{12}$$

Here, $\hat{x}$ represents the estimated state vector.

The error vector is given by:

$$e(t) = x(t) - \hat{x}(t) \tag{13}$$

Using Equation (11) and (12), error vector is obtained:

$$\begin{aligned} e(t) &= x(t) - \hat{x}(t) = x(t) - z(t) - Hy(t) = \\ & x(t) - z(t) - HCx(t) = (I - HC)x(t) - z(t) \end{aligned} \tag{14}$$

Using Equation (14), derivative of the error vector is obtained:

$$\begin{aligned} \dot{e}(t) == & (A\text{-}HCA\text{-}K_1C)e(t)\text{-}[F\text{-}(A\text{-}HCA\text{-}K_1C)]z(t)\text{-} \\ & [K_2\text{-}(A\text{-}HCA\text{-}K_1C)H]y(t) \\ & \text{-}[T\text{-}(I\text{-}HC)]Bu(t)\text{-}(I - HC)Ed(t) \end{aligned} \tag{15}$$

If the following relations hold true;

$$(HC - I)E = 0 \tag{16}$$

$$T = I - HC \tag{17}$$

$$F = A - HCA - K_1C \tag{18}$$

$$K_2 = FH \tag{19}$$

$$K = K_1 + K_2 \tag{20}$$

then the derivative of the error vector (Equation (7)) will be $\dot{e}(t) = Fe(t)$ and, then the solution of the error vector is $e(t) = e^{Ft}e(0)$. If F is chosen as a Hurwitz matrix, the error goes to zero asymptotically. Hence, $\hat{x}$ converges to x.

3. FAULT DETECTION IN FLIGHT CONTROL SYSTEM

The equations of aircraft motion are obtained from Newton's second law by employing Taylor series expansion for multivariable functions to linear functions about the equilibrium points by considering the steady reference conditions. Using the steady space representation of the linear equations is useful for choosing the input vector which controls the surface's motions that affect the value of each state variable (Mclean, 1990). Generally, aircraft motions are classified as longitudinal and lateral motions. In this paper those motions are assumed to be decoupled. Although the aircraft model is very complex, here, a simple roll flight control system for more understanding will be used.

In Fig. 1, the block diagram of a simple roll flight control system is given (Nelson, 1998).

The system is composed of a comparator, a controller, aircraft roll dynamics, and a sensor to measure the airplane's roll angle.

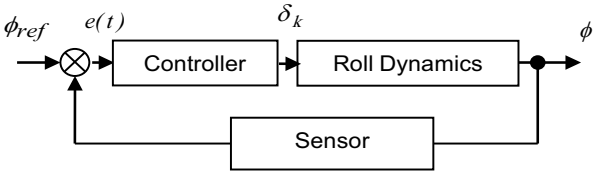


Fig 1. Block diagram of a simple roll flight control system.

Here, ϕ represents roll angle, e represents error vector, ϕ_{ref} represents desired roll angle.

Stability derivatives and aerodynamic characteristics of a small piston engine general aviation airplane are (Roskam, 2003):

$C_{l_{\delta_a}} = 0.229$, $C_{L_{\delta_r}} = 0.0147$, $C_{l_p} = -0.484$, $S = 174$ ft^2, $b = 36$ ft, $I_x = 948$ slugft^2, $Q = 49.6$ lb/ft^2, $u_0 = 220.1$ ft/s

From the numerical aerodynamic characteristics L_{δ_a} and L_p are calculated as:

$$L_{\delta_a} = \frac{QSbC_{l_{\delta_a}}}{I_x} = 75 \tag{21}$$

$$L_p = \frac{QSb^2C_{l_p}}{2I_x u_0} = -13 \tag{22}$$

The loop transfer function can be expressed as:

$$\frac{\phi}{e} = \frac{L_{\delta_a}K_p}{s(s - L_p)} = \frac{75K_p}{s(s + 13)} \tag{23}$$

The gain K_p equals 0.05 and ϕ_{ref} is the unit step input. Output of the system is changed after 50th second for sensor fault simulation. In this case, output of the system is obtained as in Fig. 2. The sampling time is 0.2 second.

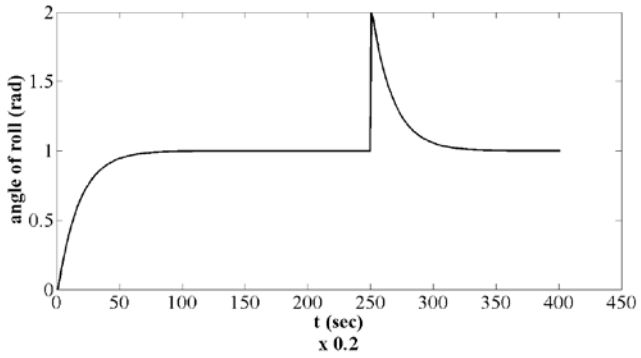


Fig 2. The output of the system. Sensor fault occurs at t=50 secs.

The error vector is obtained as in Fig. 3.

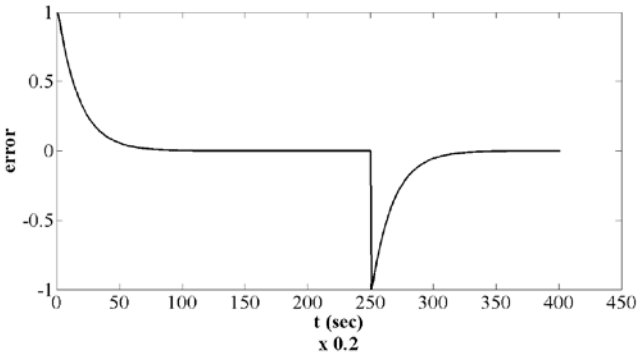


Fig 3. The error vector.

The matrices of the system are:

$$A = \begin{bmatrix} -13 & -3.75 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ C = \begin{bmatrix} 0 & 3.75 \end{bmatrix} \tag{24}$$

The control input is the deflection of the aileron as:

$$u = \delta_a \tag{25}$$

The state vector is defined as:

$$x = \begin{bmatrix} p \\ \phi \end{bmatrix} \tag{26}$$

p is roll rate and ϕ is roll angle. F and T are chosen as:

$$F = \begin{bmatrix} -10 & 0 \\ 0 & -10 \end{bmatrix},\ T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{27}$$

The matrices G and L of the observer are obtained as:

$$G = \begin{bmatrix} -1 \\ 2.67 \end{bmatrix},\ L = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{28}$$

Using parameters of the observer, the residuals are obtained as in Fig. 4. It is seen that after the 250th iteration (50 seconds), the residuals has increased and the fault is detected.

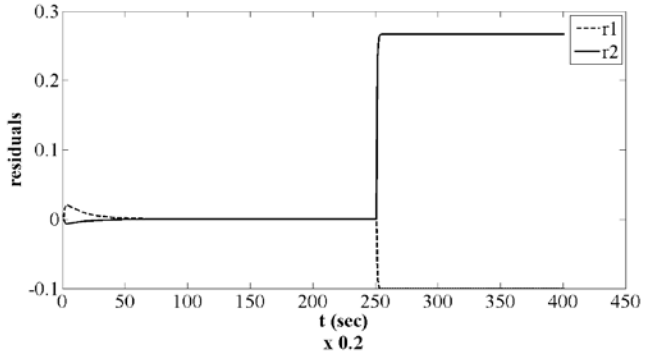


Fig. 4. The residuals.

4. FAULT TOLERANT CONTROL

After the fault detection and isolation, rapid accommodation is also important. Fault tolerant control can be accomplished in two ways: passive and active. Passive approaches make use of robust control techniques to ensure that a closed-loop system remains insensitive to certain faults using constant controller parameters and without use of on-line fault information. In active approaches, a new control system is redesigned using desirable properties of performance and robustness that were important in the original system, but with the reduced capability of the impaired system in mind (Patton, 1997).

In this study, first the fault is detected, and then the structure of the controller is changed for accommodation. P, PD and PI controllers are used and the performances are compared with each other. Rise time, settling time and steady state error are used as basic design criterion.

The aim of the controller of the closed loop control systems is to produce an output following a reference input. The controllers such as P, PD, PI and PID are widely used in practical applications with some variations depending on the plant or process structure. These controllers have some advantages and disadvantages. P type controller is known by its simplicity. The advantage of I type controller is that the output is proportional to the accumulated error. Thus, the error can be eliminated by the controller. The advantage of D type controller is that it will provide corrections before the error becomes large. P type controller's main disadvantage is that there may be a fix steady state error. The disadvantage of I controller is that the system is less stable due to the additional pole at the origin. The disadvantage of D controller is that if the error is constant it will not produce a control output.

Firstly, after the fault detection, a P controller is designed for a unity feedback flight control system. The closed loop transfer function is obtained as:

$$G(s) = \frac{75K_p}{s(s+13)} \tag{29}$$

The gain this point can be determined from the magnitude criteria as follows:

$$\frac{|K_p|\,75}{|s||s+13|} = 1 \tag{30}$$

where $s = -10$. A value of K_p is obtained as 0.4. Using this value, the output of the accommodated system is given by Fig. 5.

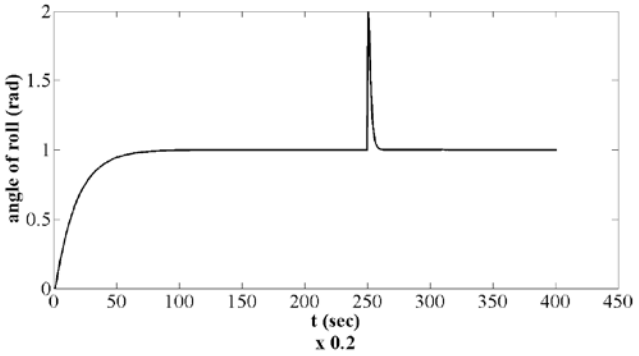


Fig 5. Output response of the feedforward plus feedback control system to a unit step disturbance K_p = 0.4.

Secondly, after the fault detection, a PD controller is designed and unity feedback is used. (Feedforward plus feedback control system). The feedforward transfer function is obtained as:

$$G(s) = \frac{75(K_p + K_d s)}{s(s+13)} \tag{31}$$

The closed loop transfer function can be shown to have the following form:

$$\frac{\phi(s)}{\phi_{ref}(s)} = \frac{75(K_p + K_d s)}{s^2 + (75K_d + 13)s + 75K_p} \tag{32}$$

Error of steady state is obtained as:

$$K_k = \lim_{s\to 0} G(s) = \infty,\ e_{ss} = \frac{1}{1+K_k} = 0 \qquad (33)$$

The system is suitable for unit step input.

On the other hand, the characteristic polynomial is obtained as:

$$s^2 + (75K_d + 13)s + 75K_p = 0 \qquad (34)$$

If $K_p = 1$ is chosen and damping ratio is 0.707, $K_d = -0.01$ is obtained. Using these values, the output of the accommodated system is given by Fig. 6.

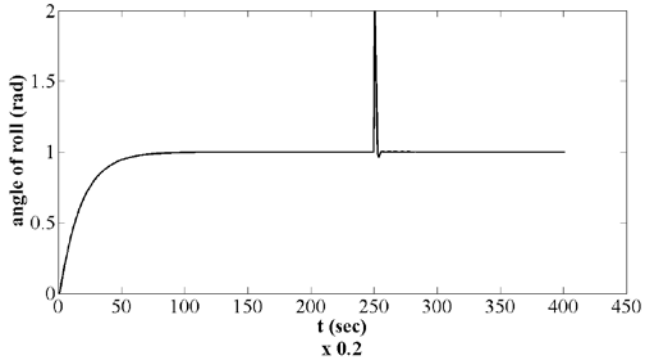


Fig 6. Output response of the feedforward plus feedback control system to a unit step disturbance K_p = 1, K_d = -0.01.

Finally, after the fault detection, a PI controller is designed and unity feedback is used. (Feedforward plus feedback control system). The feedforward transfer function is obtained as:

$$G(s) = \frac{75K_p(s + K_i/K_p)}{s^2(s+13)} \qquad (35)$$

Error of steady state is obtained as:

$$K_k = \lim_{s\to 0} G(s) = \infty,\ e_{ss} = \frac{1}{1+K_k} = 0 \qquad (36)$$

The system is suitable for unit step input.

On the other hand, the characteristic polynomial is obtained as:

$$s^3 + 13s^2 + 75K_p s + 75K_i = 0 \qquad (37)$$

If stability testing of Routh-Hurwitz is used, $K_p > 0.077K_i$ is obtained. Here, $K_i = 5$ and $K_p = 10$ are chosen. Using these values, the output of the accommodated system is given by Fig. 7.

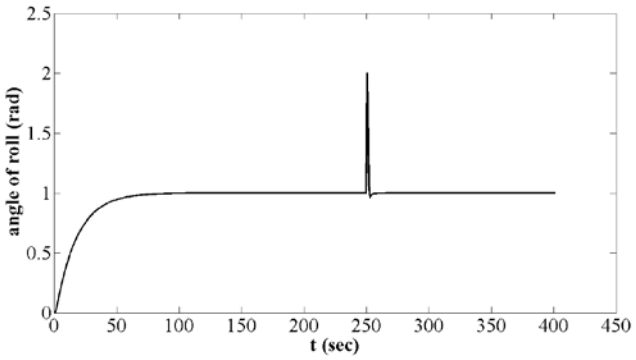


Fig 7. Output response of the feedforward plus feedback control system to a unit step disturbance K_p = 10, K_i = 5.

Performance of three controllers is given by Fig. 8.

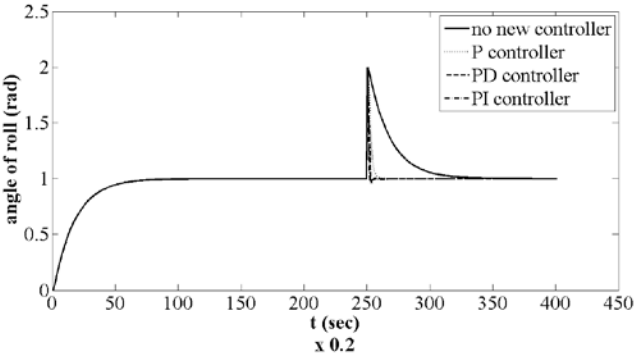


Fig 8. Output responses of the feedforward plus feedback control system to a unit step disturbance for different contollers.

5. CONCLUSION

In this study, fault detection and accommodation for sensor faults are implemented by using aircraft's flight control model. The fault detection and isolation is based on observers and accommodation is obtained by means of some conventional controllers. P, PD and PI controllers are used for accommodation.

Larger values typically mean faster response since the larger the error, the larger the proportional term compensation. An excessively large proportional gain will lead to process instability and oscillation. Larger K_p decreased the rise time and the steady state error in the simulations.

PD and PI controllers have very similar results. PD controller decreased the rise and settling times. PI controller decreased the rise and settling times, and significantly decreased the steady state error. A large fault effect results in more significant effects.

In this study, it was found that the PI controller is the more proper one because the rise time and settling time are minimum. Moreover, the steady state error goes to zero rapidly.

REFERENCES

Amato, F., Cosentino, C., Mattei, M., Paviglianiti, G. (2006). A direct/functional redundancy scheme for fault detection and isolation on an aircraft. *Aerospace Science and Technology*, Vol. 10, Issue: 4, pp. 338–345.

Aykan, R., Hajiyev, C., Caliskan, F. (2005). Kalman filter and neural network-based icing identification applied to A-340 aircraft dynamics. *Aircraft Engineering and Aerospace Technology*, Volume: 77, No. 1, pp. 23-33.

Guan, Y., Saif, M. (1991). A Novel Approach To The Design Of Unknown Input Observers. *IEEE Transactions on Automatic Control*, Volume: 36, Issue: 5, pp. 632-635.

Hajiyev, C., Caliskan, F. (2001). Integrated sensor/actuator FDI and reconfigurable control for fault-tolerant flight control system design. *The Aeronautical Journal*, Volume: 105, No. 1051, pp. 525-533.

Hajiyev, C., Caliskan, F. (2003). *Fault Diagnosis And Reconfiguration In Flight Control Systems*, Kluwer Academic Publishers, United Kingdom.

Hajiyev, C., Caliskan, F. (2005). Sensor and control surface/actuator failure detection and isolation applied to F-16 flight dynamic. *Aircraft Engineering and Aerospace Technology*, Vol. 77, No. 2, pp.152-160.

Hammouri, H. Kinnaert, M., El Yaagoubi E. H. (1999). Observer-based approach to fault detection and isolation for nonlinear systems. *IEEE Transactions on Automatic Control*, Vol. 44, Issue: 10, pp. 1879–1884.

Kabore, P., Othman, S., Mckenna, T. F., Hammouri, H. (2000). Observer-based fault diagnosis for a class of non-linear systems-Application to a free radical copolymerization reaction. *International Journal of Control*, Volume 73, Issue 9, pp. 787 – 803.

Kabore, P., Wang, H. (2001). Design of Fault Diagnosis Filters and Fault Tolerant Control for a Class of Nonlinear Systems. *IEEE Transactions On Automatic Control*, Vol. 46, No. 11, pp. 1805-1810.

Kiyak, E., Cetin, O., Kahvecioglu, A. (2008). Aircraft sensor fault detection based on unknown input observers. *Aircraft Engineering and Aerospace Technology*, Vol. 80, Issue: 5, pp. 545-548.

Mclean, D. (1990). *Automatic Flight Control Systems*, Prentice-Hall.

Nelson, R. C. (1998). *Flight Stability and Automatic Control*, McGraw Hill.

Patton, R. J. (1997). Fault-tolerant control: The 1997 situation. *In Proceedings of the 3rd IFAC symposium on fault detection, supervision and safety for technical processes*, pp. 1033–1055.

Roskam, J. (2003). *Airplane flight dynamics and automatic flight controls*, Roskam Aviation and Engineering Corp.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Performance Comparison of Different Types of Controllers for the Control of the Pitch Angle of an Aircraft

G. Iyibakanlar*. E. Kiyak.**

**Anadolu University, Civil Aviation School, Eskisehir, 26470, Eskisehir, Turkey*
(Tel: +90 222 335 05 80 / 6821; e-mail: giyibaka@anadolu.edu.tr)
***Anadolu University, Civil Aviation School, Eskisehir, 26470, Eskisehir, Turkey*
(Tel: +90 222 335 05 80 / 6879; e-mail: ekiyak@anadolu.edu.tr)

Abstract: Nowadays, many systems keep working properly thanks to realization of some control activities. The principle control operation is the designing of the controllers. In this study, the comparison of the P, PID and fuzzy controllers are realized by utilizing the pitch angle control of an airplane. As the governing performance attributes of the system comparison, the settling time, the steady state error and overshoot value has been taken into consideration. In classical controllers, the outputs having the preferred performance could be obtained by adjusting the gain; whereas, for the fuzzy controllers better values could be obtained by depending on the number of input and output functions and the number of the rules for the simulations.

Keywords: Control system design, Conventional control, Controllers, Flight control, Fuzzy logic

1. INTRODUCTION

Generally, the aim of close loop control is to control the outputs in a predetermined fashion by the feedback of the measured output. Figure 1 displays the block diagram of closed loop control system. Here, comparator compares the desired input value and output value measured by the sensor and produces an error signal (e(t)). Controller uses this error signal as the input and produces a decision signal depending on its own control type. Actuator is a dynamic component that uses this decision signal and produces a correction signal so as to minimize the error signal. It is clear that sensor both measures the value obtained in system output and converts it to a different signal types when needed. Thus, desired input is obtained in the output (Yuksel, 2001).

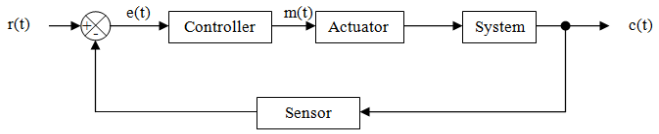


Fig 1. Closed-loop control system block diagram

The first thing necessary to design an automatic control system is a mathematical system suitable for the process or the whole system. The mathematical model of the controller can be developed and applied later on.

Automatic control system design can be realized according to two different criteria. According to the design based on "time domain", there are two important situations to consider in a system. The first one is "transient response" which depends on the stability of the system. A system is stable if every bounded input produces a bounded output. Transient response can be obtained by calculating rise time, overshoot rate and the time needed to reach steady-state situation. The second important situation for a system is steady-state response, which is determined by measuring steady-state error (Kuo, 2002). Figure 2 displays these concepts on unit step response of a second-degree closed loop system. It is very difficult to make a valid design for time domain for the systems higher than second-degree formations.

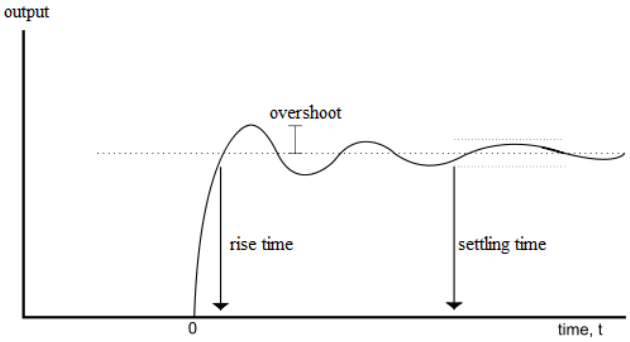


Fig 2. Step response of a second-order system

For systems with higher than second degree, frequency domain-based design is preferred, which utilizes certain methods such as Bode Diagram, Nyquist diagram and Nichols Chart (Kuo, 2002).

As for the use of PID in aviation, we can see that it was used in the algorithm of flight control system of a helicopter. It is clear that desired success level is achieved through PID controller regardless of the necessity to know the certain mathematical model of the system (Musial et al., 2010).

PID controllers were used in an application using an unmanned helicopter in which PID controller was used for

position control and PI controller for speed control (Frost et al., 2000).

In order to determine PID parameters, integral square error criterion method was used. Thanks to this PID controller used in longitudinal dynamic model of unmanned air vehicle, a fast transition to steady-state and a bit overshoot have been observed when a certain input is considered (Turkoglu, 2008).

Single input – single output and multi-layered PID controllers as well as an estimate controller that minimizes cost function were used in order to complete a mission given to an unmanned air vehicle (Kim and Shim, 2003).

In this method called "Convex Combination Method", Liu suggests designing experimental observations to develop performance criteria for various controllers. Later, if desired performance cannot be achieved through simple controllers, a new solution is searched through the combination of those. This method was realized first by choosing a PI controller for performance criteria regarding acceptable overshoot, quick response and small steady-state error in a pitch control system of an airplane and later by utilizing a convex PID controller together with this particular controller (Liu, 2003).

When classical controllers do not suffice, fuzzy logic controller, neural network-based controller and genetic algorithm-based controller can also be used. In a study that used fuzzy controller providing stability during the whole flight, it was concluded that fuzzy logic controller's parameters (membership functions, fuzzy rules etc.) might be improved (Gonsalves and Zacharias, 1994). It was claimed that a particular flight mission was completed successfully by a restructuring process conducted by fuzzy controller for an elevator breakdown in flight control system (Copeland and Rattan, 1994). Similarly, it was claimed that simple missions can be completed by a microprocessor that uses fuzzy controllers in unmanned air vehicles (Bickraj et al., 2006).

In the current study, P, PID and Fuzzy controllers are compared for the pitch angel of an airplane. As the distinctive performance characteristics enabling comparisons, settling time, steady-state error and overshoot were determined.

2. CONTROLLERS

The most important phase of designing a control system is to determine the structure and the components of the controller. As for the control structure, serial, feedback, or various combinations; such as state feedback, serial feedback and forward feedback can be used. PID controllers are often used in the structure preferred. In addition to classical controllers, certain advanced control algorithm designs such as fuzzy logic, neural networks and genetic algorithm can also be used.

2.1 Classical controllers

The most commonly used control systems in industry is PID type controllers. In such a control system, control process is carried out in three different ways. The goal here is to keep constant the predetermined values regarding overshoot time, settling time and steady-state error as well as to ensure system stability (Yuksel, 2001).

PID control is a control effect which combines the advantages of these three basic control effects into one single system. Integral effect reduces a possible constant steady-state error in a system to zero. In addition, derivational effect increases quick response for the stability of the system according to whether only PI control effect is used or not. Accordingly, PID control component provides a quick response which has zero steady-state error in the system.

PID control system is more complex and more expensive compared to other similar systems. Here, a desired control can be achieved by making appropriate settings regarding K_p , K_i and K_d parameters. If these coefficients are not suitable, we cannot make use of the advantages of PID control. All three control parts are explained below in detail:

Proportional term:

The proportional response can be adjusted by multiplying the error by a constant K_p , called the proportional gain.

$$P(t) = K_p e(t) \tag{1}$$

Integral term:
The magnitude of the contribution of the integral term to the overall control action is determined by the integral gain, K_i .

$$I(t) = K_i \int_0^t e(t)dt \tag{2}$$

Derivative term:

The magnitude of the contribution of the derivative term to the overall control action is termed the derivative gain, K_d .

$$D(t) = K_d \frac{de(t)}{dt} \tag{3}$$

The effects mentioned above are formulized as *m(t)* control signal as a whole as follows:

$$m(t) = K_p e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt} \tag{4}$$

Basically, when a dynamic structure needs controlling, PI controllers are suitable for this process, which can be explained with first degree differential equations. For instance, such controllers can be used to control water level in a water tank. PID controllers, on the other hand, are more appropriate for the situations when dynamic structure is a

secondary issue, such as the times when friction matters. As the system gets more complex, simple PID-type controllers may not be sufficient as well.

2.2 Supervision algorithms

Fuzzy logic

Fuzzy logic controller is a control type in which verbal and intuitional nature of human beings is utilized. This type of controllers consists of fuzzification, fuzzy rules and defuzzification units.

Fuzzification unit is the first unit of fuzzy process system. The knowledge received as the input for the unit in the forms of certain or feedback results become fuzzy after various changes. In other words; each piece of information is assigned a membership value, and verbalized, and later sent to fuzzy rules. The information that reaches fuzzy rules is combined with rule process. Data is already available in database in the unit. The logical suggestions mentioned here can be developed through numerical values depending on the structure of the problem as well. In the last phase, the results obtained through the use of logical decision suggestions suitable for the structure of the problem are sent to defuzzification unit. Finally, one more scale change is made for fuzzy cluster relations that are sent to defuzzification unit and each of these fuzzy pieces of information is converted to real numbers (Yen et al., 1995; Chen and Pham, 2001).

With the emergence of microcomputers, the use of such controllers also expanded, especially for the applications regarding the systems where mathematical model is not applied appropriately (Yuksel, 2001).

Neural networks

Neural Networks consist of information processing centers, called neurons, which are intensely connected to each other and work in harmony in order to imitate human brain. In fact, process units are like a transfer equation. They receive information, initiate transactions by applying transfer function and produce an output. How information will be processed by a structure highly depends on transfer function, how it is connected to other networks and its own synaptic weights.

Neural network is formed for a specific purpose and it learns through experiences just like human beings do. Neural Networks change their own structures and weights due to repeated inputs. Neural Networks can easily adapt itself to new situations just like the nervous systems of living creatures do. In other words, its structure might change according to internal and external stimulants and learning occurs accordingly. Connection weights are also taken into consideration during decision making phase. Although process units seem to be functioning alone, in fact a lot of neural networks operate simultaneously and display a “distributed and parallel computing” example (Neural networks, 2010).

Genetic algorithm

Producing practical and easy solutions for a variety of problems, simple genetic algorithm is the combination of three genetic processes; namely reproduction, crossover and mutation. These processes, which continue throughout a generation, end when the difference between optimal values obtained for the generations in maximization, when minimization problem is zero or when these values approximate to a certain predetermined value. In addition, genetic algorithms can be stopped after being repeated as many times as the certain number of generations that was determined at the beginning of the program. The higher the number is, the value to be obtained is more likely to be the optimal solution to the function (Isık, 2006).

3. COMPARISON OF CONTROLLERS FOR THE CONTROL OF THE PITCH ANGLE OF AN AIRCRAFT

The automatic pilot block diagram for a pitch angle control system of an aircraft is as follows (Nelson, 1998):

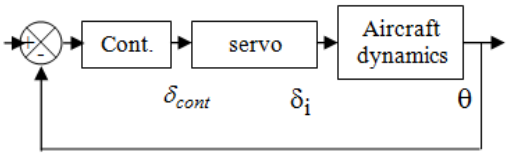


Fig 3. Pitch angle control system

The aircraft dynamics for a four-engine airliner with large fuselage with regards to a system given in Figure 3 is $\frac{\theta}{\delta_i} = \frac{-1.16}{s^2 + 0.339s + 0.75}$ and servo dynamic is $\frac{\delta_i}{\delta_{den}} = \frac{-1}{s+10}$. The forward loop transfer function of the system is (Mclean, 1990):

$$\frac{\theta}{\delta_{den}} = \frac{1.16}{(s+10)(s+0.17+0.85i)(s+0.17-0.85i)} \quad (5)$$

For this particular system, P and PID type controllers, as classic controllers, and fuzzy logic controller, as advanced control algorithm, will be used to complete the design, and later the results obtained will be compared.

The root locus of the system was obtained as shown in Figure 4. Here, horizontal axis represents real axis of the function in focus and vertical axis, the imaginary axis.

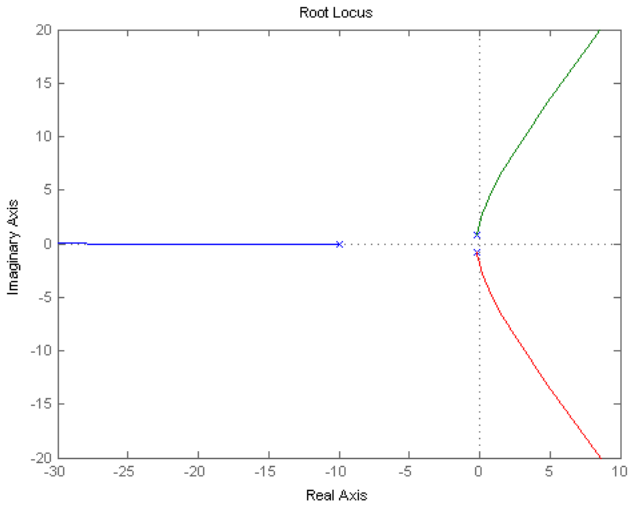


Fig 4. System root locus

With the help of root locus drawn, K gain for a particular point on root locus can be calculated through magnitude criteria.

Table 1 displays the gains for the controllers obtained by using Ziegler-Nichols method.

Table 1 Controller gains for P, PI and PID (Nelson, 1998)

Controller	K_p	K_i	K_d
P	$0.5K_{p_u}$		
PI	$0.45K_{p_u}$	$\frac{0.45K_{p_u}}{0.83T_u}$	
PID	$0.6K_{p_u}$	$\frac{0.6K_{p_u}}{0.5T_u}$	$0.6K_{p_u}(0.125T_u)$

If integral and derivation controller gains are assumed to be "0", forward loop transfer function is:

$$G(s)=\frac{1.16K_p}{(s+10)(s^2+0.339s+0.75)} \quad (6)$$

When Figure 4 is examined, it is seen that cut points of imaginary axis are $\mp 2.01i$ In order to obtain K_{p_u} in Table 1 following magnitude criteria can be applied:

$$\frac{|1.16|K_{p_u}}{|s+10||s+0.17+0.85i||s+0.17-0.85i|}=1 \quad (7)$$

If $s=2.0li$ is assumed for equation (7),

$$K_{p_u}=29.53 \quad (8)$$

is obtained. On the other hand, T_u is:

$$T_u=\frac{2\pi}{\omega}=\frac{2\pi}{2.01}=3.14 \quad (9)$$

According to values above, the following calculations were made: $K_p=14.77$ for P type controller; and $K_p=17.72$, $K_i=11.32$ and $K_d=6.93$ for PID type controller. The unit step response obtained by using these gain values are displayed in Figure 5.

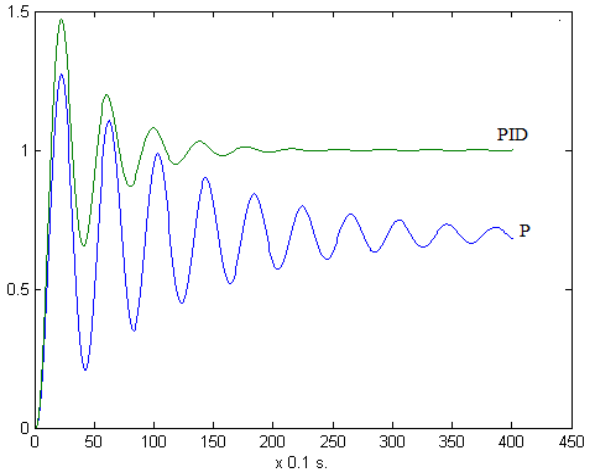


Fig 5. Step response of P and PID Controllers

Steady-state error for the unit step input for a system with P-type controller can be calculated analytically:

$$e_{ss}=\frac{1}{1+K_k} \quad (10)$$

Here K_k is position error coefficient, which can be obtained as follows (Yuksel, 2001):

$$K_k=\lim_{s\to 0}\frac{1.16K_p}{s^3+10.339s^2+4.14s+7.5} \quad (11)$$

The result here is $e_{ss}=0.3$.

Due to the integral effect in PID-type controller, $K_k=\infty$ and $e_{ss}=0$. When Figure 5 is examined, it is seen that steady-state error for P-type controller is quite high (approximately 30 % of the input) and settling time is quite long.

Settling time for PID-type controller, however, is relatively shorter and steady-state error is "zero".

The error in the system, the change in the error and membership functions belonging to the output were defined in Figure 6, Figure 7 and Figure 8. Intuitional method was used while determining membership functions.

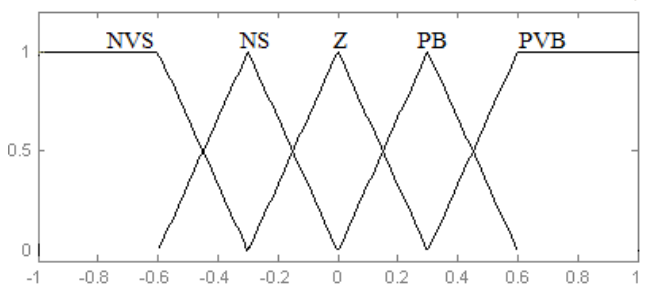


Fig 6. Membership functions for the error

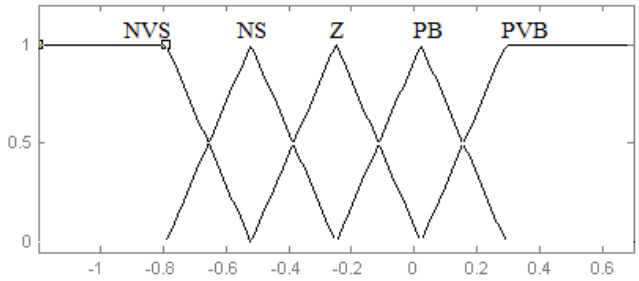


Fig 7. Membership functions for the derivative error

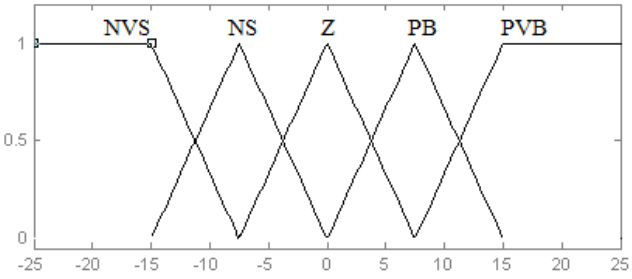


Fig 8. Membership functions for the output

Fuzzy rules for the system were formed as shown in Table 2.

Table 2 Rule base table for the system

$\dot{e}$ \ e	NVS	NS	Z	PB	PVB
NVS	NVS	NVS	NS	NS	Z
NS	NVS	NS	NS	Z	PB
Z	NS	NS	Z	PB	PB
PB	NS	Z	PB	PB	PVB
PVB	Z	PB	PB	PVB	PVB

Unit step response of the system was obtained as in Figure 9 by using the fuzzy rules in Table 2 and the membership functions in Figure 6-8.

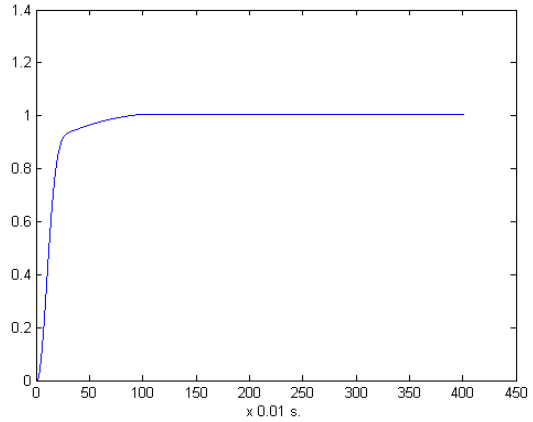


Fig 9. Step response of the fuzzy logic controller

When Figure 9 is examined, it is observed that the settling time for the system with fuzzy controller is shorter than the system with P and PID controller given in Figure 5. As for overshoot, it is seen that the system using fuzzy controller has better performance than other two controller types.

5. CONCLUSIONS

In this study, the designs for pitch angle control, which are quite significant for flight control systems, were examined by using both classic P and PID controllers and fuzzy controllers and the results were compared accordingly.

It was found that settling time obtained through P-type controller is considerably longer and steady-state error is higher.

In addition, settling time for the system with PID controller is relatively shorter than that of the system with P- type controller. A well-designed PID controller reflects together the advantageous characteristics of systems with PI and PD controllers that are designed separately.

Similarly, settling time for a system with fuzzy logic controller is shorter than those of the systems with P and PID controller. As for overshoot, the system with fuzzy logic controller has a better performance than those with P and PID controllers. A well-designed fuzzy controller can produce an output with a better performance. This study did not use unknown inputs occurring due to meteorological conditions and failures in aircraft dynamics. The dynamic model used in the study is the non-linear model linearized through Taylor series. When dynamic model is non-linear, desired performance cannot be achieved through a simple controller (classic and linear PID). In this situation, non-linear PID design should be preferred.

In real life, many systems have a nonlinear model. The solution here might be the linearization of nonlinear equations under certain conditions. When linearization is not applied, on the other hand, it is not possible to control the system by using a basic controller. The best solution in such situations might be the use of higher level control algorithms. Fuzzy logic and artificial neural networks controllers and

nonlinear controllers can be considered for the further studies on the topic of the current study.

REFERENCES

Bickraj, K., Yenilmez, A., Li, M., Tansel, İ. N. (2006). Fuzzy Logic Based Integrated Controller for Unmanned Aerial Vehicles, Conference on Recent Advances in Robotics, FCRAR, Florida.

Chen, G., Pham, T. T. (2001). Introduction to Fuzzy Sets, Fuzzy Logic, and Fuzzy Control Systems, CRC Press, Florida.

Copeland, R. P., Rattan, K. S. (1994). A Fuzzy Logic Supervisor for Reconfigurable Flight Control Systems. IEEE.

Frost, C. R., Tischler, M. B., Bielefield, M., La Montagne, T. (2000). Design and Test of Flight Control Laws for the Kaman Burro Unmanned Aerial Vehicle, American Institute of Aeronautics and Astronautics.

Gonsalves, P. G., Zacharias, G. L. (1994). Fuzzy Logic Gain Scheduling for Flight Control. IEEE.

http://e-bergi.com/2008/Subat/Yapay-Sinir-Aglari

http://pdv.cs.tu-berlin.de/MARVIN/

Işık, Y. (2006). Flight Control System Design Using Genetic Fuzzy Control, PhD Thesis, Anadolu University, Eskisehir.

Kim, H. J., Shim, D. H. (2003). A flight control system for aerial robots: algorithms and experiments. Control Engineering Practice 11.

Kuo, B. C. (2002). Automatic Control System, Literatur Publishers.

Liu, H. T. (2003). PID Type Control for Multiple Performance: A Flight Control Study, Proceedings of the American Control Conference, Denver, Colorado.

Mclean, D. (1990). Automatic Flight Control Systems, Prentice Hall.

Musial, M., Brandenburg, U. W., Hommel, G. (2010). Development of a Flight Control Algorithm for the Autonomously Flying Robot MARVIN.

Nelson, R. C. (1998). Flight Stability and Automatic Control, McGraw-Hill.

Neural networks. (2010).

Turkoglu, K., Ozdemir, U., Nikbay, M., Jafarov, E. M. (2008). PID Parameter Optimization of an UAV Longitudinal Flight Control System, Proceedings of World Academy of Science, Engineering and Technology, Volume 35.

Yen, J., Langari, R., Zadeh, L. A. (1995). Industrial Applications of Fuzzy Logic and Intelligent Systems. IEEE Press, New York.

Yuksel, I. (2001). *Automatic Control System Dynamics*, Uludag University.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fault detection and estimation in networked control systems ⋆

I. Peñarrocha * **and R. Sanchis** *

* *Departament d'Enginyeria de Sistemes Industrials i Disseny, Universitat Jaume I de Castelló, Spain; e-mail: ipenarro@esid.uji.es.*

Abstract:
In this paper, the fault detection and estimation in discrete-time systems whose output measurements from different sensors are acquired through a network is addressed. An observer that estimates the states and the faults is used in which the updating gain depends on the sensors availability. The process is assumed to be affected by disturbances of known bounds, and the proposed approach allows to achieve a compromise between the time response of the fault detector and the disturbances attenuation (related to the faults detection threshold). Two strategies are proposed: in the first one, the threshold is fixed, and the response time is minimized. In the second one, the response time is fixed, and the threshold is minimized. In both cases, the observer design is based on $\mathcal{H}_\infty$ norms minimization via LMI.

Keywords: Fault detection and isolation; Observer-based fault detection; Unknown input observer; LMI; Networked control systems; Threshold; False-alarm rate;

1. INTRODUCTION

The model based fault detection has been widely studied in the literature. In Chen and Patton [1999] a good survey of the different approaches can be found. One well known strategy consists of using an observer to estimate some process variables (typically states or outputs), and then to build a residual from the estimation error, whose evaluation leads to the fault detection, isolation and/or identification. Another approach is based on using a filter to estimate the fault signal. The estimated fault is then evaluated to detect, isolate or identify faults. One of the approaches to design both types of estimators is based on the use of norms of the transfer functions from disturbances and faults to the estimation error (see Rank and Niemann [1999]).

With respect the residual based approach, the most widely used norms are the $\mathcal{H}_\infty$ from disturbances to residual, and the $\mathcal{H}_-$ norm from faults to residual. A minimization of the quotient or the squared difference of those norms is usually proposed for the design, leading after complex manipulations to a set of LMI that must be solved iteratively. Some examples of this approach are Ding et al. [2000], Henry and Zolghadri [2005], Hou and Patton [1996], Wang et al. [2007], Zhong et al. [2003]. In Henry and Zolghadri [2006] this approach is compared to the fault estimation one, concluding that both approaches can reach a similar performance.

With respect the fault estimation approach, it is based on the use of the $\mathcal{H}_\infty$ norms from the disturbance and faults to the fault estimation error, leading after relatively simple manipulations to a LMI problem. In Henry and Zolghadri [2006] a fault estimation filter is proposed, where the design is based on imposing those norms to be lower than two constants. However, no suggestions are made in how to select those constants (in the example one of the constants is fixed to an arbitrary value, and the other one is minimized). In Gao et al. [2008] a fault estimator filter is proposed, and for the design, only the norm of the fault estimation error with respect the disturbance is minimized. This leads in general to a very slow fault estimation dynamics (in order to filter out the disturbances). Furthermore, the process input is not considered in the fault estimation filter equation, leading to changes in the estimation due to changes in process input.

On the other hand, in many practical applications the measurements are not available in a regular periodic basis due to, for instance, the use of a network that transmits the information of the sensors with data dropout, or because of the use of slow sensors or destructive sample analyzing devices whose measurements are scarcely available in time. This situation has recently been dealt with in some works, but assuming a periodic sampling framework, as in Izadi et al. [2005], Li et al. [2005], Wang et al. [2008]. Other works, as Gao et al. [2008], He and Zhou [2008], Zhang et al. [2004] have dealt with the irregular measurement case due to network constraints from a stochastic perspective, assuming the data availability follows a stochastic model. However, these works assume that the outputs are all available (or not) at the same time, and propose a time invariant fault detection filter that is made robust to the missing measurements probability distribution , leading to a conservative design.

In the present paper, a time variant fault estimator is proposed, that uses a simple extended model of the process to build an observer that includes the fault signals. The estimator gain depends on the availability of measure-

⋆ This work was supported by CICYT project number DPI2008-06731-C02-02/DPI

ments from the different sensors, that are assumed to be connected to the network at different nodes (and hence each sensor measurement could be available at different instants). The estimator is based on the use of an integrator in the fault equation, leading to a simpler strategy than the one considered in Henry and Zolghadri [2006]. The known process input is taken into account and hence its effect is fully compensated (for perfect model assumed), as a difference with Gao et al. [2008]. The network communication constraint between the sensors and the fault detection estimator is modelled by a simple probability of successful data transmission every period, leading to a finite set of possible scarce measurement scenarios, defined by the number of periods between consecutive measurements and the available sensors, with their associated probability of occurrence. The estimator gains are then defined as a function of those scenarios, guaranteing the fulfilment of the required expected performance by means of the solution of a set of LMI. Besides, the norm of the fault estimation error with respect disturbances as well as with respect the faults are considered, leading to two strategies to design the fault estimator if the bounds of disturbances norms are assumed to be known. In the first strategy, the threshold of the minimum detectable fault is fixed and then the response time of the fault estimator is minimized. In the second strategy, the response time of the fault estimator is fixed, and the fault detection threshold is minimized. Both approaches allow to achieve a desired compromise between disturbance attenuation (and hence minimum detectable fault) and fault detector time response. In both cases, a convex LMI minimization problem is solved. The layout of the paper is as follows: section 2 describes the problem, including the process model, the network operation assumed (and the resulting measurement availability pattern) and the proposed observer equation, in section 3 the estimation error dynamics is obtained, section 4 describes the design of the observer based on $\mathcal{H}_\infty$ optimization via LMI, section 5 shows an illustrative example, and in section 6 the conclusions are summarized.

2. PROBLEM STATEMENT

The ZOH discrete equivalent of the continuous linear time invariant process is assumed to be described by the equation

$$\boldsymbol{x}[t+1] = \boldsymbol{A}\,\boldsymbol{x}[t] + \boldsymbol{B}_u\,\boldsymbol{u}[t] + \boldsymbol{B}_w\,\boldsymbol{w}[t] + \boldsymbol{B}_f\,\boldsymbol{f}[t], \quad \text{(1a)}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^{n_u}$ are the known (or measured) inputs, $w \in \mathbb{R}^{n_w}$ is the state disturbance, $\boldsymbol{f} \in \mathbb{R}^{n_f}$ is the fault vector, and t is control period time counter. Each sensor, or group of sensors, is assumed to be connected to a different network node. It takes a measurement every period t and tries to send it to the controller and fault detection estimator node. If the measurement can not be sent during the control period t it is discarded (lost) and a new measurement is taken. A constant probability, β_i, is assumed for the event of the sensor i transmitting a measurement to the controller during one control period. In the case of successful transmission, a delay lower than one control period is assumed. As a consequence of the assumed network operation, the measurements are only available at some periods $t = t_k$ (that will depend on the probabilities β_i). The measurements are also assumed to be affected by noise and possible faults, hence the measurement equation is

$$m_{i,k} = \boldsymbol{c}_i\,\boldsymbol{x}_k + \boldsymbol{d}_i\,\boldsymbol{u}_k + v_{i,k} + \boldsymbol{h}_i\,\boldsymbol{f}_k, \;\; i = 1,\ldots,n_m \quad \text{(1b)}$$

where $m_{i,k}$ is the measurement of the i-th sensor at the k-th sampling instant (not all sensors are available simultaneously) $v_{i,k}$ is the i-th sensor noise, n_m is the number of sensors, and $\boldsymbol{x}_k$ and $\boldsymbol{u}_k$ represent the state and the input at instant $t = t_k$, i.e. $\boldsymbol{x}_k = \boldsymbol{x}[t_k]$, $\boldsymbol{u}_k = \boldsymbol{u}[t_k]$ (subindex $k \in \mathbb{N}$ is used to represent the value of a signal at a scarce sampling instant $t = t_k$). The system (1) is assumed to be detectable.

The values of different sensors m_i are received at different sampling instants, but at least one element is assumed to be available at instant $t = t_k$. The number of control input periods from t_{k-1} to t_k is denoted with $N_k = t_k - t_{k-1}$ and, therefore, $t_k = \sum_{i=1}^{k} N_i$ represents the instant in which the t-th input update occurs and the k-th sample with the values of the available sensors is received.

Let us define the sensor availability factor $\alpha_{i,k}$ of the i-th sensor for every sampling instant $t = t_k$ as

$$\alpha_{i,k} = \begin{cases} 1, \text{ if } m_i \text{ is received at } t = t_k, \\ 0, \text{ if } m_i \text{ is not received at } t = t_k, \end{cases}$$

Let us now define the availability matrix as

$$\boldsymbol{\alpha}_k = \mathrm{diag}\{\alpha_{1,k},\ldots,\alpha_{n_m,k}\}. \quad (2)$$

If all measurements are received at instant $t = t_k$, then $\boldsymbol{\alpha}_k = \boldsymbol{I}$. Depending on the measurements successful transmission pattern, there can be different values for matrix $\boldsymbol{\alpha}_k$ and they are assumed to belong to a known set

$$\boldsymbol{\alpha}_k \in \Xi = \{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r\}. \quad (3)$$

In the general case, any combination of available sensor measurements is possible, leading to $r = 2^{n_m} - 1$. A new parameter s_k is introduced to define the sampling scenario as the combination $(N_k, \boldsymbol{\alpha}_k)$ at the time of the k-th sampling reception. The parameters N_k and $\boldsymbol{\alpha}_k$ can now be written as a function of s_k: $N_k = N(s_k)$, and $\boldsymbol{\alpha}_k = \boldsymbol{\alpha}(s_k)$.

The controller's probability of receiving an output measurement from sensor i in one period, β_i, can be expressed in terms of the output availability factor as:

$$\beta_i = P(\alpha_i[t] = 1). \quad (4)$$

The complementary probability of failing on receiving an output sample from sensor i is $P(\alpha_i[t] = 0) = 1 - \beta_i$. With these definitions, the probability of having a sampling scenario defined by $N(s_k)$ and $\boldsymbol{\alpha}(s_k)$ is given by

$$P\{s_k\} = \left(\prod_{i=1}^{n_m}(1-\beta_i)^{N(s_k)-1}\right)\prod_{i=1}^{n_m}(1-\beta_i)^{1-\alpha_i(s_k)}(\beta_i)^{\alpha_i(s_k)} \quad (5)$$

Remark 1. Let $\alpha_i[t]$ $(i = 1,\ldots,n_m)$ be binomial variables with $P\{\alpha_i[t] = 1\} = \beta_i$. Let us call $N \in \mathbb{N}$ the number of periods between two consecutive instants with measurements reception, i.e., such that $\alpha_i[t+j] = 0$ for $i = 1,\ldots,n_m$ and $j \in \{1, N-1\}$. For a given $\varepsilon \in (0,1)$, if $N_{\max} \in \mathbb{N}$ is chosen to fulfill

$$N_{\max} > \frac{\ln(\varepsilon)}{\sum_{i=1}^{n_m}\ln(1-\beta_i)}, \quad (6)$$

then, $P\{N_j > N_{\max}\} < \varepsilon$.

With the previous result, the possible scenarios with a probability greater than $1-\boldsymbol{\varepsilon}$ are all the possible combinations of $\boldsymbol{\alpha}_k \in \Xi$ and $N(s_k) \in \mathcal{N} = \{1, \ldots, N_{\max}\}$, and can be characterized by the probability (5). The sampling scenario simply enumerates all of those combinations leading to a set $s_k \in \mathcal{S} = \{1, \ldots, N_{\max} \cdot 2^{n_m-1}\}$. Note that the smaller the value of $\boldsymbol{\varepsilon}$, the larger the value of $N_{\max}$, and, therefore, the number of possible scenarios to take into account.

3. FAULT ESTIMATION

In order to estimate the fault vector, an extended order model is used. The fault vector evolution can be written as $\boldsymbol{f}[t+1] = \boldsymbol{f}[t] + \Delta\boldsymbol{f}[t]$, where $\Delta\boldsymbol{f}[t]$ is the variation of the fault from instant t to $t+1$. The system dynamics can be rewritten as

$$\begin{bmatrix}\boldsymbol{x}[t+1]\\ \boldsymbol{f}[t+1]\end{bmatrix} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}_f\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}[t]\\ \boldsymbol{f}[t]\end{bmatrix} + \begin{bmatrix}\boldsymbol{B}_u\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{u}[t] + \begin{bmatrix}\boldsymbol{B}_w & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{w}[t]\\ \Delta\boldsymbol{f}[t]\end{bmatrix}. \tag{7}$$

where a new state vector is introduced including the fault, and its variation from one period to the next is considered as a disturbance. The following notation is introduced for the extended order model:

$$\bar{\boldsymbol{x}}[t+1] = \bar{\boldsymbol{A}}\,\bar{\boldsymbol{x}}[t] + \bar{\boldsymbol{B}}_u\,\boldsymbol{u}[t] + \bar{\boldsymbol{B}}_w\bar{\boldsymbol{w}}[t]. \tag{8a}$$

The measurement equation is now written as

$$m_{i,k} = \underbrace{[\boldsymbol{c}_i\ \boldsymbol{h}_i]}_{\bar{\boldsymbol{c}}_i}\bar{\boldsymbol{x}}_k + \boldsymbol{d}_i\,\boldsymbol{u}_k + v_{i,k},\quad i = 1, \ldots, n_m. \tag{8b}$$

where a new vector $\bar{\boldsymbol{c}}_i$ has been introduced. The order of the extended order system will be noted as $\bar{n} = n + n_f$. It is assumed that the system (8) is detectable for any of the possible measurement patterns. This extended order model is used to estimate the system faults as follows. Initially, the state is observed running the model in open loop, leading to

$$\hat{\bar{\boldsymbol{x}}}[t^-] = \bar{\boldsymbol{A}}\,\hat{\bar{\boldsymbol{x}}}[t-1] + \bar{\boldsymbol{B}}_u\,\boldsymbol{u}[t-1], \tag{9a}$$

where a null disturbance and fault change are considered as their best estimations with the available information. Depending on the availability of a new measurement at $t = t_k$ (i.e. some m_i has been received successfully), the estimated state is updated by

$$\hat{\bar{\boldsymbol{x}}}[t_k] = \hat{\bar{\boldsymbol{x}}}[t_k^-] + \sum_{i=1}^{n_m}\boldsymbol{\ell}_{i,k}\,(m_{i,k} - \bar{\boldsymbol{c}}_i\,\hat{\bar{\boldsymbol{x}}}[t_k^-] - \boldsymbol{d}_i\,\boldsymbol{u}_k)\,\alpha_{i,k}. \tag{9b}$$

where $\boldsymbol{\ell}_{i,k}$ is the gain vector used to update the estimated state with the measurement $m_{i,k}$. If there is no measurement available, the best estimation is the one obtained running the model in open loop, i.e. $\hat{\bar{\boldsymbol{x}}}[t] = \hat{\bar{\boldsymbol{x}}}[t^-]$. Finally, the fault estimation vector is obtained from the extended vector state as

$$\hat{\boldsymbol{f}}[t] = [\boldsymbol{0}_{n_f\times n}\ \boldsymbol{I}]\ \hat{\bar{\boldsymbol{x}}}[t] = \boldsymbol{C}_f\,\hat{\bar{\boldsymbol{x}}}[t], \tag{9c}$$

where $\boldsymbol{0}_{n_f\times n}$ is a null matrix of the indicated order.

The dynamics of the fault estimation error depends on the matrix gain

$$\boldsymbol{L}_k = [\boldsymbol{\ell}_{1,k}\ \boldsymbol{\ell}_{2,k}\ \cdots\ \boldsymbol{\ell}_{n_m,k}] \tag{10}$$

defined at measuring instants ($t = t_k$), that must be designed to assure: the observer stability, robustness to the irregular data availability, a proper attenuation of the disturbances and measurement noises and a fast tracking of the real fault. In order to design the fault detector (9) with these properties, the prediction error dynamic equation must be obtained.

Lemma 2. (Fault estimation error dynamics). The fault estimation error dynamics of the algorithm (9) applied to system (8) when there is at least one measurement available every N_k input periods (with N_k time variant depending on the network traffic), is described by the linear time-variant system

$$\tilde{\bar{\boldsymbol{x}}}_k = \left(\boldsymbol{I} - \sum_{i=1}^{n_m}\boldsymbol{\ell}_{i,k}\,\bar{\boldsymbol{c}}_i\,\alpha_{i,k}\right)\bar{\boldsymbol{A}}^{N_k}\,\tilde{\bar{\boldsymbol{x}}}_{k-1} - \sum_{i=1}^{n_m}\boldsymbol{\ell}_{i,k}\,v_{i,k}\,\alpha_{i,k} \tag{11a}$$

$$+\left(\boldsymbol{I} - \sum_{i=1}^{n_m}\boldsymbol{\ell}_{i,k}\,\bar{\boldsymbol{c}}_i\,\alpha_{i,k}\right)\sum_{j=1}^{N_k}\bar{\boldsymbol{A}}^{j-1}\bar{\boldsymbol{B}}_w\bar{\boldsymbol{w}}[t_k - j]$$

$$\tilde{\boldsymbol{f}}_k = \boldsymbol{C}_f\,\tilde{\bar{\boldsymbol{x}}}_k \tag{11b}$$

that is updated every measuring instant. The estimation error vector is defined when a measurement is available ($t = t_k$) as $\tilde{\bar{\boldsymbol{x}}}_k \equiv \tilde{\bar{\boldsymbol{x}}}[t_k] = \bar{\boldsymbol{x}}[t_k] - \hat{\bar{\boldsymbol{x}}}[t_k]$, while the fault estimation error is defined as $\tilde{\boldsymbol{f}}_k = \boldsymbol{f}[t_k] - \hat{\boldsymbol{f}}[t_k]$.

Proof. The proof is based on the recursive application of (9a) for packet dropout, and application of (9b) when successful reception of measurements occur.

The goal of the present work is to find a procedure to design the matrix $\boldsymbol{L}_k$ such that the system (11) attains prescribed stability, disturbance attenuation and fault tracking conditions. The design of gain $\boldsymbol{L}_k$ is addressed by assuming a different matrix gain for each pair $(N_k, \boldsymbol{\alpha}_k)$, i.e., for each possible value of the network sampling parameter s_k. The calculation of matrices $\boldsymbol{L}_k$ is done off-line and gives as a result a finite set of gains

$$\boldsymbol{L}_k = \boldsymbol{L}(s_k) \in \boldsymbol{\mathcal{L}} = \{\boldsymbol{L}(1), \boldsymbol{L}(2), \ldots, \boldsymbol{L}(q)\}. \tag{12}$$

where q is the number of possible sampling scenarios. Every time a new measurement is received (with values of one or more sensors), a different gain $\boldsymbol{L}_k$ (depending on the value s_k) is applied to update the state estimation with equation (9b).

Remark 3. If a new vector gathering the noise at measurement instant and the extended disturbances between received measurements is defined as

$$\mathbb{W}_k = \left[\boldsymbol{v}_k^T,\ \bar{\boldsymbol{w}}[t_k-1]^T,\ \bar{\boldsymbol{w}}[t_k-2]^T,\ \cdots,\ \bar{\boldsymbol{w}}[t_k-\beta]^T\right]^T,$$

with $\beta = \max\{\mathcal{N}\}$, (the maximum number of intersampling periods N_k), and defining the observer gain matrix as in (12), the estimation error dynamics (11) can be written parametrically as

$$\tilde{\bar{\boldsymbol{x}}}_k = \boldsymbol{\mathcal{A}}(s_k)\,\tilde{\bar{\boldsymbol{x}}}_{k-1} + \boldsymbol{\mathcal{B}}(s_k)\mathbb{W}_k \tag{13a}$$

$$\tilde{\boldsymbol{f}}_k = \boldsymbol{C}_f\,\tilde{\bar{\boldsymbol{x}}}_k, \tag{13b}$$

where

$$\boldsymbol{\mathcal{A}}(s_k) = (\boldsymbol{I} - \boldsymbol{L}(s_k)\,\boldsymbol{\alpha}(s_k)\,\bar{\boldsymbol{C}})\,\bar{\boldsymbol{A}}^{N(s_k)},$$

$$\boldsymbol{\mathcal{B}}(s_k) = \left[-\boldsymbol{L}(s_k)\,\boldsymbol{\alpha}(s_k)\quad (\boldsymbol{I} - \boldsymbol{L}(s_k)\,\boldsymbol{\alpha}(s_k)\,\bar{\boldsymbol{C}})\boldsymbol{\Lambda}(N(s_k))\right].$$

With the introduction of parameter s_k the error dynamics is represented as a linear time parametric varying system. As the parameter has only a set of previously known possible values (the sampling scenarios), the error dynamics behaves as a jump linear system and $\mathcal{H}_\infty$ LMI based approaches can be applied.

4. $\mathcal{H}_\infty$ FAULT ESTIMATOR DESIGN

In order to design a stable fault detector that estimates as fast as possible the faults and minimizes the false-alarm ratio, the knowledge about the disturbances $\boldsymbol{w}$ and measurement noises $\boldsymbol{v}$ is taken into account, and its attenuation is dealt with using $\mathcal{H}_\infty$ performance. The fault magnitude is assumed to be previously unknown, although the fault detection threshold is used as a design parameter.

Two different design strategies are addressed. The first one consists of fixing the fault detection threshold and then finding the fault estimator that minimizes the response time, while imposing an adequate disturbances attenuation level in order to avoid the fault estimation to reach the predefined threshold in the absence of fault. The second strategy consists of fixing the desired time response of the fault detector and then maximizing disturbances attenuation and finding the threshold that defines the minimum detectable fault.

Theorem 4. (Robust $\mathcal{H}_\infty$ performance). Consider the fault estimation algorithm (9) applied to system (1) where the measurements of each sensor are received every period with a known probability β_i. Assume there is at least one received measurement every $N_k < N_{\max}$ periods (with $N_{\max}$ fulfiling (6) for a given small $\boldsymbol{\varepsilon}$), and that the sensors availability every time there is a successful transmission is given by the matrix $\boldsymbol{\alpha}_k \in \Xi$. Assume there are q different possible combinations of N_k and α_k. For given constants $\gamma_{v_1},\dots,\gamma_{v_{n_m}}$, $\gamma_{w_1},\dots,\gamma_{w_{n_w}}$, $\gamma_{f_1},\dots,\gamma_{f_{n_f}}$, assume that there exist matrices $\boldsymbol{P}=\boldsymbol{P}^\top \in \mathbb{R}^{\bar{n}\times\bar{n}}$, $\boldsymbol{X}(j)\in\mathbb{R}^{\bar{n}\times n_m}$, $j=1,\dots,q$ such that

$$\begin{bmatrix} \bar{\boldsymbol{P}} & \boldsymbol{M}_A & \boldsymbol{M}_B \\ \boldsymbol{M}_A^\top & \boldsymbol{P}-\boldsymbol{C}_f^\top\boldsymbol{C}_f & \boldsymbol{0} \\ \boldsymbol{M}_B^\top & \boldsymbol{0} & \boldsymbol{\Gamma} \end{bmatrix} \succ 0, \tag{14}$$

with

$$\bar{\boldsymbol{P}} = \operatorname{diag}\{\boldsymbol{P},\dots,\boldsymbol{P}\} \tag{15a}$$

$$\boldsymbol{M}_A = \begin{bmatrix} \sqrt{p_1}\left(\boldsymbol{P}-\boldsymbol{X}(1)\boldsymbol{\alpha}(1)\bar{\boldsymbol{C}}\right) & \boldsymbol{A}^{N(1)} \\ \vdots & \\ \sqrt{p_q}\left(\boldsymbol{P}-\boldsymbol{X}(q)\boldsymbol{\alpha}(q)\bar{\boldsymbol{C}}\right) & \boldsymbol{A}^{N(q)} \end{bmatrix},$$

$$\boldsymbol{M}_B = \begin{bmatrix} -\sqrt{p_1}\boldsymbol{X}(1)\boldsymbol{\alpha}(1) & \sqrt{p_1}(\boldsymbol{P}-\boldsymbol{X}(1)\boldsymbol{\alpha}(1)\bar{\boldsymbol{C}})\boldsymbol{\Lambda}(N(1)) \\ \vdots & \\ -\sqrt{p_q}\boldsymbol{X}(q)\boldsymbol{\alpha}(q) & \sqrt{p_q}(\boldsymbol{P}-\boldsymbol{X}(q)\boldsymbol{\alpha}(q)\bar{\boldsymbol{C}})\boldsymbol{\Lambda}(N(q)) \end{bmatrix}$$

where

$$\boldsymbol{\Gamma} = \operatorname{diag}\{\boldsymbol{\Gamma}_v,\boldsymbol{\Gamma}'_w\}, \boldsymbol{\Gamma}'_w = \operatorname{diag}\{r_{N_{\max}}\boldsymbol{\Gamma}_{\bar{w}},\dots,r_1\boldsymbol{\Gamma}_{\bar{w}}\},$$
$$\boldsymbol{\Gamma}_v = \operatorname{diag}\{\bar{N}_1\gamma_{v_1},\dots,\bar{N}_{n_m}\gamma_{v_{n_m}}\}, \boldsymbol{\Gamma}_{\bar{w}} = \operatorname{diag}\{\boldsymbol{\Gamma}_w,\boldsymbol{\Gamma}_f\}$$
$$\boldsymbol{\Gamma}_w = \operatorname{diag}\{\gamma_{w_1},\dots,\gamma_{w_{n_w}}\}, \boldsymbol{\Gamma}_f = \operatorname{diag}\{\gamma_{f_1},\dots,\gamma_{f_{n_f}}\}.$$

and the probabilities p_j, r_j and $\bar{N}_i$ are defined as follows. p_j is the probability of having a sampling scenario $s_k=j$ given by

$$p_j = \left(\prod_{i=1}^{n_m}(1-\beta_i)^{N(j)-1}\right)\prod_{i=1}^{n_m}(1-\beta_i)^{1-\alpha_i(j)}\beta_i^{\alpha_i(j)},$$

r_j is the probability of having $j-1$ samples without measurements, $r_j = \prod_{i=1}^{n_m}(1-\beta_i)^{j-1}$, and $\bar{N}_i$ is the expected value of the time between received measurements for each sensor: $\bar{N}_i = \sum_{j=1}^{N_{\max}} j\cdot(1-\beta_i)^{j-1}\beta_i$. Assume also that

$$\lim_{K\to\infty}\sum_{k=0}^{K}\bar{N}_i v_{i,k}^2 = \lim_{T\to\infty}\sum_{t=0}^{T} v_i[t]^2 \tag{16}$$

$$\lim_{K\to\infty}\sum_{k=0}^{K}\sum_{j=1}^{N_{\max}} r_j w_i[t+j-1]^2 = \lim_{T\to\infty}\sum_{t=0}^{T} w_i[t]^2 \tag{17}$$

Then, defining the fault detector gain depending on the sampling scenario index j as $\boldsymbol{L}(j)=\boldsymbol{P}^{-1}\boldsymbol{X}(j)$, the fault estimation defined by algorithm (9) fulfils:

- Converges asymptotically to zero in average in the absence of disturbances and faults.
- Under zero initial conditions and no faults, the fault estimation error at sampling instants is bounded by
$$\mathbb{E}\|\tilde{\boldsymbol{f}}_k\|^2_{RMS} < \|\boldsymbol{\Gamma}_v^{1/2}\boldsymbol{v}[t]\|^2_{RMS} + \|\boldsymbol{\Gamma}_w^{1/2}\boldsymbol{w}[t]\|^2_{RMS} \tag{18}$$
- Under zero initial conditions and no disturbances, the fault estimation error at sampling instants is bounded by $\mathbb{E}\|\tilde{\boldsymbol{f}}_k\|_2^2 < \|\boldsymbol{\Gamma}_f^{1/2}\Delta\boldsymbol{f}[t]\|_2^2$

Proof. Introducing $\boldsymbol{X}(j)=\boldsymbol{P}\boldsymbol{L}(j)$ in (14), applying Schur complements and multiplying inequality by $[\tilde{\bar{\boldsymbol{x}}}_{k-1}^\top\ \mathbb{W}_k^\top]$ on the left, and by its transpose on the right, one obtains

$$\sum_{j=1}^{q}\left(p_j\underbrace{(\boldsymbol{\mathcal{A}}(j)\tilde{\bar{\boldsymbol{x}}}_k+\boldsymbol{B}_c(j)\mathbb{W}_k)^\top}_{\star}\boldsymbol{P}(\star)\right) - \tilde{\bar{\boldsymbol{x}}}_{k-1}^\top\boldsymbol{P}\tilde{\bar{\boldsymbol{x}}}_{k-1}$$
$$+\tilde{\boldsymbol{f}}_{k-1}^\top\tilde{\boldsymbol{f}}_{k-1} - \mathbb{W}_k^\top\boldsymbol{\Gamma}\mathbb{W}_k < 0,$$

Then under null disturbances, noises and faults, and defining the Lyapunov function $\mathcal{V}_k = \tilde{\bar{\boldsymbol{x}}}_k^\top\boldsymbol{P}\tilde{\bar{\boldsymbol{x}}}_k$, the previous expression leads to $\mathbb{E}\{\mathcal{V}_k\} < \mathcal{V}_{k-1}$ that assures asymptotical convergence in average of the extended state estimation error (and hence of the fault estimation).

Now, if a null initial condition is assumed ($\tilde{\bar{\boldsymbol{x}}}_0=\boldsymbol{0}$), adding from $k=1$ to $k=K$ one obtains

$$\mathbb{E}\{\mathcal{V}\}_K + \sum_{k=1}^{K}\left(\tilde{\boldsymbol{f}}_{k-1}^\top\tilde{\boldsymbol{f}}_{k-1} - \mathbb{W}_k^\top\boldsymbol{\Gamma}\mathbb{W}_k\right) < 0. \tag{19}$$

As $\boldsymbol{P}\succ 0$, then $\mathbb{E}\{\mathcal{V}\}_K > \boldsymbol{0}$, leading to

$$\sum_{k=1}^{K}\left(\tilde{\boldsymbol{f}}_{k-1}^\top\tilde{\boldsymbol{f}}_{k-1} - \mathbb{W}_k^\top\boldsymbol{\Gamma}\mathbb{W}_k\right) < 0. \tag{20}$$

Introducing the definitions of $\boldsymbol{\Gamma}$ and $\mathbb{W}_k$ in the previous expression it leads to

$$\sum_{k=1}^{K}\left(\tilde{\boldsymbol{f}}_{k-1}^\top\tilde{\boldsymbol{f}}_{k-1} - \sum_{i=1}^{n_m}\gamma_{v_i}\bar{N}_i v_{i,k}^2 - \right. \tag{21}$$
$$\left.\sum_{j=0}^{N_{\max}-1} r_{j+1}\left(\sum_{i=1}^{n_w}\gamma_{w_i}w_i[t_{k-1}+j]^2 + \sum_{i=1}^{n_f}\gamma_{f_i}\Delta f_i[t_{k-1}+j]^2\right)\right) < 0.$$

If the faults are assumed to be null, using equations (16) and (17), dividing the above expression by K and taking limits when K tends to ∞, the RMS norm of the signals is obtained, and therefore, the fault estimation error is bounded by

$$\|\tilde{\boldsymbol{f}}_k\|^2_{RMS} < \sum_{i=1}^{n_m}\gamma_{v_i}\|v_i[t]\|^2_{RMS} + \sum_{i=1}^{n_w}\gamma_{w_i}\|w_i[t]\|^2_{RMS} =$$
$$= \|\boldsymbol{\Gamma}_v^{1/2}\boldsymbol{v}[t]\|^2_{RMS} + \|\boldsymbol{\Gamma}_w^{1/2}\boldsymbol{w}[t]\|^2_{RMS},$$

This bound is related to the minimum size of a fault to be distinguishable from the disturbances effects.

On the other hand, if null disturbances and measurement noise are assumed, taking limits as K tends to ∞ in equation (21) one obtains

$$\|\tilde{\boldsymbol{f}}_k\|_2^2 < \sum_{i=1}^{n_f} \gamma_{f_i} \|\Delta f_i[t]\|_2^2 = \|\boldsymbol{\Gamma}_f^{1/2} \Delta \boldsymbol{f}[t]\|_2^2, \tag{22}$$

If the fault is a step signal, then $\Delta \boldsymbol{f}$ is an impulse, and $\boldsymbol{\Gamma}_f$ is related to the number of samples needed by the fault detector to reach the real fault value.

Remark 5. The previous theorem demonstrates the stability of the fault estimator and gives the gains $\boldsymbol{L}(j)$ to be applied when $N_k \in \{1, \ldots, N_{\max}\}$. Nevertheless some scenarios with larger N_k can be possible (with a small probability lower than $\boldsymbol{\varepsilon}$) and the stability and the gain to be applied are not stablished. However, the values of $\boldsymbol{L}(j)$ for a given combination of sensors converges to a given value as N increases. Therefore, if $N_{\max}$ is large enough ($\boldsymbol{\varepsilon}$ low enough), when a value $N_k > N_{\max}$ occurs, the gain corresponding to the combination of sensors and $N_{\max}$ is proposed to be applied.

On the other hand, if the value of N_k is larger than $N_{\max}$ for some sporadic instants, the predictor is still stable if the following condition is fulfilled:

$$\sigma_{\max}(\boldsymbol{A}) \left(1 - \prod_{i=1}^{n_m} (1 - \beta_i)\right) < 1 \tag{23}$$

where $\sigma_{\max}(\boldsymbol{A})$ denotes the maximum singular value. This condition is always true for stable systems. For unstable systems the condition will hold if the transmission probability is sufficiently high. This can be demonstrated as follows. Between measuring instants the fault estimator runs in open loop. Then, for a given sampling instant t_k, the expected value of the state prediction error norm at the next sampling instant (when some sensor l is available) can be expressed as

$$\mathbb{E}||\tilde{\boldsymbol{x}}_{k+1}|| = \beta_l \sum_{j=1}^{\infty} \prod_{i=1}^{n_m} (1 - \beta_i)^{j-1} ||\boldsymbol{A}^j \tilde{\boldsymbol{x}}_k||$$

This value can be bounded by

$$\mathbb{E}||\tilde{\boldsymbol{x}}_{k+1}|| < \frac{\beta_l}{\prod_{i=1}^{n_m}(1-\beta_i)} \sum_{j=1}^{\infty} \left(\prod_{i=1}^{n_m} (1 - \beta_i) \sigma_{\max}(\boldsymbol{A})\right)^j ||\tilde{\boldsymbol{x}}_k||,$$

that is a finite value only if (23) is fulfilled.

The previous theorem allows to define two strategies in the design of the fault estimator, assuming that the norms of disturbances and measurement noise are known.

The first strategy consists of fixing the minimum fault to be detected ($f_{\min}$), and then minimizing the response time of the estimator. For this purpose, the following minimization problem

$$\text{Minimize } \text{trace}(\boldsymbol{\Gamma}_f), \text{ subject to LMI (14), and}$$
$$\sum_{i=1}^{n_m} \gamma_{v_i} \|v_i[t]\|_2^2 + \sum_{i=1}^{n} \gamma_{w_i} \|w_i[t]\|_2^2 < (f_{\min}/p)^2 \tag{24}$$

along variables γ_{v_i}, γ_{w_i}, $\boldsymbol{\Gamma}_f$, $\boldsymbol{P}$ and $\boldsymbol{X}(s_k)$, leads to the fastest fault detector assuring, with a probability depending on p, that the fault estimation will be under $f_{\min}$ for no faults. When a fault occurs, the time needed by the fault detector to reach the fault value is proportional to γ_{f_i}, according to equation (22).

The second design strategy is based on fixing the response time (by fixing a constant $\gamma_{f,\min}$), and minimizing the threshold (minimum detectable fault). In this case, the minimization problem

$$\text{Minimize } \sum_{i=1}^{n_m} \gamma_{v_i} \|v_i[t]\|_2^2 + \sum_{i=1}^{n} \gamma_{w_i} \|w_i[t]\|_2^2$$
$$\text{subject to LMI (14), and } \text{trace}(\boldsymbol{\Gamma}_f) < \gamma_{f,\min} \tag{25}$$

along variables γ_{v_i}, γ_{w_i}, $\boldsymbol{\Gamma}_f$, $\boldsymbol{P}$ and $\boldsymbol{X}(s_k)$, leads to the lowest fault threshold for a given response time (determined by $\gamma_{f,\min}$). The threshold is defined as

$$(f_{\min}/p)^2 = \sum_{i=1}^{n_m} \gamma_{v_i} \|v_i[t]\|_2^2 + \sum_{i=1}^{n} \gamma_{w_i} \|w_i[t]\|_2^2.$$

where p depends on the desired probability of false alarms. The fault detection logic is simply

$$\begin{cases} \text{if } \|\sum_{k-R+1}^{k} \hat{\boldsymbol{f}}_i / R\|^2 < f_{\min}^2 \rightarrow \text{no fault} \\ \text{if } \|\sum_{k-R+1}^{k} \hat{\boldsymbol{f}}_i / R\|^2 > f_{\min}^2 \rightarrow \text{fault} \end{cases} \tag{26}$$

where the false alarm probability depends on R and p. For example, if the disturbances are independent gaussian signals, for $R = 1$ and $p = 3$ the false alarm probability is about 0.3% in every measurement, while for $R = 3$ and $p = 2$ it is less than 0.1%.

5. EXAMPLES

Consider an industrial continuous-stirred tank reactor with process matrices (Gao et al. [2008]):

$$\boldsymbol{A} = \begin{bmatrix} 0.9719 & -0.0013 \\ -0.0340 & 0.8628 \end{bmatrix}, \boldsymbol{B}_u = \begin{bmatrix} -0.0839 & 0.0232 \\ 0.0761 & 0.4144 \end{bmatrix},$$
$$\boldsymbol{B}_w = \boldsymbol{B}_u, \boldsymbol{B}_f = \begin{bmatrix} -0.0839 \\ 0.0761 \end{bmatrix}, \bar{\boldsymbol{C}} = \begin{bmatrix} 1 & 0 & 0.01 \\ 0 & 1 & 0.01 \end{bmatrix},$$
$$\boldsymbol{d}_1 = \boldsymbol{d}_2 = [0 \ 0]\,.$$

It is assumed that the probability of receiving a measurement from the first sensor is $\beta_1 = 0.4$, while for the second one it is $\beta_2 = 0.3$ (they have different network access priorities). Taking a value of $\boldsymbol{\varepsilon} = 0.0001$, the maximum sampling period to take into account (fulfilling (6)) is $N_{\max} = 11$. The measurement noises variances are 0.005^2 and 0.05^2, while the white noise input disturbances variances are 0.001^2.

Assume that faults larger than 0.2 must be detected as fast as possible. For this purpose, the threshold $f_{\min} = 0.4$ (with $p = 2$) is fixed and minimization problem (24) is solved leading to a set of 33 gain matrices $L(s_k)$ and an index $\gamma_f = 39.3$. In figure 1 a simulation is shown for this fault detector where a fault of size 0.8 has been considered. As can be observed, the time needed to reach the 95% of the real fault value is about 11 control periods (with the second measurement).

Assume now that a detector about 10 times slower than the previous one can be admissible in order to reduce the threshold and hence the minimum detectable fault. For this purpose, an index $\gamma_f = 500$ is fixed and minimization problem (25) is solved, leading to a set of 33 observer gain matrices $L(s_k)$ and a threshold of $f_{\min} = 0.142$ (with $p = 2$). In figure 2 a simulation is shown where an small fault of 0.3 units has been applied in order to show the time response of the fault detector and the accuracy on

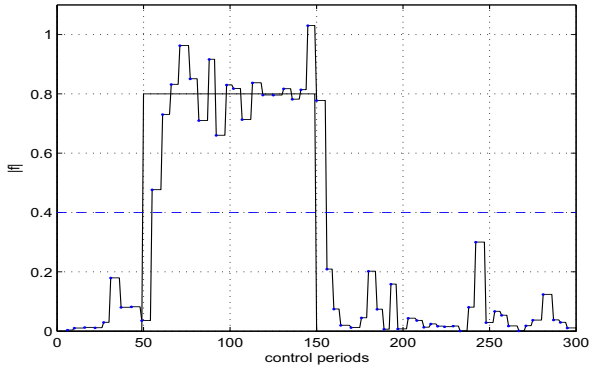


Fig. 1. Fault(–), fault estimation (- -), sampling times (·), and threshold. $f_{min} = 0.4$.

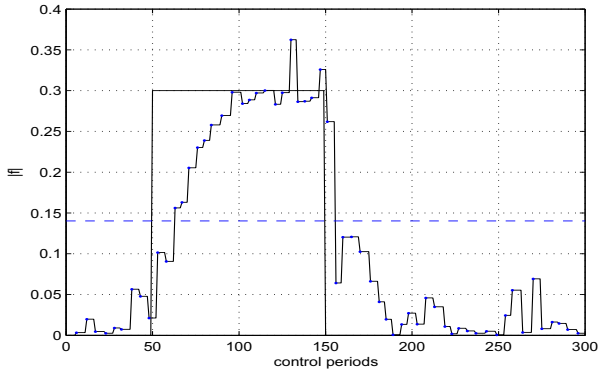


Fig. 2. Fault(–), fault estimation (- -), sampling times (·), and threshold. $\gamma_f = 500$.

fault detection. It can be seen how fault errors due to disturbances and noise measurements are below the given threshold. The time needed by the fault detector to reach the 95% of the real fault value is about 50 control periods (11 avaliable measurements).

Comparing both designs, it is easy to see that the use of a threshold of $f_{\min} = 0.142$ in the first fault detector would have led to several false alarms, while using a threshold of $f_{\min} = 0.4$ in the second detector, the fault would'nt have been detected.

6. CONCLUSIONS

In this work, the design of a fault detection and estimation algorithm for linear systems with several sensors connected with the fault detector through a network, is addressed. The fault detector is addressed as an observer whose states and faults estimation is updated with the scarcely available measurements. $\mathcal{H}_\infty$ performance has been taken into account, making the observer robust to the disturbances and the irregular data availability with its associated probability.

Two different design strategies have been proposed, showing a compromise between disturbances attenuation and convergence speed of the fault estimator. On one hand, an strategy that minimizes the response time for a given threshold has been proposed. On the other hand, an strategy that minimizes the fault threshold (i.e., maximizes the disturbances attenuation) for a given convergence speed has been addressed.

The proposed strategies give as a result a gain scheduled observer where the observer estimation update gain is a function of the sensor availability and its probability. Some simulations have illustrated the design compromise, demonstrating the validity of the proposed approach.

The use of different thresholds (possibly adaptive), for independently detecting each element of the fault vector, will be investigated in future works.

REFERENCES

Chen, J. and Patton, R. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers.

Ding, S., Jeinisch, T., Frank, P., and Ding, E. (2000). A unified approach to the optimization of fault detection systems. *International Journal of adaptive contorol and signal processing*, 14, 725–745.

Gao, H., Chen, T., and Wang, L. (2008). Robust fault detection with missing measurements. *International Journal of Control*, 81(5), 804–819.

He, X. and Zhou, Z.W.D. (2008). Fault detection for networked systems with incomplete measurements. *Proceedings of the 17th World Congress*, 13557–13562.

Henry, D. and Zolghadri, A. (2005). Design of fault diagnosis filters: A multi-objective approach. *Journal of The Franklin Institute*, 342, 421–446.

Henry, D. and Zolghadri, A. (2006). Norm-based design of robust fdi schemes for uncertain systems under feedback control: Comparison of two approaches. *Control Engineering Practice*, 14(9), 1081–1097.

Hou, M. and Patton, R. (1996). An LMI approach to $\mathcal{H}_-/\mathcal{H}_\infty$ fault detection observers. *UKACC International Conference on Control*, (427), 305–310.

Izadi, I., Zhao, Q., and Chen, T. (2005). An optimal scheme for fast rate fault detection based on multirate sampled data. *Journal of Process Control*, 15, 307–319.

Li, W., Shah, S., and Xiao, D. (2005). Kalman filters for non-uniformly sampled multirate systems. *16th IFAC world congress*.

Rank, M. and Niemann, H. (1999). Norm-based design of robust fdi schemes for uncertain systems under feedback control: Comparison of two approaches. *International Journal of Control*, 72(9), 773–783.

Wang, J., Yang, G., and Liu, J. (2007). An LMI approach to $\mathcal{H}_-$ index and mixed $\mathcal{H}_-\mathcal{H}_\infty$ fault detection observer design. *Automatica*, 43, 1656–1665.

Wang, Y., Ding, S.X., Ye, H., Wei, L., Zhang, P., and Wang, G. (2008). Fault detection of networked control systems with packet based periodic communication. *International Journal of Adaptive Control and Signal Processing*.

Zhang, P., Ding, S., Frank, P., and Sader, M. (2004). Fault detection of networked control systems with missing measurements. *5th Asian Control Conference*, 2, 1258–1263.

Zhong, M., Ding, S., Lam, J., and Wang, H. (2003). An LMI approach to design robust fault detection filter for uncertain LTI systems. *Automatica*, 39, 543–550.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Optimization of a Water For Injection Control System for a Pharmaceutical Plant

A. Visioli * **M. Ammannito** ** **M. Caselli** *** **M. Incardona** ****

* *Dipartimento di Ingegneria dell'Informazione, University of Brescia, Italy*
e-mail: antonio.visioli@ing.unibs.it
** *Eli Lilly, Sesto Fiorentino, Florence (Italy)*
e-mail: ammannito_massimiliano@lilly.com
*** *ER Sistemi, Parma (Italy)*
e-mail: michele.caselli@ersistemi.it
**** *formerly with Stilmas SpA, Settala, Milan (Italy), now with Pharmagel Engineering SpA, Lodi (Italy)*
e-mail: marco.incardona@pharmagel.it

Abstract: The use of a Proportional-Integral-Derivative (PID) plus feedforward technique for a temperature control loop in a Water For Injection (WFI) utility in a pharmaceutical plant is described in this paper. In particular, the process considered consists in cooling the WFI which has to fill a tank for the production of insulin. In this context, the use of an optimal feedforward action allows to fill the tank in a short time by satisfying at the same time the tight temperature constraints.

1. INTRODUCTION

Proportional-Integral-Derivative (PID) controllers are the controllers most adopted in industry due to the good cost/benefit ratio they are able to provide. In fact, they can provide satisfactory performances for a wide range of processes, despite their ease of use. Tuning and automatic tuning techniques have been developed to help the operators to select appropriate values for the parameters in such a way that less and less specific knowledge is required to use them (O'Dwyer, 2006; Leva et al., 2001). However, it is well known that the performances of these controllers much depend, in addition to the tuning of the PID parameters, to the appropriate implementation of those additional functionalities, such as anti-windup, set-point filtering, feedforward, and so on (Åström and Hägglund, 2006; Visioli, 2006). Methodologies for the effective design of such a features are nowadays easier and easier to implement, due to the increase of computational power available in industrial plants.

In this context, a design method for a PID plus feedforward controller has been proposed in (Visioli, 2004), aiming at achieving a minimum output transition time, subject to actuator constraints, when a set-point change is required. The technique relies on assuming a first-order-plus-dead-time (FOPDT) model of the process and on applying the maximum actuator effort for a determined time interval. Then, in order to cope with the unavoidable model uncertainties, the PID parameters are tuned according to any conventional method and the reference signal for the closed-loop is determined by filtering appropriately the step reference signal. In this way, a high performance is obtained in set-point following task without impairing the load rejection capabilities.

In this paper we show an interesting application of this PID plus feedforward methodology. In particular, a plant for the production of insulin has been considered, in which the distribution of Water For Injection (WFI) plays a key role. This kind of water is a highly purified water, with requirements defined by three standards mainly used in the pharmaceutical industry (US Pharmacopea, EU Pharmacopea, and Japanese Pharmacopea). In order to keep a low microbiological content, WFI is kept at a high temperature and continuously recirculating using closet circuits (called loops), starting from a storage tank. However, production lines and machines in several production steps require WFI at a temperature much lower than the one at which the WFI is recirculated in the loop. For this reason fast set-point changes are often required and therefore this kind of control task is suitable to be tackled with the PID plus feedforward technique proposed in (Visioli, 2004).

The paper is organized as follows. In Section 2 the PID plus feedforward technique is reviewed. Then, the plant and the control task are described respectively in Section 3 and 4. Experimental results are given in Section 5. Finally, conclusions are drawn in Section 6.

2. PID PLUS FEEDFORWARD DESIGN

Typically, a PID plus feedforward control scheme, for the purpose of improving the set-point response, is implemented as shown in Figure 1 (Åström and Hägglund, 2006), where $M(s)$ is a reference model that gives the desired response of a set-point change and $G(s)$ is chosen as

$$G(s) = \frac{M(s)}{\tilde{P}(s)}. \quad (1)$$

where $\tilde{P}(s)$ is the minimum-phase part of the process transfer function $P(s)$. However, it is difficult with this scheme to address explicitly the actuator constraints and therefore to achieve a minimum-time process variable transition. Indeed, a transition of the process output y from the value y_0 to the value y_1 in a minimum time interval subject to the actuator constraint can be obtained by implementing te following methodology. In this section, for the sake of clarity and without loss of generality, it will be assumed $y_0 = 0$ and $y_1 > 0$.

First, as it is standard practice in an industrial context, the process is described by a FOPDT model, i.e.:

$$P(s) = \frac{K}{Ts+1}e^{-Ls}. \quad (2)$$

Many different methods have been devised in order to determine the model parameters by means of a simple identification experiment (see for example (Åström and Hägglund, 2006; Visioli, 2006)). Based on this model, the output u_{ff} of the (nonlinear) feedforward block FF is defined as follows:

$$u_{ff}(t) = \begin{cases} \bar{u}_{ff} & \text{if } t < \tau \\ y_1/K & \text{if } t \geq \tau \end{cases} \quad (3)$$

where $\bar{u}_{ff}$ is the maximum output value of the actuator and the value of τ is determined, after trivial calculations, in such a way that the process output y (that is necessarily zero until time $t = L$) is y_1 at time $t = \tau + L$. It results:

$$\tau = -T \ln \left(\frac{\bar{u}_{ff} - y_1/K}{\bar{u}_{ff}} \right). \quad (4)$$

In this way, if the process is described perfectly by model (2), an output transition in the time interval $[L, \tau + L]$ occurs. Then, at time $t = \tau + L$ the output settles at value y_1 because of the constant value assumed by $u_{ff}(t)$ for $t \geq \tau$.

Formally, we have that the nominal system output is

$$y(t) = \begin{cases} 0 & \text{if } t < L \\ \bar{u}_{ff}\,(1 - \exp(-t/T)) & \text{if } L < t < \tau + L \\ y_1 & \text{if } t \geq \tau \end{cases} \quad (5)$$

Obviously, in order to cope with unavoidable model uncertainties, a feedback action has to be provided. This is accomplished by implementing a PID controller that can be tuned, in principle, by any conventional method O'Dwyer (2006). However, as high set-point following performances are ensured by the application of the feedforward action, it is suggested to select the PID gains in order to achieve a satisfactory performance in the rejection of load disturbances. In fact, the deviations due to the modelling errors between the desired and the actual output can be treated as the effect of a load disturbance Wallen (2000). The feedback PID controller is applied starting from time $t = \tau + L$.

Then, a suitable reference signal y_f has to be applied to the closed-loop system. It is desired that y_f be equal to the desired process output (5) that would be obtained in case the process is modelled perfectly by expression (2).

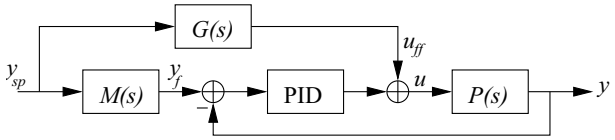


Fig. 1. The typical PID plus feedforward control scheme for the set-point following task.

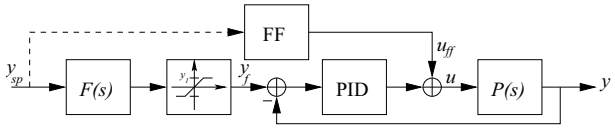


Fig. 2. The PID plus nonlinear feedforward control scheme for the set-point following task.

Thus, the step reference signal y_{sp} of amplitude y_1 has to be filtered by the system

$$F(s) = \frac{K\bar{u}_{ff}/y_1}{Ts+1}e^{-Ls} \quad (6)$$

and then saturated at the level y_1.

3. DESCRIPTION OF THE PLANT

The Eli Lilly site in Sesto Fiorentino (Florence - Italy) is dedicated to the production of insulin (injectable product). As all sites dedicated to parenteral production, Water For Injection represents one of the most important production utilities. As already mentioned, this kind of water is a highly purified water, with requirements defined by three standards mainly used in the pharmaceutical industry (US Pharmacopea, EU Pharmacopea, and Japanese Pharmacopea). These specifications are satisfied by generating WFI with different production processes all based on water distillation (Disi and Owens, 2001). In any case, one of the main requirement is represented by the low microbiological content (max 10 CFU). In order to maintain this condition, WFI is kept at a high temperature (usually around 88°C) and continuously recirculating using closet circuits (called loops), starting from the storage tank. These pipes supply the different distribution ports into the production areas and then come back to the storage tank. Stagnant water condition and the environmental temperature has to be avoided because this represents the optimal conditions for the microbiological growth in the water.

However, production lines and machines in several production steps require WFI at a temperature lower than the one at which the WFI is recirculated in the loop. Thus, in order to supply properly these production lines and machines, branches from the main loop with heat exchangers have to be installed, to be activated when the production machines required cold water. In general, this concept can be implemented in several different modes, each one with specific advantages and disadvantages (Myers et al., 2001). The architecture employed in the plant considered in this work is based on the so-called sub-loops, namely, lines coming from the main header where a cooler is installed; then, the pipe supplies the distribution ports before rejoining the main header. When the production line does not require cold water, the cooler is not activated and the sub-loop is flushed with hot water that eventually

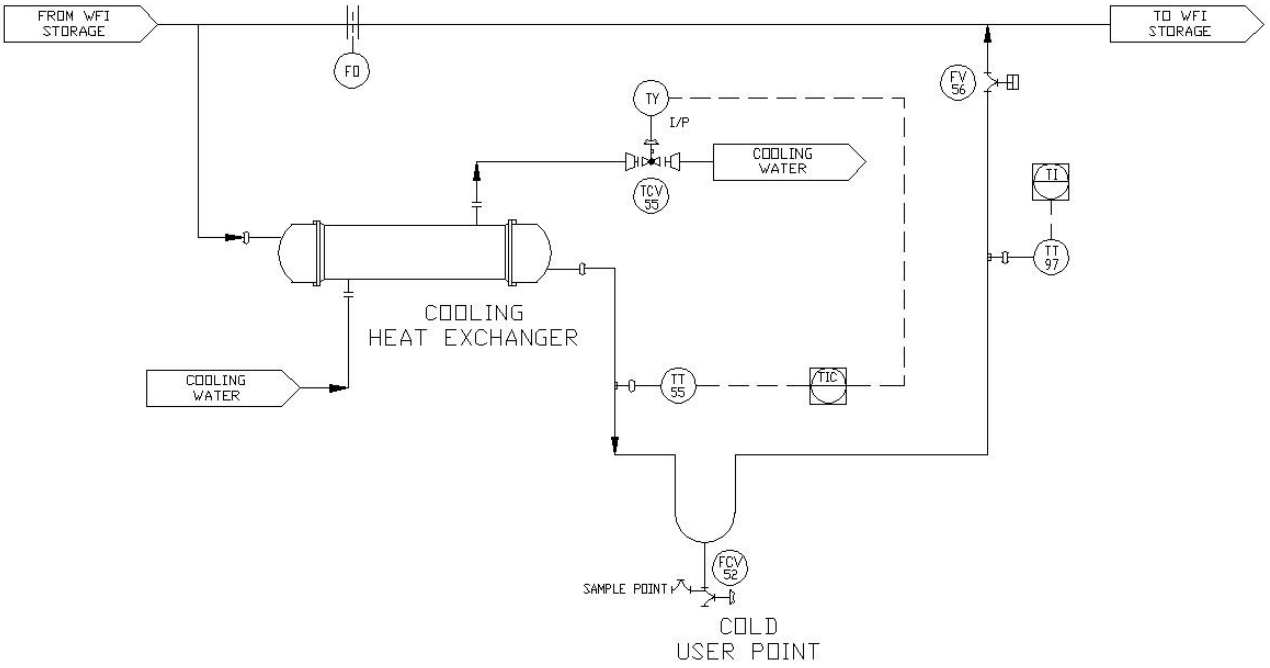


Fig. 3. The (simplified) Process and Instrumentation diagram of the considered plant.

rejoins the main flow in the main header. In this way the sub-loop is continuously flushed with hot water in order to avoid the risk of microbiological growth. When cold water is required (namely, a set-point change is applied), the heat exchanger is employed to cool down the WFI. Only when the required temperature value is attained, the distribution port valve opens (namely, a step load disturbance occurs) and the WFI flows out of the loop to supply the process machine. During the transition period, before the temperature target value is attained, the WFI is re-circulated back into the main loop.

4. CONTROL TASK

A (simplified) Process and Instrumentation (P&I) diagram of the considered plant is shown in Figure 3. The aim of the control system is to cool down the WFI, by means of the heat exchanger actuated by the valve TCV55, as fast as possible when a request of cold water is performed by the production line (in this case it is a tank where the insulin is produced). In fact, from a production standpoint the cooling down time represents a dead time that has to be reduced as much as possible. In addition, in the time interval between the cool water request and the opening of the valve, WFI at a temperature lower than 88°C is sent back to the main loop with the effect of decreasing the average temperature into the main loop, worsening the conditions for the preservation of the WFI characteristics into the loop.

Thus, in this situation, in order to reduce the amount of cold water sent back to the main loop, the valve installed on the return pipe from the sub-loop to the main header is almost totally closed, maintaining only a small flow to keep the sub-loop flushed. Hence, when the required temperature is obtained within the sub-loop, the opening of the distribution port valve has the effect to increase the WFI flow in the heat exchanger, and this represents a significant disturbance on the process which has to be compensated properly so that the temperature control specifications are still kept (otherwise the valve has to be closed again).

From the control design viewpoint it appears that a set-point following task has to be addressed when the temperature has to be lowered and a load disturbance rejection task has to be addressed when the distribution port valve opens. Specifically, the distribution port valve FCV52 can be opened when the WFI temperature set-point of 17°C is attained both before (TT55) and after the sample point (TT97) at it is maintained for at least 10 s. This strategy has been implemented in order to ensure that the temperature specifications on the production line are fully satisfied. The maximum control error is ±1.5°C, namely, if the WFI temperature exceeds the value of 18.5°C or it is lower than 15.5°C, then the control system must close the distribution port valve.

5. EXPERIMENTAL RESULTS

The PID plus feedforward control scheme has been implemented by means of standard control hardware and software. In particular, a Programmable Logic Controller has been employed and the control algorithm has been written by means of standard software components (blocks). Indeed, the feedforward action has been implemented by setting the PID controller in manual mode and by using a counter block. Then, a bumpless transition is performed when the PID controller is applied.

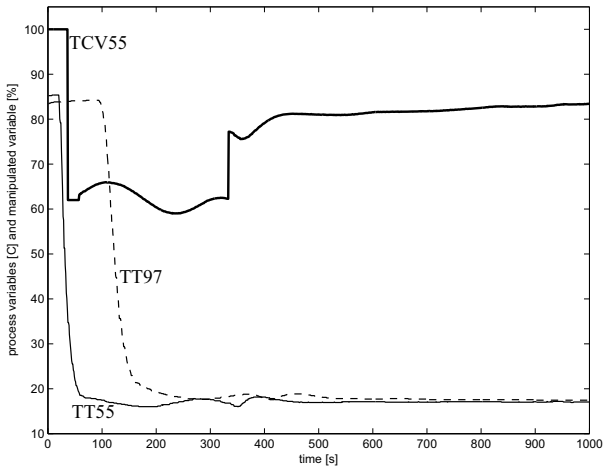


Fig. 4. The experimental results obtained with the PI plus feedforward method.

In any case, as a first step, a FOPDT model of the process has been estimated by applying the tangent method Visioli (2006) to an open-loop step response, where the temperature TT55 before the sample point has been considered as the process variable (obviously, the input of the system is the heat exchanger valve TCV55). In fact, it is supposed that the temperature TT97 will assume the same values of TT55 after a dead time and can be therefore controlled in open loop.
The estimated FOPDT model is

$$P(s) = -\frac{0.75}{0.2s+1}e^{-0.27s}, \tag{7}$$

where the time constant and the dead time are expressed in minutes. By applying expression (4) and by considering $y_1 = 88 - 17 = 71$ and $\bar{u}_{ff} = 100$ (note that the input is expressed in percentage), we have $\tau = 0.586$ min, i.e., $\tau = 35$ s. A PI controller with transfer function

$$C(s) = K_p\left(1 + \frac{1}{T_i s}\right) \tag{8}$$

has been employed. The parameters have been selected in order to achieve a satisfactory load disturbance rejection performance. After a few experiments, they have been fixed as $K_p = 0.9$ and $T_i = 11.25$. The result of the experiment with the use of the PID plus feedforward conrol law are shown in Figure 4, where the two process variables TT55 and TT97 are shown together with the control variable TCV55. In particular, after $\tau = 35$ s the control variable is set to its nominal steady-state value and then the PI controller is applied. At time $t = 323$ s the set-point response control requirements are met and therefore, after 10 s, the sample port valve opens.

In order to improve the load rejection response, a feedforward action (which consists in this case of pure gain) has been employed also in this context, according to the standard scheme of Figure 5, where $H(s)$ represents the transfer function between the disturbance and the controlled temperature.
It can be seen that the required temperature is maintained despite the load disturbance and therefore the tank is filled in a minimum time.
The proposed strategy has been therefore proven to be effective.

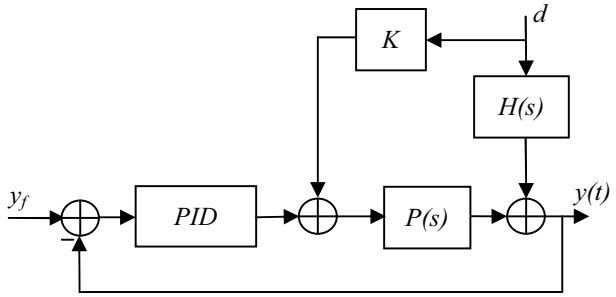


Fig. 5. The typical PID plus feedforward control scheme for load disturbance rejection.

6. CONCLUSIONS

In this paper we have shown that a novel PID plus feedforward control strategy can be employed effectively for the optimization of a WFI control system in a pharmaceutical plant. The relative ease of implementation has been highlighted and the presented results have demonstrated the effectiveness of the methodology.

REFERENCES

K. J. Åström and T. Hägglund. *Advanced PID Control*. ISA Press, Research Triangle Park, USA, 2006.
S. Disi and B. Owens. Final treatment options: Water for injection (wfi). In *Pharmaceutical Engineering Guides for New and Renowated Facilities - Volume 4: Water and Steam System, ISPE, Tampa (FL)*, 2001.
A. Leva, C. Cox, and A. Ruano. Hands-on PID autotuning: a guide to better utilisation. Technical report, IFAC Technical Brief, available at www.ifac-control.org, 2001.
R. Myers, G. Gray, B. Bader, R. Brozek, J. Cox, P. Skinner, and G. Geisler. Storage and distribution systems. In *Pharmaceutical Engineering Guides for New and Renowated Facilities - Volume 4: Water and Steam System, ISPE, Tampa (FL)*, 2001.
A. O'Dwyer. *Handbook of PI and PID Tuning Rules*. Imperial College Press, 2006.
A. Visioli. *Practical PID Control*. Springer, London, UK, 2006.
A. Visioli. A new design for a PID plus feedforward controller. *Journal of Process Control*, 14:455–461, 2004.
A. Wallen. *Tools for autonomous process control*. PhD thesis, Lund Institute of Technology, Lund, S, 2000.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Diagnosis for the Reliability Improvement of Embedded Systems using 3D Laser Vibrometer

First O. Bennouna*. Second H. Chafouk*
Third J. P. Roux***

**IRSEEM (Institut de Recherche en Systèmes Electroniques EMbarqués)*
Technopôle du Madrillet, Avenue Galilée, BP 10024, 76801 Saint Etienne du Rouvray, FRANCE
(Tel: +33232915824; e-mail: bennouna@esigelec.fr).

***CEVAA (Centre d'Etudes Vibro-Acoustiques pour l'Automobile)*
2 rue Joseph Fourier, 76800 Saint Etienne du Rouvray, France (e-mail: jp.roux@cevaa.com)

Abstract: This paper concerns the diagnosis of electronic systems based on vibration data. The electronic system used is a simplified printed circuit board. Measurements on PCBs are done using a 3D laser vibrometer. The diagnosis procedure combines wavelet transforms and artificial neural networks in order to improve the reliability of the electronic system.

Keywords: Embedded diagnosis, reliability, vibration, wavelet transform, artificial neural network, laser vibrometer.

1. INTRODUCTION

Constraints imposed by the environment play a major role in the reliability of embedded systems used in aerospace applications, automotive or rail. Compared to other parameters such as humidity or temperature, vibrations are heavily involved in the problems of mechanical reliability. In the case of electronic printed circuit boards, excessive vibration can lead to failures such as welding rupture, which can impact significantly operation and safety. In this case, diagnosis and fault detection techniques can be combined to guarantee optimal performance of the process.

The study of the influence of these vibrations is usually done in two ways: The first consists on a modal approach generally limited to the frequency measurement or calculation of the first bending and torsion modes (Cifuentes et al., 1995). They are considered as the most damaging cases during the lifecycle of the card. In the case where a vibration mode is identified as potentially dangerous, we generally search to increase its frequency so that it can't be excited in operating conditions. A second way to protect embedded systems against the influence of vibrations, is to control the evolution of data during their life cycle (Gu et al., 2007). The analysis of the vibration response of these systems can detect, localize and identify the fault. Its presence logically leads to changes in the system structure (i.e. stiffness, mass, damping); so the changes monitoring of these characteristic parameters can provide important information. The data analysis can be done in the temporal domain, frequency domain or time/frequency domain. In the last case, several articles on the wavelet transform have recently appeared. (Bayissa et al., 2007) have used the continuous wavelet transform CWT to identify structural faults on a plate, when (Han et al., 2005) worked on the same problem using wavelet packet transform WPT.

This article concerns the diagnosis of embedded systems in vehicles, based on the study of vibration signals. Different instrumentations are planned on the basis of measurements with and without contact. The wavelet transform is used to decompose the measured signals in the form of indicators (Bayissa et al., 2007) (Han et al., 2005), which constitute the input of neural networks for classification and decision. Significant results of diagnosis applied on laser vibrometer data are presented, and can consider many perspectives of implementation.

This article is organized as follows: Section 2 describes the diagnosis procedure including tools used (wavelet transform and neural networks). Tests, instrumentation and the corresponding results will be presented in Section 3. A brief summary and the innovative aspects of this project are given in the end of the paper.

2. DIAGNOSIS APPROACH

The diagnosis procedure proposed in this paper uses two techniques that are the wavelet transforms and neural networks. The first will extract indicators (energy and entropy) that will be the entries of the second to make decision and classification. It replaces the traditional methods that use either time signals with an approach generally based on correlation, or frequency analysis which is a type of Fourier transform for detecting or locating faults. This procedure is illustrated in the following figure:

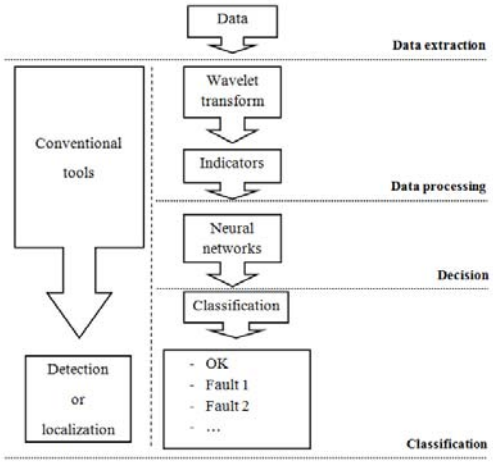


Fig. 1. Diagnosis approach.

2.1 Wavelet Transform

The wavelet transform can characterize a signal in time / frequency domains, and thus bearing the disadvantage of the Fourier transform which is the loss of temporal information. For this, there are several types: the Continuous Wavelet Transform (CWT - see appendix A), the discrete (DWT - see appendix B), and Wavelet Packet Transform (WPT). Each transformation has its advantages and drawbacks.

In this paper, the WPT is used because it provides a complete level by level decomposition of the studied signal. It also allows upgrading the problem of the DWT, which focuses essentially on low frequency bands.

Consider a temporal signal y(t), the wavelet packet function is given by:

$$\psi^i_{a,b}(t) = 2^{a/2}\psi^a(2^a t - b), \qquad i = 1,2,3,... \tag{1}$$

The integers i, a and b are respectively the modulation, scale and translation parameters. The wavelet packet coefficients $c^i_{a,b}(t)$ can be obtained using the following equation:

$$c^i_{a,b}(t) = \int_{-\infty}^{+\infty} y(t)\psi^i_{a,b}(t)\,dt \tag{2}$$

The wavelet packet component of the signal $y^i_a(t)$ can be represented by a linear combination of wavelet packet functions $\psi^i_{a,b}(t)$ as follows:

$$y^i_a(t) = \sum_{b=-\infty}^{+\infty} c^i_{a,b}(t)\psi^i_{a,b}(t) \tag{3}$$

Thus, the original signal can be reconstructed using the following expression:

$$y(t) = \sum_{i=1}^{2^a} y^i_a(t) \tag{4}$$

For a robust representation of the signal, indicators can be created. In this paper, energy and entropy are used. Thus, the components of wavelet packet energy $E_{y^i_a}$ can be calculated by:

$$E_{y^i_a} = \int_{-\infty}^{+\infty} y^i_a(t)^2\,dt \tag{5}$$

The energy can be calculated as follows:

$$E_y = \sum_{i=1}^{2^a} E_{y^i_a} \tag{6}$$

Similarly, the entropy can be calculated by:

$$S_{WT} = \sum_{j=1}^{n} p_j \,.\, \ln\left[\frac{1}{p_j}\right] = -\sum_{j=1}^{n} p_j \,.\, \ln[p_j] \tag{7}$$

With:

$$p_j = \frac{E_j}{E_{tol}} \tag{8}$$

p_j is the ratio of energy, E_j is the energy at scale j, and E_{tot} is the total energy of the signal.

2.2 Artificial Neural Networks

In recent years, the use of Artificial Neural Networks (ANN) has been extended to many domains such as prediction, classification and pattern recognition. In industrial applications using complex systems, the most used neural network is BPNN (Back Propagation Neural Network). It is a multilayer perceptron; its general structure is given in the figure bellow:

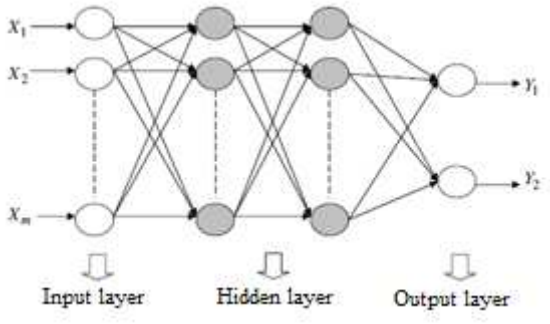


Fig. 2. BPNN's architecture.

3. APPLICATION TO EMBEDDED ELECTRONIC CARDS

The diagnosis procedure presented in the previous paragraph was applied on simplified electronic cards. Different types of damage were made on cards, or directly on the components using mechanical tools (unsoldered pins, damaged pins …). Concerning instrumentation, different configurations were tested. The use of piezoelectric sensors, glued directly on the card and on component, has been tested successfully (Bennouna et al., 2008).

The case of accelerometers glued on the surface card with an emitter placed on the component has also been tested successfully, and this according to different placements of receivers (Bennouna et al., 2009).

Whatever the sensors used, the emission signals are pulses with different lengths and amplitudes (20-40µs and 1 to 10 volts) The signals are digitized to 24 bits using dedicated workstations, for a sampling frequency of 51.2 kHz.

In this paper, laser vibrometer datas were used. They provide a matrix of response on the studied card in order to test a wide variety of placement transducers, and therefore offer optimal solutions. Examples of realized instrumentation and associated processing are summarized in Figure 3.

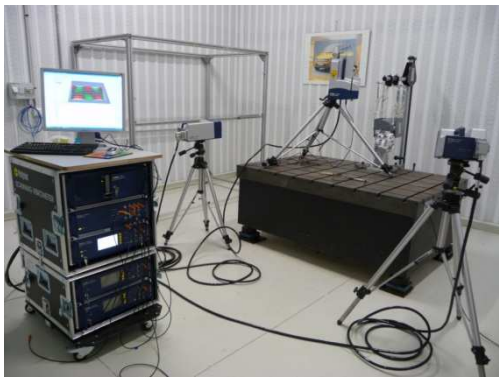


Fig. 3a. Laser vibrometer.

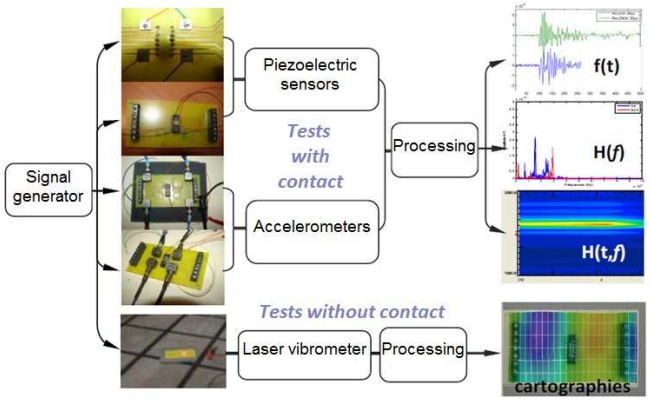


Fig. 3b. Brief Summary of tested instrumentation.

Testing support are PCBs with 14 pin welded in the centre position, and two terminals located on both sides of the component. The cards size is 100mm x 50mm x 1.5mm (Figure 4).

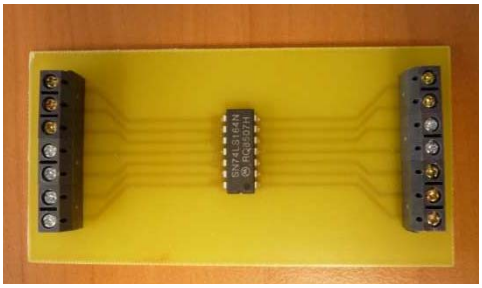


Fig. 4. Testing support.

Figure 5 illustrates the case of a 55 mesh points measured by a laser vibrometer, used as the basis of presented work.

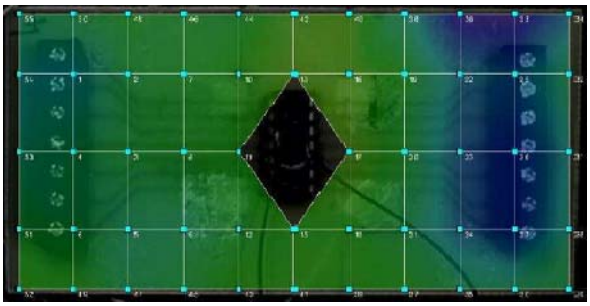


Fig. 5. 55 mesh points measured by laser vibrometry.

The figure below shows an example of signal decomposition to energy and entropy wavelet scales, which are the two indicators used. Three PCBs are tested here: PCB1 which is damage free, PCB2 where a welding point is unsoldered, and PCB3 with the presence of two faults together. Results show that the entropy appears as an indicator of the presence of defects, while energy can characterises them.

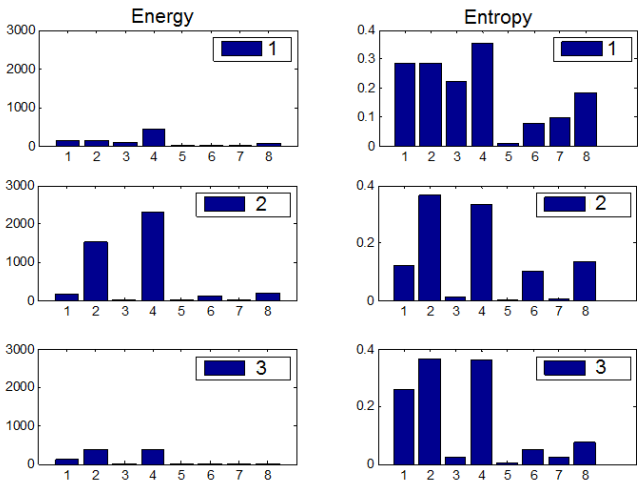


Fig. 6. Decomposition of energy and entropy scales (mother wavelet: db3).

The classification and decision part is made by a BPNN. It contains three layers: the first contains the indicators previously mentioned; the second is composed of 8 neurons, and the third with one neuron which can answer the

following question: Is there any presence of default? If so, it can identify and localize this fault. The following figure shows the model of BPNN:

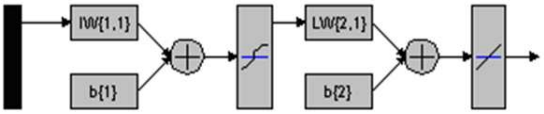


Fig. 7. BPNN's model.

Figure 8 illustrates an example of the obtained results on the 55 mesh points previously presented. Thus, for PCB3 where faults are presented by the stars, the diagnosis procedure can detect the presence of a restricted fault zone (here represented in red). This technique is promising because it can detect the presence of default and overall localize the position on the PCB.

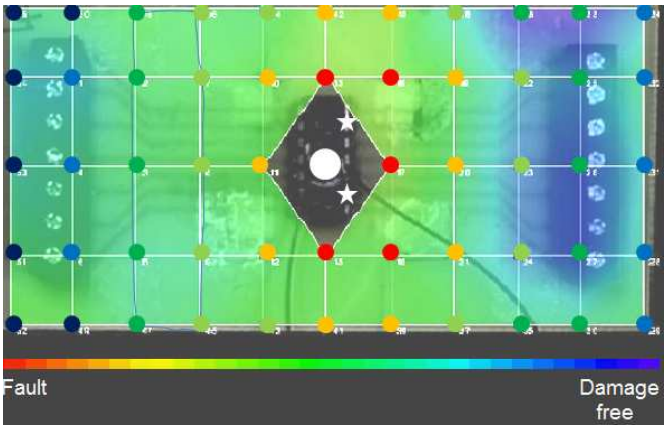


Fig. 8. Example of the diagnosis procedure results.

Based on this initial validation work with a simplified mesh, tests on densified meshes will help to evaluate the effect of the used number points reported to the precision of detection and localization.

In all cases, the ability to make a damage diagnosis map appears really promising, and it has the advantage to provide simplified defects detection. Recent studies also point in this direction (Banerjee et al., 2009), except that the obtained mapping in the cited case, is based on a calculation of damage index for each point. The principle of fault location based on an identification of channels between transmitters and receivers, also seems interesting, and will be logically tested in our case even if the frequency domain of this work is significantly different. Nevertheless, in previous work, we tested the use of damage indices (Bennouna et al., 2008) which had provided positive results for the default identification. The addition of this criterion as the input of the artificial neural network may be considered.

4. CONCLUSIONS

In this paper, a diagnosis procedure, based on a signal processing (wavelet transform), decision and classification (neural networks), has been presented. The first tool is used to decompose the signal to indicators (the energy and entropy have been used here) which are inputs of the neural network for decision and classification.

The use of different technologies and sensors configurations help to test the robustness of the diagnosis procedure. The first measurements by laser vibrometer can easily simulate the use of different sets of receiver points. On this basis, the obtained optimized configurations can be tested with the implantation of transducers on the embedded system, for example, in the case of vibration endurance.

Future work will be based on the use of densified mesh to estimate the detection precision. New cards will be tested with a weld defect placed in various locations, and lack of structural type (cracking) of various lengths that will allow us to test the performance of diagnosis aspects. In parallel, compared the use of piezoelectric accelerometers masses and different sensitivities should identify the order parameter in the implementation of in-situ sensors (preferred sensitivity or mass?). This last question is crucial in the case of practical implementation of the diagnosis.

Acknowledgements: The work presented in this article was created through funding from the Institute CARNOT ESP. The authors wish to thank everyone who helped us to have this support.

REFERENCES

Banerjee, S., Ricci, F., Monaco, E. and Mal, A. (2009). A wave propagation and vibration-based approach for damage identification in structural components. *Journal of Sound and Vibration*, vol. 322, pp.167-183.

Bayissa, W. L., Haritos, N. and Thelandersson, S. (2007). Vibration-based structural damage identification using wavelet transform. *Mechanical Systems and Signal Processing.*

Bennouna, O., Chafouk, H., Robin, O. and Roux, J.-P. (2008). A diagnosis approach combining wavelet transform and artificial neural networks. In Proceeding of the 9th conference on Sciences & techniques of Automatic control & computer engineering, Sousse, Tunisia, 20 to 22 December.

Bennouna, O., Chafouk, H., Robin, O., and Roux, J.-P. (2010). Embedded diagnosis based on vibration data. *International Journal of Adaptive and Innovative Systems*, Vol. 1 (No. 3/4), pp. 285-296.

Cifuentes, A.O., Shulga, N. and Neff, C.A. (1995). A sensitivity study of printed wiring board vibrations using a statistical method. *Journal of Sound and Vibration*, Vol. 181 (No. 4), pp. 593-604.

Gu, J., Barker, D. and Pecht, M. (2007). Prognostics implementation of electronics under vibration loading. *Microelectronics reliability*, Vol. 47, pp. 1849-1856.

Han, J. G., Ren, W. X. and Sun, Z. S. (2005). Wavelet packet based damage identification of beam structures. *International Journal of Solids and Structures*, Vol. 42, pp. 6610-6627.

Appendix A. CWT

The CWT of a signal y(t) is given by the following equation:

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} y(t) \psi^* \left(\frac{t-b}{a}\right) dt$$

where

T(a,b) are the wavelet coefficients,

a and b are respectively the scale (dilatation) and translation (position) parameters,

y(t) is the vibration response signal,

ψ^* is the complex conjugate of the mother wavelets function ψ.

The energy density functions, also known as wavelet power spectra, in the time/scale (or time/frequency) domain are:

$$C(t,f) = \left|T(a,b)\right|^2$$

For damage identification, the total energy of the time-frequency density function of the signal is calculated. It is represented by the Zeroth-Order Moment (ZOM), and given by the following equation:

$$ZOM = \int \int_{-\infty}^{+\infty} C(t,f)\, dt\, df$$

Appendix B. DWT

The DWT is faster (time calculation) than the CWT. Indeed, the mother wavelet is dilated in discrete values avoiding thus redundancy. Generally, the scale parameter a is a power of two; wavelets have the form $\psi(2^k t+l)$, where k and l are integers.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Smith Predictor Based Control of Continuous-Review Perishable Inventory Systems with a Single Supply Source

P. Ignaciuk*, A. Bartoszewicz**

Institute of Automatic Control, Technical University of Łódź
Stefanowskiego 18/22 St., 90-924 Łódź, Poland
(e-mail: *przemyslaw.ignaciuk@p.lodz.pl,**andrzej.bartoszewicz@p.lodz.pl)

Abstract: In this paper we address the problem of efficient control of continuous-review perishable inventory systems. In the considered systems the goods at a distribution center used to fulfill unknown, variable demand deteriorate at a constant rate, and are replenished with delay from a remote supply source. We develop a new supply policy which incorporates the Smith predictor to counteract the adverse effects of delay. The proposed policy guarantees that the assigned storage space at the distribution center is never exceeded which means that the cost of emergency storage is eliminated. Moreover, we show that with appropriately chosen controller parameters all of the demand imposed at the distribution center is realized from the readily available resources.

Keywords: inventory control, perishable inventory systems, time-delay systems, Smith predictor.

1. INTRODUCTION

It follows from the extensive review papers documenting the research work in the past (Nahmias, 1982; Rafaat, 1991; Goyal and Giri, 2001; Ortega and Lin, 2004; Sarimveis et al., 2008; Karaesmen et al., 2008) that certain areas of inventory control are not sufficiently addressed at the formal design level. This concerns in particular a large and very important class of problems related to the management of perishable commodities (food, drugs, gasoline, etc.). The main difficulty in developing control schemes for perishable inventories stems from the necessity of conducting an exact analysis of product lifetimes. The design problem becomes cumbersome in the situation when the product demand is subject to significant uncertainty and inventories are replenished with non-negligible delay, which frequently happens in modern supply chains. In such circumstances, in order to maintain high service level and at the same time keep stringent cost discipline, when placing an order it is necessary not only to account for the demand during procurement latency but also for the stock deterioration in that time.

Since the stock accumulation of perishables cannot be represented as a pure integrator, the effects of order procurement delay cannot be adequately accounted for by introducing the notion of work-in-progress or inventory position variables (constituting the sum of the on-hand and on-order goods), as has been done in a number of successful research works for nondecaying inventories, e.g. (Blanchini et al., 2000, Boccadoro et al., 2008). In contrast to our earlier results devoted exclusively to periodic-review inventory systems with nondeteriorating stock (Ignaciuk and Bartoszewicz, 2010a, b), in this work we analyze continuous-review systems with random lifetime of the stored goods. In order to solve the stability problems related to nonnegligible delay (see e.g. (Hoberg et al., 2007) for a discussion of the influence of delay on the dynamics of the traditional inventory systems), we propose to apply the Smith predictor (Smith, 1959). The designed control strategy is demonstrated to establish nonnegative and bounded ordering signal, which is a crucial requirement for the practical implementation of any replenishment rule. It is also shown that in the inventory system governed by the proposed policy the stock level never exceeds the assigned warehouse capacity, which means that the potential necessity for an expensive emergency storage outside the company premises is eliminated. At the same time, we demonstrate that the stock is never depleted, which implies full demand satisfaction from the readily available resources and 100% service level.

2. PROBLEM FORMULATION

We consider an inventory system where the goods at a distribution center used to fulfill the customers' (or retailers) demand are acquired with delay from a supply source. Such setting, illustrated in Fig. 1, is frequently encountered in production-inventory systems where a common point (distribution center), linked to a factory or an external, strategic supplier, is used to provide goods for another production stage or a distribution network. The task is to design a control strategy which, on one hand, will minimize the holding and shortage costs, and, on the other hand, will ensure smooth flow of goods despite unpredictable changes in market conditions.

3.1 System model

The imposed demand (the number of items requested from the distribution center) is modeled as an *a priori* unknown, bounded function of time $d(t)$, where t denotes time. We assume that demand can follow any statistical distribution as long as $0 \le d(t) \le d_{\max}$, where $d_{\max}$ is a positive constant.

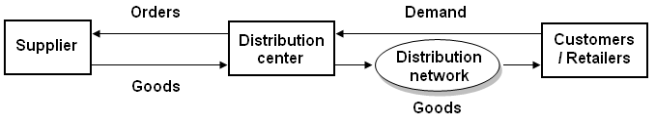


Fig. 1. Inventory system with a strategic supplier.

If there is a sufficient number of items at the distribution center to satisfy the imposed demand, then the actually met demand $h(t)$ (the number of items sold to customers or sent to retailers in the distribution network) will be equal to the requested one. Otherwise, the imposed demand is satisfied only from the arriving shipments, and the additional demand is lost (we assume that the sales are not backordered, and the excessive demand is equivalent to a missed business opportunity). Thus,

$$0 \le h(t) \le d(t) \le d_{\max}. \tag{1}$$

The on-hand stock used to fulfill the market demand deteriorates when kept in the distribution center warehouse at a constant rate σ, $0 \le \sigma < 1$. It is replenished with delay $L_p > 0$ from a remote supply source. Denoting the quantity ordered from the supplier at time t by $u(t)$, and the received shipment by $u_R(t)$, we have

$$u_R(t) = u(t - L_p). \tag{2}$$

Consequently, the stock balance equation can be written in the following way

$$\dot{y} = -\sigma y(t) + u_R(t) - h(t) = -\sigma y(t) + u(t - L_p) - h(t). \tag{3}$$

According to the stock balance equation, the on-hand stock decreases due to the realized sales represented by function $h(\cdot)$, and the decay characterized by factor σ. It is refilled from the goods acquired from the supplier $u_R(\cdot)$. For the sake of further analysis it is convenient to represent (3) in an integral form. We assume that initially the warehouse is empty, i.e. $y(0) = 0$, and the first orders are placed at $t = 0$, i.e. $u(t) = 0$ for $t < 0$. Solving (3) for $y(\cdot)$, we obtain (see the Appendix)

$$y(t) = \int_0^t e^{-\sigma(t-\tau)} u_R(\tau)\, d\tau - \int_0^t e^{-\sigma(t-\tau)} h(\tau)\, d\tau. \tag{4}$$

Since $u_R(t) = u(t - L_p)$ and $u(t < 0) = 0$, we can rewrite (4) in the following form

$$\begin{aligned} y(t) &= \int_0^t e^{-\sigma(t-\tau)} u(\tau - L_p)\, d\tau - \int_0^t e^{-\sigma(t-\tau)} h(\tau)\, d\tau \\ &= \int_0^{t-L_p} e^{-\sigma(t-L_p-\tau)} u(\tau)\, d\tau - \int_0^t e^{-\sigma(t-\tau)} h(\tau)\, d\tau. \end{aligned} \tag{5}$$

Note that in order to adequately model the stock accumulation of perishable goods, a saturating integrator needs to be applied, which makes the considered system nonlinear. However, if one can ensure that the control signal is non-negative for arbitrary t, then by introducing the function representing the actually realized sales, $h(t) \le d(t)$, the stock dynamics can be reduced to linear equation (5). In the further part of the paper, we will design a control law which will be shown to satisfy the conditions $u(t) \ge 0$ and $h(t) = d(t)$. As a result, the inventory system will stay in the linear region of operation for the whole range of disturbance $0 \le d(t) \le d_{\max}$.

3.2 Transfer function representation

The linear part of the model of the considered inventory system with perishable goods can be described using transfer functions. The system block diagram is shown in Fig. 2. The saturating integrator in an internal loop represents the operation of accumulating the stock of perishables characterized by decay factor σ. The controller, with transfer function $G_C(s)$, is supposed to steer the on-hand stock level $y(t)$ towards the reference value y_{ref}, such that a high level of demand satisfaction is achieved.

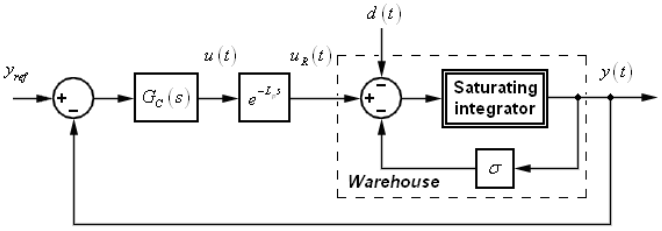


Fig. 2. System model.

3. PROPOSED CONTROL STRATEGY

The principal obstacle in providing efficient control in the considered class of systems is the latency in procuring orders. Indeed, each non-zero order placed at the supplier at instant t will appear at the distribution center with lead-time L_p at instant $t + L_p > t$ which may lead to oscillations, or even cause instability. In order to satisfactorily counteract the adverse effects of delay in the analyzed system with perishable goods, it is not sufficient to introduce inventory position variables (constituting the sum of on-hand stock and open orders), or the notion of work-in-progress, as it is usually done in the traditional inventory systems (Blanchini et al., 2000, Warburton, 2007; Boccadoro et al., 2008). This is due to the fact that the pure sum of open orders (or work-in-progress) does not account for the stock degradation within lead-time. To overcome the delay problem, in this work we propose to apply the Smith predictor (Smith, 1959), which proved a successful method of dead-time compensation in many engineering areas (Palmor, 1996). The basic idea behind the Smith predictor is to simulate the behavior of a remote plant inside the controller structure, thus eliminating the delay from the main feedback loop. The proposed control strategy, employing the Smith predictor for dead-time compensation is illustrated in Fig. 3.

The control structure consists of the primary plant controller $C(s)$ and the Smith predictor built on the linearized model of the plant $G^*(s) = 1/(s + \sigma)$. With the primary controller selected as the proportional control law $C(s) = K$, where K is a positive constant, we obtain the transfer function of the overall control structure $G_C(s)$,

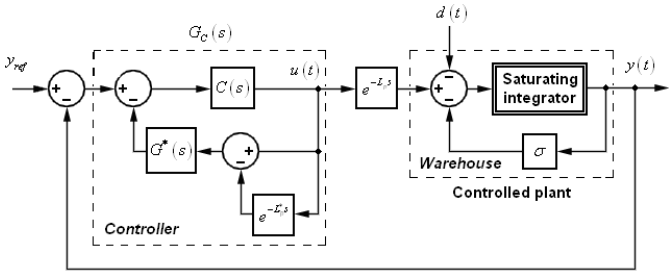


Fig. 3. Controller structure.

$$G_C(s) = \frac{C(s)}{1+C(s)\left[G^*(s)-e^{-L_p s}G^*(s)\right]} = \frac{K}{1+KG^*(s)\left(1-e^{-L_p s}\right)}. \quad (6)$$

In the linear region of operation the plant dynamics is fully represented by the transfer function $G(s) = 1/(s + \sigma)$. If the system parameters used by the controller match those of the true object, i.e. when $e^{-L_p^* s}G^*(s) = e^{-L_p s}G(s)$, then we can write the closed-loop transfer functions:

a) with respect to the reference input $Y_{ref}(s) = y_{ref} / s$

$$\frac{Y(s)}{Y_{ref}(s)} = \frac{K}{s+\sigma+K}e^{-L_p s}, \quad (7)$$

b) with respect to the disturbance $D(s) = \mathcal{L}(d(t))$,

$$\frac{Y(s)}{D(s)} = -\frac{1}{s+\sigma} + \frac{K}{s+\sigma+K}e^{-L_p s}. \quad (8)$$

It is clear from (7) and (8) that the term related to delay is eliminated from the characteristic equation (the denominator of the closed-loop transfer function). Consequently, since $K > 0$ and $\sigma \geq 0$, the closed-loop system under nominal operating conditions is stable for arbitrary lead-time and any bounded disturbance.

4. PROPERTIES OF THE PROPOSED STRATEGY

Before we state the properties of the proposed inventory policy (6), it is convenient to present it in time domain. We assume that the controller has the exact knowledge of the system parameters. Taking into account the initial conditions, we can write the control law in time domain by direct inspection of the block diagram shown in Fig. 3 in the following form

$$u(t) = K\left[y_{ref} - y(t)\right] - K\left[\int_0^t e^{-\sigma(t-\tau)}u(\tau)\,d\tau - \int_0^{t-L_p} e^{-\sigma(t-L_p-\tau)}u(\tau)\,d\tau\right]. \quad (9)$$

This control law can be interpreted as to generate orders in proportion to the difference between the current on-hand stock and its reference value $K(y_{ref} - y(t))$ decreased by the amount of open orders quantified by the rate of deterioration within the last lead-time (the terms in the square brackets).

The properties of the proposed control strategy will be given in three Theorems, and strictly proved. The first theorem shows that the ordering signal generated by the controller is always nonnegative and bounded, which is a crucial prerequisite for the implementation of any cost-efficient inventory management policy. The second proposition specifies the upper bound of the on-hand stock, which constitutes the smallest warehouse capacity required to store all the incoming shipments. Finally, the third theorem shows how to select the stock reference value in order to guarantee that all of the imposed demand will be fulfilled from the readily available resources at the distribution center thus ensuring the maximum service level.

Theorem 1: The ordering signal generated by controller (9) applied to system (3) satisfies the following inequalities

$$K\frac{\sigma y_{ref}}{\sigma+K} \leq u(t) \leq K y_{ref}. \quad (10)$$

Moreover, there exists a time instant t_0 such that for any $t \geq t_0$

$$u(t) \leq K\frac{\sigma y_{ref} + d_{\max}}{\sigma+K}. \quad (11)$$

Proof: Substituting (5) into (9) we get

$$u(t) = K\left[y_{ref} - \int_0^t e^{-\sigma(t-\tau)}u(\tau)\,d\tau + \int_0^t e^{-\sigma(t-\tau)}h(\tau)\,d\tau\right]. \quad (12)$$

Consequently,

$$\begin{aligned}\dot{u} &= -K\frac{d}{dt}\left\{e^{-\sigma t}\int_0^t e^{\sigma\tau}\left[u(\tau)-h(\tau)\right]d\tau\right\} \\ &= K\left\{\sigma e^{-\sigma t}\int_0^t e^{\sigma\tau}\left[u(\tau)-h(\tau)\right]d\tau - e^{-\sigma t}e^{\sigma t}\left[u(t)-h(t)\right]\right\} \\ &= K\left\{\sigma\int_0^t e^{-\sigma(t-\tau)}\left[u(\tau)-h(\tau)\right]d\tau - \left[u(t)-h(t)\right]\right\}.\end{aligned} \quad (13)$$

It follows from (12) that

$$\sigma K\int_0^t e^{-\sigma(t-\tau)}\left[u(\tau)-h(\tau)\right]d\tau = \sigma\left[Ky_{ref} - u(t)\right]. \quad (14)$$

Hence, we can rewrite (13) as

$$\begin{aligned}\dot{u} &= \sigma\left[Ky_{ref} - u(t)\right] - K\left[u(t)-h(t)\right] \\ &= \sigma K y_{ref} - (\sigma+K)u(t) + Kh(t).\end{aligned} \quad (15)$$

Investigating $\dot{u} = 0$ we get

$$u(t) = K\frac{\sigma y_{ref} + h(t)}{\sigma+K}. \quad (16)$$

According to constraint (1) the minimum satisfied demand equals zero. At the initial time $u(0) = Ky_{ref} > 0$. Therefore, since $0 \leq \sigma < 1$ and $h(\cdot) \geq 0$, we get from (16) that $u(\cdot)$ de-

creases as long as it is bigger than $K[\sigma y_{ref} + h(\cdot)]/(\sigma + K)$, and it never falls below $K\sigma y_{ref}/(\sigma + K)$. Moreover, there exists a time instant t_0 when $u(\cdot)$ reaches the level of $K[\sigma y_{ref} + d_{\max}]/(\sigma + K)$ for the first time. Since $h(\cdot) \le d_{\max}$, we get from (16) that for all $t \ge t_0$

$$u(t) \le K\frac{\sigma y_{ref} + d_{\max}}{\sigma + K}.$$

This conclusion ends the proof. □

Theorem 2: If policy (9) is applied to system (3), then the on-hand stock at the distribution center never exceeds the level of y_{ref} for $\sigma = 0$ and

$$y_{\max} = \frac{K}{\sigma + K}\left[y_{ref} + \frac{d_{\max}}{\sigma}\left(1 - e^{-\sigma L_p}\right)\right] \quad \text{for } \sigma > 0. \tag{17}$$

Proof: Applying (12) to the stock balance equation (3) we get

$$\begin{aligned}\dot{y} &= -\sigma y(t) + Ky_{ref} \\ &\quad - K\left[\int_0^{t-L_p} e^{-\sigma(t-L_p-\tau)}u(\tau)\,d\tau - \int_0^{t} e^{-\sigma(t-\tau)}h(\tau)\,d\tau\right] \\ &\quad + K\int_0^{t-L_p} e^{-\sigma(t-L_p-\tau)}h(\tau)\,d\tau - K\int_0^{t-L_p} e^{-\sigma(t-\tau)}h(\tau)\,d\tau \\ &\quad - K\int_{t-L_p}^{t} e^{-\sigma(t-\tau)}h(\tau)\,d\tau - h(t).\end{aligned} \tag{18}$$

Using (5) we can notice that the term in the square brackets in (18) actually equals $y(t)$. Consequently, we have

$$\begin{aligned}\dot{y} &= Ky_{ref} - (\sigma + K)y(t) - h(t) \\ &\quad + K\left(e^{\sigma L_p} - 1\right)\int_0^{t-L_p} e^{-\sigma(t-\tau)}h(\tau)\,d\tau - K\int_{t-L_p}^{t} e^{-\sigma(t-\tau)}h(\tau)\,d\tau.\end{aligned} \tag{19}$$

Closer investigation of $\dot{y} = 0$ leads to

$$\begin{aligned}y(t) &= \frac{Ky_{ref}}{\sigma + K} + \frac{K}{\sigma + K}\left(e^{\sigma L_p} - 1\right)\int_0^{t-L_p} e^{-\sigma(t-\tau)}h(\tau)\,d\tau \\ &\quad - \frac{K}{\sigma + K}\left[\int_{t-L_p}^{t} e^{-\sigma(t-\tau)}h(\tau)\,d\tau + \frac{h(t)}{K}\right].\end{aligned} \tag{20}$$

It follows from (20) that since $K > 0$, $\sigma \ge 0$, and $h(\cdot) \ge 0$, the biggest value of $y(\cdot)$ is expected when $h(\tau) = d_{\max}$ for $\tau \le t - L_p$ and $h(\tau) = 0$ in the interval $(t - L_p, t]$. We get immediately from (20) that for $\sigma = 0$ (the case of nondeteriorating stock) $y(t) \le y_{ref}$. Evaluating the first integral in (20) for $\sigma > 0$ we obtain

$$\begin{aligned}&\int_0^{t-L_p} e^{-\sigma(t-\tau)}h(\tau)\,d\tau \le d_{\max}\int_0^{t-L_p} e^{-\sigma(t-\tau)}d\tau \\ &= d_{\max}e^{-\sigma t}\int_0^{t-L_p} e^{\sigma\tau}d\tau = d_{\max}\frac{e^{-\sigma t}}{\sigma}\left(e^{\sigma\tau}\right)\Big|_0^{t-L_p} \\ &= d_{\max}\frac{e^{-\sigma t}}{\sigma}\left[e^{\sigma(t-L_p)} - 1\right] = \frac{d_{\max}}{\sigma}\left[e^{-\sigma L_p} - e^{-\sigma t}\right] \le \frac{d_{\max}}{\sigma}e^{-\sigma L_p}.\end{aligned} \tag{21}$$

Consequently, applying (21) to (20), we arrive at

$$y(t) \le \frac{K}{\sigma + K}\left[y_{ref} + \frac{d_{\max}}{\sigma}\left(1 - e^{-\sigma L_p}\right)\right]. \tag{22}$$

This ends the proof. □

It follows from Theorem 2 that if the warehouse of size $y_{\max}$ specified by (17) is assigned at the distribution center, then all the incoming shipments can be stored locally, and any cost associated with emergency storage is eliminated. Apart from the efficient warehouse space management, a successful inventory control strategy in modern supply chain is expected to achieve a high level of demand satisfaction. The proposition formulated below shows how the reference stock level should be selected so that $y(t) > 0$, which implies that all of the demand imposed on the distribution center is satisfied from the readily available resources.

Theorem 3: If policy (9) is applied to system (3), and the reference stock level is selected as

$$y_{ref} > d_{\max}\left(L_p + 1/K\right) \quad \text{for } \sigma = 0, \tag{23}$$

$$y_{ref} > d_{\max}\left[\left(1 - e^{-\sigma L_p}\right)/\sigma + 1/K\right] \quad \text{for } \sigma > 0, \tag{24}$$

then the on-hand stock level at the distribution center is strictly positive for any $t > L_p$.

Proof: Note that $e^{\sigma L_p} - 1 > 0$. Hence, considering (1) and (20), we can expect the smallest on-hand stock level in the circumstances when $h(\tau) = 0$ for $\tau \le t - L_p$ and $h(\tau) = d_{\max}$ for τ belonging to the interval $(t - L_p, t]$. The warehouse is empty for $t \le L_p$. In the case of the system with nonndeteriorating stock ($\sigma = 0$) we get immediately from (20)

$$y(t) \ge y_{ref} - d_{\max}\left(L_p + 1/K\right). \tag{25}$$

Thus, based on assumption (23) we have $y(t) > 0$ for $\sigma = 0$. Evaluating the second integral in (20) for $t > L_p$ in the case when $h(t) = d_{\max}$ and $\sigma > 0$, we obtain

$$\begin{aligned}&\int_{t-L_p}^{t} e^{-\sigma(t-\tau)}h(\tau)\,d\tau \le d_{\max}\int_{t-L_p}^{t} e^{-\sigma(t-\tau)}d\tau \\ &= d_{\max}e^{-\sigma t}\int_{t-L_p}^{t} e^{\sigma\tau}d\tau = d_{\max}\frac{e^{-\sigma t}}{\sigma}e^{\sigma\tau}\Big|_{t-L_p}^{t} \\ &= d_{\max}\frac{e^{-\sigma t}}{\sigma}\left[e^{\sigma t} - e^{\sigma(t-L_p)}\right] = \frac{d_{\max}}{\sigma}\left[1 - e^{-\sigma L_p}\right].\end{aligned} \tag{26}$$

Applying (26) to (20), we get the on-hand stock level $y(\cdot)$ at the instant when it is minimal

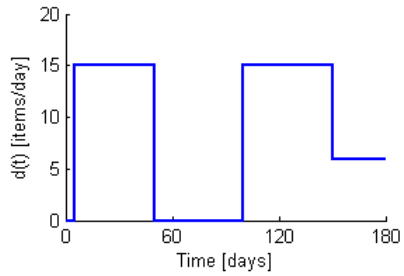


Fig. 4. Market demand.

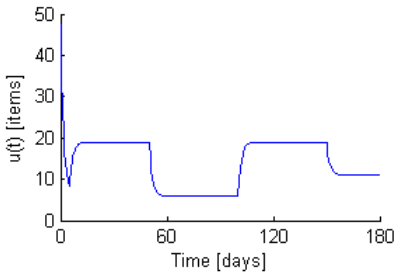


Fig. 5. Orders placed at the supplier.

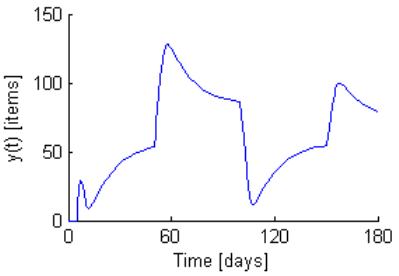


Fig. 6. On-hand stock level.

$$y(t) \geq \frac{K}{\sigma + K}\left[y_{ref} - \frac{d_{\max}}{\sigma}\left(1 - e^{-\sigma L_p}\right) - \frac{d_{\max}}{K} \right]. \quad (27)$$

If the reference stock level is adjusted according to (24), then using (27) one may conclude that

$$y(t) \geq \frac{K}{\sigma + K}\left\{ y_{ref} - d_{\max}\left[\left(1 - e^{-\sigma L_p}\right)/\sigma + 1/K \right] \right\} > 0. \quad (28)$$

This completes the proof. □

Remark: It follows from Theorem 1 that the controller generates ordering signal that is always nonnegative and bounded, which makes the considered system positive. In addition, if assumptions of Theorem 3 are fulfilled, then $h(t) = d(t)$, and the plant remains in the linear region for arbitrary demand satisfying condition (1). Considering the responses with respect to the reference input (7) and with respect to the disturbance (8) we get the overall system transfer function

$$Y(s) = \frac{Ke^{-L_p s}}{s + \sigma + K}\frac{y_{ref}}{s} - \left(\frac{1}{s + \sigma} - \frac{K}{s + \sigma + K}e^{-L_p s} \right) D(s). \quad (29)$$

Consequently, applying the final value theorem we get the steady-state stock level y_{ss} (the stock level in the presence of the steady-state demand $d_{ss} > 0$)

$$y_{ss} = \lim_{s \to 0} sY(s) = \begin{cases} y_{ref} - \dfrac{1 + KL_p}{K} d_{ss} & \text{for } \sigma = 0, \\ \dfrac{Ky_{ref}}{\sigma + K} - \left(\dfrac{1}{\sigma} - \dfrac{K}{\sigma + K} \right) d_{ss} & \text{for } \sigma > 0. \end{cases} \quad (30)$$

Equation (30) indicates that a finite steady-state error will be present at the output when a proportional control law is chosen as the primary controller $C(s)$ in (6). Typically in engineering systems, this error would need to be reduced (or eliminated), for instance by introducing a proportional-integral controller in place of the proportional one. Also a feed-forward term could be applied to compensate the effects of disturbance. However, in the considered application, y_{ref} can be assigned an arbitrary value and any steady-state error can be tolerated. What is important from the practical point of view when studying inventory control problems is the size of the required storage space and demand utilization. Theorems 2 and 3 show precisely how much storage space should be provided to accommodate all the incoming shipments (relation (17)), and how to select y_{ref} to guarantee that all the sales are realized from the readily available resources (inequalities (23) and (24)).

4. NUMERICAL EXAMPLE

The properties of the designed policy (9) are verified in simulations conducted for the model of perishable inventory system described in Section 2. The system parameters are set in the following way: lead-time $L_p = 5$ days, inventory decay factor $\sigma = 0.07$ day^{-1}, and the maximum daily demand at the distribution center $d_{\max} = 15$ items/day. The actual demand follows the pattern illustrated in Fig. 4, which reflects abrupt seasonal changes in a half-year trend. The controller gain is adjusted to $K = 0.5$. In order to ensure full demand satisfaction, the stock reference level is set according to the guidelines of Theorem 3 as $y_{ref} = 95 > 93$ items. This results in the required storage space calculated according to (17) $y_{\max} = 139$ items.

The orders generated by controller (9) in response to the demand pattern from Fig. 4 are shown in Fig. 5, and the resultant on-hand stock in Fig. 6. We can see from the graphs depicted in Fig. 5 that the proposed controller quickly responds to the sudden changes in the demand trend without oscillations or overshoots in the ordering signal. For $t \geq t_0 = 2$ days the order quantity remains in the interval [5.83, 19.01 items/day], precisely as dictated by Theorem 1. The knowledge about the range of order changes helps in establishing long-term relationship between the distribution center and the supplier, and facilitates capacity planning down the supply chain (at the supplier and its subcontractors). We can see from Fig. 6 that the stock level does not increase beyond $y_{\max} = 139$ items, which means that the assigned warehouse capacity is sufficient to store the goods at the distribution center at all times. Moreover, the on-hand stock never falls to zero after the initial phase which implies full demand satisfaction and 100% service level.

5. CONCLUSIONS

In this work we designed a new inventory management policy for continuous-time inventory systems with perishable goods. The proposed policy employs the Smith predictor for compensating the adverse effects of order procurement delay. As a result, the system stability is guaranteed for arbitrary delay and any bounded demand pattern. The ordering signal generated by the designed policy smoothly adapts to the demand changes, and thus it is easy to follow by the supplier.

The ordering signal is proved to remain finite and always nonnegative, which is crucial for the practical implementation of any inventory management scheme. It is also demonstrated in the paper that the stock level resulting from the application of the proposed policy does not increase beyond the precisely determined warehouse capacity, which eliminates the need for costly emergency storage and facilitates capacity planning at the distribution center. Finally, it is shown how to select controller parameters to achieve full satisfaction of the unknown market demand.

ACKNOWLEDGMENT

This work was financed by the Polish State budget in the years 2010–2012 as a research project N N514 108638 "Application of regulation theory methods to the control of logistic processes". P. Ignaciuk gratefully acknowledges financial support provided by the Foundation for Polish Science (FNP). He is also a scholarship holder of the project entitled "Innovative Education without Limits – Integrated Progress of the Technical University of Łódź" supported by the European Social Fund.

REFERENCES

Blanchini, F., Pesenti, R, Rinaldi, F., and Ukovich, W. (2003). Feedback control of production-distribution systems with unknown demand and delays. *IEEE Transactions on Robotics and Automation*, *16*(3), 313–317.

Boccadoro, M., Martinelli, F., and Valigi, P. (2008). Supply chain management by H-infinity control," *IEEE Transactions on Automation Science and Engineering*, *5*(4), 703–707.

Goyal, S.K., and Giri, B.C. (2001). Recent trends in modeling of deteriorating inventory. *European Journal of Operational Research*, *134*(1), 1–16.

Hoberg, K., Bradley, J.R., and Thonemann, U.W. (2007). Analyzing the effect of the inventory policy on order and inventory variability with linear control theory. *European Journal of Operational Research*, *176*(3), 1620–1642.

Ignaciuk, P., and Bartoszewicz, A. (2010). LQ optimal sliding mode supply policy for periodic review inventory systems," *IEEE Transactions on Automatic Control*, *55*(1), 269–274.

Ignaciuk, P., and Bartoszewicz, A. (2010). LQ optimal and reaching law based sliding modes for inventory management systems," *International Journal of Systems Science*, *41* (*in press*).

Karaesmen, I., Scheller-Wolf, A., and Deniz, B. (2008). Managing perishable and aging inventories: review and future research directions. In: Kempf, K., Keskinocak, P., and Uzsoy, R. (eds.). *Handbook of production planning*. Dordrecht: Kluwer.

Nahmias, S. (1982). Perishable inventory theory: a review. *Operations Research*, *30*(4), 680–708.

Ortega, M., and Lin, L. (2004). Control theory applications to the production-inventory problem: a review. *International Journal of Production Research*, *42*(11), 2303–2322.

Palmor, Z.J. (1996). Time-delay compensation – Smith predictor and its modifications. In Levine, W.S. (ed.) The Control Handbook, CRC Press, 224–237.

Rafaat, F. (1991). Survey of literature on continuously deteriorating inventory models. *Journal of the Operational Research Society*, *42*(1), 27–37.

Sarimveis, H., Patrinos, P., Tarantilis, C.D., and Kiranoudis, C.T. (2008). Dynamic modeling and control of supply chain systems: a review. *Computers & Operations Research*, *35*(11), 3530–3561.

Smith, O.J.C. (1959). A controller to overcome dead time. *ISA Journal*, *6*(2), 28–33.

Warburton, R.D.H. (2007). An optimal, potentially automatable ordering policy. *International Journal of Production Economics*, *107*(2), 483–495.

APPENDIX

We solve differential equation (3) with the initial conditions: $y(0) = 0$, and $u_R(t) = u(t - L_p) = 0$ for $t < L_p$. First we consider the homogeneous equation

$$\dot{y} + \sigma y(t) = 0, \tag{31}$$

which leads to

$$y(t) = y(0)e^{-\sigma t}. \tag{32}$$

In order to determine the nonhomogeneous solution we assume $y(t)$ in the following form

$$y(t) = q(t)e^{-\sigma t}, \tag{33}$$

where $q(t)$ is a function differentiable with respect to time. Differentiating both sides of (33) we obtain

$$\dot{y} = \dot{q}e^{-\sigma t} - \sigma q(t)e^{-\sigma t}. \tag{34}$$

Substituting (33) and (34) into (3), we get

$$\dot{q}e^{-\sigma t} = u_R(t) - h(t). \tag{35}$$

Solving (35) for $q(t)$ yields

$$q(t) = \int_0^t e^{\sigma\tau}\left[u_R(\tau) - h(\tau)\right]d\tau + C, \tag{36}$$

where C is the constant of integration. Substituting (36) into (33), we arrive at

$$\begin{aligned} y(t) &= \left\{\int_0^t e^{\sigma\tau}\left[u_R(\tau) - h(\tau)\right]d\tau + C\right\}e^{-\sigma t} \\ &= Ce^{-\sigma t} + \int_0^t e^{-\sigma(t-\tau)}\left[u_R(\tau) - h(\tau)\right]d\tau. \end{aligned} \tag{37}$$

Applying the initial condition $y(0) = 0$, we have $C = 0$, and

$$y(t) = \int_0^t e^{-\sigma(t-\tau)}\left[u_R(\tau) - h(\tau)\right]d\tau. \tag{38}$$

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Smoothing in Multiple Model Change Detection for Stochastic Systems ⋆

Ivo Punčochář * **Jindřich Duník** * **Miroslav Šimandl** *

* *Department of Cybernetics and Research Centre DAR, Faculty of Applied Sciences, University of West Bohemia, Univerzitní 8, 306 14 Plzeň, Czech Republic (e-mails: ivop@kky.zcu.cz (I. Punčochář), dunikj@kky.zcu.cz (J. Duník), simandl@kky.zcu.cz (M. Šimandl)).*

Abstract: The paper focuses on the application of smoothing in multiple model change detection for stochastic systems. In a typical multiple model change detection scheme, the decisions are made based on filtering estimates of an unmeasured state of an observed system. Since better estimates of the state lead to finer decisions, a higher quality of estimates is of great interest. The way to improve the decisions considered in this paper consists in deferring decisions and using more precise smoothing estimates instead of the filtering ones. As a result, the decisions of a higher quality are obtained at the cost of delaying that decisions. The approach introduces a new level of the compromise between the quality of decisions and the delay for detection that is inherent in all change detection methods.

Keywords: Stochastic systems, detection systems, change detection, optimal estimation, smoothing.

1. INTRODUCTION

The problem of change detection arises in many application areas ranging from quality control to fault detection. Recently, it has received a great deal of attention because of increasing requirements on safety, reliability, and low maintenance costs. The fundamental goal is to find a detector that processes available measurements and generates decisions about changes in an observed system.

Most change detection methods employ a model of the system to perform the detection. According to the complexity of these models, the change detection methods can be divided into two groups: signal-based methods and model-based methods. While the signal-based methods (Isermann, 1984) use only simple assumptions about the measured signals, the model-based methods (Jones, 1973; Basseville and Nikiforov, 1993) utilize more complex models. The model-based methods make it possible to detect a wider range of changes in the system because more detailed information about system behavior is used. A detector usually consists of a residual generator and a decision generator that are connected in cascade. The residual generator processes the measurements and generates residual signals that are close to zero before a change and significantly deviate from zero when the change occurs. The decision generator analyzes, mostly statistically, these residual signals and generates decisions about changes.

Recently, increasing attention has been drawn to more advanced methods called active change detection methods (Blackmore and Williams, 2006; Šimandl and Punčochář, 2009). Contrary to the above mentioned change detection methods, that can be denoted as passive, the active change detection methods improve the quality of decisions by designing and applying a suitable input signal to the observed system. The difference between the passive and active change detection methods is illustrated in Fig. 1. The passive detector **PD** uses the measurements $\mathbf{z}_k$ obtained from the system **S** for finding the decision $\mathbf{d}_k$, whereas the active detector **AD** uses the output $\mathbf{y}_k$ and generates, in addition to the decision $\mathbf{d}_k$, the input $\mathbf{u}_k$ for improving the quality of the decisions. Instead of pursuing these advanced active change detection methods, this paper focuses on pointing out a possible improvement of passive multiple model change detection for stochastic systems.

The multiple model approach represents a favorite and successful modeling tool for describing systems that can undergo abrupt changes. Besides change detection, the multiple model approach is widely used in application areas such as fault detection (Gustafsson, 2000), state estimation (Ackerson and Fu, 1970; Blom and Bar-Shalom, 1988; Šimandl and Královec, 2000), adaptive control (Athans et al., 2006) and target tracking (Bar-Shalom et al., 2001). Within change detection, the individual models describe different modes of system behavior and the aim is to decide which of the models describes the system behavior best. Typically, the decision concerning the current time step based on filtering estimates is required. However, there are situations in which it pays off to defer the decision by a few time steps in favor of increasing the quality of that decision. A higher quality of the deferred decision follows from using smoothing estimates instead of the filtering ones. Since the deferred decision at the current time step relates to a past time step, the range of

⋆ This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, project No. 1M0572, and by the Czech Science Foundation, project No. GA102/08/0442.

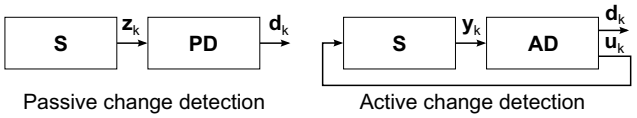


Fig. 1. Schema of passive and active change detection

applications is somewhat limited. Some typical examples in which it is reasonable to use deferred decisions are:

- A detector with event-driven smoothing. On-line applications in which change detection is based on the filtering estimates. The smoothing estimates are used for a more precise identification of the actual time of the change whenever the change is detected.
- A detector with deferred decisions. On-line applications in which it pays off to respond to a change with a delay. It typically applies to systems of which the individual components are spatially distributed.
- Batch data analysis. Off-line applications in which historical data are processed and analyzed in batch.

The goal of the paper is to design a smoothing based change detector in the multi model framework for stochastic systems and demonstrate the benefits that follow from employing smoothing in multiple model change detection.

The paper is organized as follows. The multiple model description of the observed system, a formal description of the optimal passive detector, and a design criterion are introduced in Section 2. Section 3 reviews information processing strategies and deals with the derivation of the optimal passive detector law. Section 4 focuses on the smoothing algorithm that is used for improving multiple model change detection. The presentation of an optimal filtering algorithm, which is an inherent part of smoothing, is followed by the optimal smoothing algorithm itself and closed with a short discussion. The results are illustrated by means of a numerical example in Section 5.

2. PROBLEM FORMULATION

Consider the system **S** described at each time step $k \in \mathcal{T} = \{0, 1, \ldots, F\}$ by the jump Markov linear Gaussian model

$$\mathbf{x}_{k+1} = \mathbf{A}_{\mu_k}\mathbf{x}_k + \mathbf{G}_{\mu_k}\mathbf{w}_k, \quad (1)$$
$$\mathbf{z}_k = \mathbf{C}_{\mu_k}\mathbf{x}_k + \mathbf{H}_{\mu_k}\mathbf{v}_k, \quad (2)$$

where $\mathbf{z}_k \in \mathbb{R}^{n_z}$ denotes an output, $\bar{\mathbf{x}}_k = \left[\mathbf{x}_k^T, \mu_k\right]^T$ is an unmeasured state of the system consisting of the variables $\mathbf{x}_k \in \mathbb{R}^{n_x}$ and $\mu_k \in \mathcal{M} = \{1, 2, \ldots, N\}$. The vector $\mathbf{x}_k$ is continuous in values and represents a common state of individual linear Gaussian models. The scalar μ_k is the index of linear Gaussian model that represents the system at the time step k. It is assumed that the number of models $N \geq 2$ and the matrices $\mathbf{A}_{\mu_k}$, $\mathbf{C}_{\mu_k}$, $\mathbf{G}_{\mu_k}$, $\mathbf{H}_{\mu_k}$ of appropriate dimensions are known. The switching between the models is described by known transition probabilities $\pi_{i,j} = P(\mu_{k+1} = j | \mu_k = i)$ for all $i, j \in \mathcal{M}$. The noises $\mathbf{w}_k \in \mathbb{R}^{n_x}$ and $\mathbf{v}_k \in \mathbb{R}^{n_z}$ are mutually independent zero-mean white Gaussian noises with identity covariance matrices (i.e. $\mathcal{N}\{\mathbf{0}, \mathbf{E}_n\}$). The initial condition $\mathbf{x}_0$ is described by the Gaussian probability density function $\mathcal{N}\{\hat{\mathbf{x}}_{0|-1}, \mathbf{P}_{x,0|-1}\}$ and the variable μ_0 is described by the probability $P(\mu_0)$. The variables $\mathbf{x}_0$ and μ_0 are mutually independent and also independent of the noises.

The passive detector **PD** is a dynamic system that generates the decisions based on all information available at the current time step. Such a system can formally be described at each time step $k \in \mathcal{T}$ by

$$d_k = \sigma_k\left(\mathbf{I}_0^k\right), \quad (3)$$

where $d_k \in \mathcal{M}$ denotes a decision, $\sigma_k(\mathbf{I}_0^k)$ is an unknown function and $\mathbf{I}_0^{k\,T} = [\mathbf{z}_0^{k\,T}, d_0^{k-1\,T}]$ is an information vector containing all information that has been received up the current time step k. Note that the notation $\mathbf{z}_i^j = [\mathbf{z}_i^T, \mathbf{z}_{i+1}^T, \ldots, \mathbf{z}_j^T]^T$ is used for expressing the time sequence of variables or functions. If $j < i$ then the result of the indexing operation is the empty sequence.

The aim is to find a sequence of the functions σ_0^F such that the passive detector provides the highest quality of decisions, which is usually judged by means of the total cost caused by wrong decisions. As it is discussed in the introduction, there are some situations in which the decision at the current time step should inform which model was valid at a past time step. To include these situations into the problem formulation, the optimal passive generator is designed by minimizing the criterion

$$J(\sigma_\ell^F) = \mathrm{E}\left\{\sum_{k=\ell}^{F} L_k^{\mathrm{d}}(\mu_{k-\ell}, d_k)\right\}, \quad (4)$$

where $\mathrm{E}\{\cdot\}$ denotes the expectation operator with respect to all random variables, $L_k^{\mathrm{d}}(\mu_{k-\ell}, d_k)$ is a real-valued non-negative cost function representing a detection objective, and $\ell \geq 0$ is a lag (i.e. a delay between the decision d_k and the time step to which it is referring). Note that for the case $\ell > 0$, the optimal passive detector is not defined for the time steps $k = 0, \ldots, \ell - 1$.

3. OPTIMAL DETECTOR DESIGN

3.1 Information processing strategies

The problem introduced in the previous section represents a dynamic optimization problem. As it has been discussed for the optimal stochastic control (Bar-Shalom and Tse, 1974) or the optimal active change detection and control (Šimandl and Punčochář, 2009), such an optimization problem can be solved using three basic information processing strategies (IPS's). The open-loop (OL) IPS assumes that only an a priori information and no future information will be used. The open-loop feedback (OLF) IPS assumes that an additional information will be available in a priori unknown time steps and this information will be used together with the a priori information. The closed-loop (CL) IPS supposes that an additional information will be received and utilized at all future time steps.

It generally holds that $J^{\mathrm{OL}} \geq J^{\mathrm{OLF}} \geq J^{\mathrm{CL}}$, where the superscript denotes the particular IPS. For the problem considered in this paper, the use of the OLF IPS leads to the same result as the utilization of the CL IPS would lead (i.e. $J^{\mathrm{OL}} \geq J^{\mathrm{OLF}} = J^{\mathrm{CL}}$). To show the equivalence between the CL IPS and the OLF IPS, the CL IPS will be used for deriving the optimal passive detector law and the result will be confronted with the passive detector obtained using the OLF IPS.

3.2 Derivation of optimal passive detector law

The use of the CL IPS for solving the given dynamic optimization problem leads to dynamic programming. Thus the problem is considered backward in time from the final time step F and the optimal passive detector can be obtained by solving the following backward recursive equation for time steps $k = F, F-1, \ldots, \ell$

$$V_k^* \left(\mathbf{I}_0^k\right) = \min_{d_k \in \mathcal{M}} \mathrm{E}\left\{L_k^{\mathrm{d}}\left(\mu_{k-\ell}, d_k\right) + V_{k+1}^*\left(\mathbf{I}_0^{k+1}\right) | \mathbf{I}_0^k, d_k\right\}, \quad (5)$$

where $\mathrm{E}\{\cdot|\cdot\}$ denotes the conditional expectation operator and $V_k^*(\mathbf{I}_0^k)$ is the Bellman function that expresses the minimum expected cost incurred from the current time step k to the final time step F. The initial condition for the backward recursive equation is $V_{F+1}^* = 0$ and the value of the criterion (4) can be expressed as $J^{\mathrm{CL}} = J\left(\sigma_\ell^{F*}\right) = \mathrm{E}\left\{V_\ell^*\left(\mathbf{I}_0^\ell\right)\right\}$.

The conditional probability $P\left(\mu_{k-\ell}|\mathbf{I}_0^k, d_k\right)$ and the conditional probability density function (pdf) $p\left(\mathbf{z}_{k+1}|\mathbf{I}_0^k, d_k\right)$, needed for evaluation of the conditional expectations in the backward recursive equation (5), can be computed using a nonlinear estimator. The exact nonlinear filter and smoother for multiple models are presented in Section 4. It can easily be shown that the probability $P\left(\mu_{k-\ell}|\mathbf{I}_0^k, d_k\right)$ and the pdf $p\left(\mathbf{z}_{k+1}|\mathbf{I}_0^k, d_k\right)$ satisfy the identities

$$P\left(\mu_{k-\ell}|\mathbf{I}_0^k, d_k\right) = P\left(\mu_{k-\ell}|\mathbf{z}_0^k\right), \quad (6)$$

$$p\left(\mathbf{z}_{k+1}|\mathbf{I}_0^k, d_k\right) = p\left(\mathbf{z}_{k+1}|\mathbf{z}_0^k\right). \quad (7)$$

Note that these identities just reflect the fact that the decisions d_0^k can not bring any additional information since they are deterministic functions of the measurements $\mathbf{z}_0^k$.

The identities (6) and (7) can be useful for rewriting the backward recursive equation (5) into a simpler form. The Bellman function at time step $k = F$ is

$$V_F^*\left(\mathbf{I}_0^F\right) = \min_{d_F \in \mathcal{M}} \mathrm{E}\left\{L_F^{\mathrm{d}}\left(\mu_{F-\ell}, d_F\right) | \mathbf{I}_0^F, d_F\right\}. \quad (8)$$

Applying (6) the Bellman function $V_F^*\left(\mathbf{I}_0^F\right)$ takes the form

$$V_F^*\left(\mathbf{I}_0^F\right) = \min_{d_F \in \mathcal{M}} \mathrm{E}\left\{L_F^{\mathrm{d}}\left(\mu_{F-\ell}, d_F\right) | \mathbf{z}_0^F, d_F\right\}, \quad (9)$$

from which it can be seen that the right-hand side is independent of the decisions d_0^F and thus it holds that $V_F^*(\mathbf{I}_0^F) = V_F^*(\mathbf{z}_0^F)$.

Now, if it is assumed that $V_{k+1}^*(\mathbf{I}_0^{k+1}) = V_{k+1}^*(\mathbf{z}_0^{k+1})$ at a time step $k+1$, then the Bellman function at a time step k can be written as

$$V_k^*\left(\mathbf{I}_0^k\right) = \min_{d_k \in \mathcal{M}} \mathrm{E}\{L_k^{\mathrm{d}}\left(\mu_{k-\ell}, d_k\right) + V_{k+1}^*\left(\mathbf{z}_0^{k+1}\right) | \mathbf{I}_0^k, d_k\}. \quad (10)$$

Applying (6) and (7) it can be seen that the Bellman function $V_k^*\left(\mathbf{I}_0^k\right)$ is independent of the decisions d_0^k (i.e. $V_k^*(\mathbf{I}_0^k) = V_k^*(\mathbf{z}_0^k)$). Thus, the backward recursive equation can be written at each time step $k \in \mathcal{T}$ as

$$V_k^*\left(\mathbf{z}_0^k\right) = \min_{d_k \in \mathcal{M}} \mathrm{E}\left\{L_k^{\mathrm{d}}\left(\mu_{k-\ell}, d_k\right) | \mathbf{z}_0^k, d_k\right\} + \mathrm{E}\left\{V_{k+1}^*\left(\mathbf{z}_0^{k+1}\right) | \mathbf{z}_0^k, \right\}, \quad (11)$$

and the optimal decision d_k^* is given by

$$d_k^* = \sigma_k^*\left(\mathbf{z}_0^k\right) = \arg\min_{d_k \in \mathcal{M}} \mathrm{E}\left\{L_k^{\mathrm{d}}\left(\mu_{k-\ell}, d_k\right) | \mathbf{z}_0^k, d_k\right\}, \quad (12)$$

where $\sigma_k^*\left(\mathbf{z}_0^k\right)$ is a function describing the optimal passive detector. Note that the Bellman functions cannot actually be computed analytically due to intractable conditional expectations. However, it does not hamper the solution since the optimal decisions can be computed without the knowledge of the Bellman functions. The reason is that the decisions do not influence the future data.

Now, a detector based on the OLF IPS will be derived. This IPS uses all available information as if no more information will be received in the future (i.e. the OL IPS is used for the future time steps). Therefore, the following optimization problem has to be solved at each time step $k \in \mathcal{T}$

$$\min_{d_k^F} \mathrm{E}\left\{\sum_{i=k}^{F} L_i^{\mathrm{d}}\left(\mu_{i-\ell}, d_i\right) | \mathbf{z}_0^k, d_k^F\right\}.$$

Since each decision d_i influences only the cost function $L_i^{\mathrm{d}}(\cdot,\cdot)$, this minimization problem can be recast as

$$\sum_{i=k}^{F} \min_{d_i} \sum_{\mu_{k-\ell}^{i-\ell}} L_i^{\mathrm{d}}\left(\mu_{i-\ell}, d_i\right) P\left(\mu_{k-\ell}^{i-\ell}|\mathbf{z}_0^k\right).$$

From this expression, it can be seen that the optimal decision d_k^* is given by

$$d_k^* = \sigma_k^*\left(\mathbf{z}_0^k\right) = \arg\min_{d_k \in \mathcal{M}} \mathrm{E}\left\{L_k^{\mathrm{d}}\left(\mu_{k-\ell}, d_k\right) | \mathbf{z}_0^k, d_k\right\}, \quad (13)$$

which corresponds to an intuitively expected result. Since (13) is the same as (12), it is clear that the use of the CL IPS brings no improvement over the OLF IPS.

4. MULTIPLE MODEL CHANGE DETECTION IMPROVEMENT BY SMOOTHING

This section deals with the estimation algorithms that provide the probability needed in the optimal passive detector. First, the filtering algorithm is presented because it is an inherent part of the smoothing algorithm and employed whatever the lag ℓ is. Then the smoothing algorithm is provided and the section is concluded with a discusion.

4.1 Optimal filtering in multiple model change detection

This subsection summarizes the filtering algorithm for multiple Gaussian linear models which provides the conditional probability $P(\mu_0^k|\mathbf{z}_0^k)$ together with the conditional pdf's $p(\mathbf{x}_k|\mathbf{z}_0^k)$ and $p(\mathbf{z}_{k+1}|\mathbf{z}_0^k)$.

If the lag ℓ is zero, the conditional probability $P\left(\mu_k|\mathbf{z}_0^k\right)$ is needed to make the optimal decision. This conditional probability is given by the following marginalization

$$P\left(\mu_k|\mathbf{z}_0^k\right) = \sum_{\mu_0^{k-1}} P\left(\mu_0^k|\mathbf{z}_0^k\right), \quad (14)$$

where $\sum_{\mu_0^i}$ denotes the sum over all model sequences μ_0^i. The conditional probability $P(\mu_0^k|\mathbf{z}_0^k)$ can be evaluated recursively according to

$$P\left(\mu_0^k|\mathbf{z}_0^k\right) = \frac{p\left(\mathbf{z}_k|\mathbf{z}_0^{k-1}, \mu_0^k\right) P\left(\mu_0^k|\mathbf{z}_0^{k-1}\right)}{p\left(\mathbf{z}_k|\mathbf{z}_0^{k-1}\right)}, \quad (15)$$

where the predictive probability $P(\mu_0^k|\mathbf{z}_0^{k-1})$ can easily be computed as

$$P\left(\mu_0^k|\mathbf{z}_0^{k-1}\right) = P\left(\mu_k|\mu_{k-1}\right) P\left(\mu_0^{k-1}|\mathbf{z}_0^{k-1}\right), \quad (16)$$

and the predictive pdf $p(\mathbf{z}_k|\mathbf{z}_0^{k-1})$ independent of the model sequence μ_0^k is given by the relation

$$p\left(\mathbf{z}_k|\mathbf{z}_0^{k-1}\right) = \sum_{\mu_0^k} p\left(\mathbf{z}_k|\mathbf{z}_0^{k-1}, \mu_0^k\right) P\left(\mu_0^k|\mathbf{z}_0^{k-1}\right). \quad (17)$$

Given the model sequence μ_0^k, the system can be described by a t-variant linear Gaussian model, and thus the Gaussian pdf $p(\mathbf{z}_k|\mathbf{z}_0^{k-1}, \mu_0^k) = \mathcal{N}\{\hat{\mathbf{z}}_{k|k-1}\left(\mu_0^k\right), \mathbf{P}_{z,k|k-1}\left(\mu_0^k\right)\}$ can be obtained from the Kalman filter. The mean value $\hat{\mathbf{z}}_{k|k-1}\left(\mu_0^k\right)$ and the covariance matrix $\mathbf{P}_{z,k|k-1}\left(\mu_0^k\right)$ are given as

$$\hat{\mathbf{z}}_{k|k-1}\left(\mu_0^k\right) = \mathbf{C}_{\mu_k}\hat{\mathbf{x}}_{k|k-1}\left(\mu_0^{k-1}\right), \quad (18)$$

$$\mathbf{P}_{z,k|k-1}\left(\mu_0^k\right) = \mathbf{C}_{\mu_k}\mathbf{P}_{x,k|k-1}\left(\mu_0^{k-1}\right)\mathbf{C}_{\mu_k}^T + \mathbf{H}_{\mu_k}\mathbf{H}_{\mu_k}^T, \quad (19)$$

where the mean value $\hat{\mathbf{x}}_{k|k-1}(\mu_0^{k-1})$ and the covariance matrix $\mathbf{P}_{x,k|k-1}(\mu_0^{k-1})$ are the first and the second moment of the Gaussian predictive pdf $p(\mathbf{x}_k|\mathbf{z}_0^{k-1}, \mu_0^{k-1}) = \mathcal{N}\{\hat{\mathbf{x}}_{k|k-1}(\mu_0^{k-1}), \mathbf{P}_{x,k|k-1}(\mu_0^{k-1})\}$, respectively.

The predictive mean value and covariance matrix are given by the initial conditions at the zero time step, and at other time steps, they are computed within the predictive step of the Kalman filter according to the relations

$$\hat{\mathbf{x}}_{k+1|k}\left(\mu_0^k\right) = \mathbf{A}_{\mu_k}\hat{\mathbf{x}}_{k|k}\left(\mu_0^k\right), \quad (20)$$

$$\mathbf{P}_{x,k+1|k}\left(\mu_0^k\right) = \mathbf{A}_{\mu_k}\mathbf{P}_{x,k|k}\left(\mu_0^k\right)\mathbf{A}_{\mu_k}^T + \mathbf{G}_{\mu_k}\mathbf{G}_{\mu_k}^T, \quad (21)$$

where $\hat{\mathbf{x}}_{k|k}(\mu_0^k)$ and $\mathbf{P}_{x,k|k}(\mu_0^k)$ are the filtering mean value and covariance matrix, respectively. These are obtained within the filtering step of the Kalman filter by evaluating the relations

$$\hat{\mathbf{x}}_{k|k}\left(\mu_0^k\right) = \hat{\mathbf{x}}_{k|k-1}\left(\mu_0^{k-1}\right) + \mathbf{K}^{\mathrm{F}}\left(\mu_0^k\right)\left[\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}\left(\mu_0^{k-1}\right)\right], \quad (22)$$

$$\mathbf{P}_{x,k|k}\left(\mu_0^k\right) = \mathbf{P}_{x,k|k-1}\left(\mu_0^{k-1}\right) - \mathbf{K}^{\mathrm{F}}\left(\mu_0^k\right)\mathbf{C}_{\mu_k}\mathbf{P}_{x,k|k-1}\left(\mu_0^{k-1}\right), \quad (23)$$

where the Kalman gain $\mathbf{K}^{\mathrm{F}}\left(\mu_0^k\right)$ is given by

$$\mathbf{K}^{\mathrm{F}}\left(\mu_0^k\right) = \mathbf{P}_{x,k|k-1}\left(\mu_0^{k-1}\right)\mathbf{C}_{\mu_k}^T\mathbf{P}_{z,k|k-1}\left(\mu_0^k\right)^{-1}. \quad (24)$$

4.2 Optimal smoothing in multiple model change detection

The filtering estimate given by the conditional pdf $p(\mathbf{x}_k|\mathbf{z}_0^k)$ and the probability $P(\mu_k|\mathbf{z}_0^k)$ represents the optimal estimate of the state $\bar{\mathbf{x}}_k$ on the basis of all data up to the time step k. In some cases, it is, however, suitable to consider a different type of the state estimate, the smoothed estimate given by $p(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k)$ and $P(\mu_{k-\ell}|\mathbf{z}_0^k)$ with $\ell > 0$. In fact, the smoothed estimate is the estimate of the past state $\bar{\mathbf{x}}_{k-\ell}$ with respect to the last currently available measurement. Thus, compared to the filtering estimate of the state $\bar{\mathbf{x}}_{k-\ell}$, the smoothed estimate utilizes information from a larger data set.

Analogously to the filtering estimate, the smoothed probability is given by the marginalization

$$P\left(\mu_{k-\ell}|\mathbf{z}_0^k\right) = \sum_{\mu_0^{k-\ell-1}} \sum_{\mu_{k-\ell+1}^k} P\left(\mu_0^k|\mathbf{z}_0^k\right). \quad (25)$$

and the conditional pdf is of the form

$$p\left(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k\right) = \sum_{\mu_0^k} P\left(\mu_0^k|\mathbf{z}_0^k\right) p\left(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k, \mu_0^k\right), \quad (26)$$

Table 1. Summary of the smoothing algorithm

Smoothing algorithm
Initialization: The predictive mean value $\hat{\mathbf{x}}_{0|-1}$, covariance matrix $\mathbf{P}_{x,0|-1}$, probability $P(\mu_0)$, and time step $k = 0$.
Step 1a: Based on the new output $\mathbf{z}_k$, compute the filtering mean values $\hat{\mathbf{x}}_{k|k}(\mu_0^k)$, covariance matrices $\mathbf{P}_{x,k|k}(\mu_0^k)$, and probability $P(\mu_0^k|\mathbf{z}_0^k)$ using (22), (23), and (15), respectively.
Step 1b (optional): Using (14), evaluate the filtering probability $P(\mu_k|\mathbf{z}_0^k)$ and employ it for finding the optimal filtering decision d_k^* according to (12) (i.e. $\ell = 0$).
Step 2a: Compute the smoothing probability $P(\mu_{k-\ell}|\mathbf{z}_0^k)$ using (25) and determine the optimal smoothing decision d_k^* according to (12) (i.e. $\ell > 0$).
Step 2b (optional): Compute the smoothing mean values $\hat{\mathbf{x}}_{k-\ell|k}(\mu_0^k)$ and covariance matrices $\mathbf{P}_{x,k-\ell|k}(\mu_0^k)$ using (27) and (28), respectively.
Step 3: Compute the predictive mean values $\hat{\mathbf{x}}_{k+1|k}(\mu_0^k)$, covariance matrices $\mathbf{P}_{x,k+1|k}(\mu_0^k)$ and probability $P(\mu_0^{k+1}|\mathbf{z}_0^k)$ using (20), (21), and (16).
Step 4: $k \leftarrow k+1$ and return to **Step 1**.

where the pdf $p(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k, \mu_0^k)$ is Gaussian with the mean value $\hat{\mathbf{x}}_{k-\ell|k}(\mu_0^k)$ and the covariance matrix $\mathbf{P}_{x,k-\ell|k}(\mu_0^k)$, i.e. $p(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k, \mu_0^k) = \mathcal{N}\{\hat{\mathbf{x}}_{k-\ell|k}(\mu_0^k), \mathbf{P}_{x,k-\ell|k}(\mu_0^k)\}$.

The solutions to the smoothing problem, i.e. computation of the pdf $p(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k)$, can generally be divided into three groups (Anderson and Moore, 1979), namely the fixed-point smoothing, when the time instant ℓ is fixed and computation can be performed online during the experiment, the fixed-lag smoothing, when the difference $k - \ell$ is fixed and computation can be carried out online, and the fixed-interval smoothing, when the time step k is fixed and computation is run backward in time. In this paper, the fixed-lag smoothing is considered, i.e. the pdf $p(\mathbf{x}_{k-\ell}|\mathbf{z}_0^k)$ with a constant lag $\ell > 0$ is computed.

In the literature, several approaches to the computation of the smoothed mean $\hat{\mathbf{x}}_{k-\ell|k}(\mu_0^k)$ and the covariance matrix $\mathbf{P}_{x,k-\ell|k}(\mu_0^k)$ have been proposed. Among others the following smoothing approaches can be mentioned; the smoothing approach based on the augmentation of the state by the state being smoothed (Söderström, 1974) and subsequent application of a filter, the smoother based on the combination of two optimal filters (the first running forward in time and the second one running backward in time) (Fraser and Potter, 1969), or the Rauch-Tung-Striebel smoother (RTSS) (Lewis, 1986). Because of computational efficiency and ease of application, the RTSS is selected here.

The relations describing the RTSS can be summarized as follows

$$\hat{\mathbf{x}}_{k-\ell|k}(\mu_0^k) = \hat{\mathbf{x}}_{k-\ell|k-\ell}(\mu_0^k) + \mathbf{K}^{\mathrm{S}}(\mu_0^k) \times \left[\hat{\mathbf{x}}_{k-\ell+1|k}(\mu_0^k) - \hat{\mathbf{x}}_{k-\ell+1|k-\ell}(\mu_0^k)\right], \quad (27)$$

$$\mathbf{P}_{x,k-\ell|k}(\mu_0^k) = \mathbf{P}_{x,k-\ell|k-\ell}(\mu_0^k) - \mathbf{K}^{\mathrm{S}}(\mu_0^k) \times \left[\mathbf{P}_{x,k-\ell+1|k-\ell}(\mu_0^k) - \mathbf{P}_{x,k-\ell+1|k-\ell}(\mu_0^k)\right] \times \mathbf{K}^{\mathrm{S}}(\mu_0^k)^T, \quad (28)$$

where the smoother gain is given by

$$\mathbf{K}^{\mathrm{S}}(\mu_0^k) = \mathbf{P}_{x,k-\ell|k-\ell}(\mu_0^k)\mathbf{A}_{\mu_l}^T\mathbf{P}_{x,k-\ell+1|k-\ell}(\mu_0^k)^{-1}. \quad (29)$$

4.3 Discussion

In this section, some issues concerning the computational demands of the exact filtering and smoothing algorithms, the choice of the cost function and the possibility of a simplified smoothing are discussed.

The relations introduced in Section 4.1 and Section 4.2 provide the exact solutions to the filtering and smoothing problems for the multiple Gaussian linear models. Unfortunately, as the number of distinct model sequences μ_0^k grows exponentially in time, there are N^{k+1} Kalman filters required at the time step k for solving the estimation problems. To keep computational demands at a reasonable level, model sequence merging, pruning, or a combination of those is needed. There are several methods that differ in complexity, computational demands and accuracy of estimates (Ackerson and Fu, 1970; Blom and Bar-Shalom, 1988; Boers and Driessen, 2005). The approximation based on model sequence merging is used because of its simplicity and clarity. The sequences that have the same last h-step history are merged together by moment matching. The depth $h \geq 0$ is chosen as a compromise between complexity and quality of estimates. It is important to note that for smoothing the depth h has to be greater or equal to the lag ℓ to make smoothing possible.

During the merging, the conditional probability of the sequence μ_{k-h}^k is computed as

$$P\left(\mu_{k-h}^k|\mathbf{z}_0^k\right) = \sum_{\mu_0^{k-h-1}} P\left(\mu_0^k|\mathbf{z}_0^k\right), \tag{30}$$

and the filtering pdf $p(\mathbf{x}_k|\mathbf{z}_0^k, \mu_{k-h}^k)$ that has the Gaussian sum form

$$p\left(\mathbf{x}_k|\mathbf{z}_0^k, \mu_{k-h}^k\right) = \sum_{\mu_0^{k-h-1}} P\left(\mu_0^{k-h-1}|\mathbf{z}_0^k, \mu_{k-h}^k\right) \times p\left(\mathbf{x}_k|\mathbf{z}_0^k, \mu_0^k\right), \tag{31}$$

where

$$P\left(\mu_0^{k-h-1}|\mathbf{z}_0^k, \mu_{k-h}^k\right) = \frac{P\left(\mu_0^k|\mathbf{z}_0^k\right)}{P\left(\mu_{k-h}^k|\mathbf{z}_0^k\right)} \tag{32}$$

is replaced by a single Gaussian pdf such that the first two moments of the random variable $\mathbf{x}_k$ are preserved.

The properties of the optimal passive detector are determined by the cost function $L_k^{\mathrm{d}}(\mu_{k-\ell}, d_k)$. When the deferring of decisions is considered (i.e. $\ell > 0$), the cost function can also include the cost connected with deferring a decision. Then the choice of the lag ℓ is given by a compromise between the quality of detection and the delay for detection. Such a compromise is inherent to all detection methods and it is demonstrated in the second scenario of the illustrative example.

Although the whole smoothing algorithm is presented in Section 4.2, only relation (25) is needed in the optimal passive detector that provides the decisions based on the smoothing estimates. The other relations of the smoothing algorithm will be used when better estimates of the whole state are required (e.g. in the case of a batch data analysis).

5. ILLUSTRATIVE EXAMPLE

This numerical example illustrates a change in the quality of detection when the smoothing estimates are employed in the passive detector. Particular attention is paid to how the value of the criterion changes as the lag increases. There are two distinct situations to consider. In the first situation the cost of deferring the decision by one time step is zero. In the second situation there is a nonzero cost connected with deferring the decision. Although this cost can be a complex function of the state and the decision, the simplest case of the constant cost is considered in this numerical example.

The finite detection horizon is $F = 40$, and the set $\mathcal{M} = \{1, 2\}$ consist of indexes two second order stable models with the following matrices

$$\begin{aligned} &\mathbf{A}_1 = \begin{bmatrix} 0.9 & 1 \\ 0 & 0.9 \end{bmatrix}, \ \mathbf{A}_2 = \begin{bmatrix} 0.8 & 1 \\ 0 & 0.9 \end{bmatrix}, \\ &\mathbf{G}_1 = 0.01\mathbf{E}_2, \quad \mathbf{G}_2 = 0.1\mathbf{E}_2, \\ &\mathbf{C}_1 = \mathbf{C}_2 = [1\ 0], \ \mathbf{H}_1 = \mathbf{H}_2 = 0.01. \end{aligned} \tag{33}$$

The initial state $\mathbf{x}_0$ has Gaussian distribution with the mean value $\hat{\mathbf{x}}_{0|-1} = [1\ 0]^T$ and the covariance matrix $\mathbf{P}_{x,0|-1} = 0.1\mathbf{E}_2$. The initial probabilities of models are $P(\mu_0 = 1) = P(\mu_0 = 2) = 0.5$ and the switching between the models is described by the transition probabilities $\pi_{1,1} = \pi_{2,2} = 0.95$ and $\pi_{1,2} = \pi_{2,1} = 0.05$. To include both the situations, the cost function $L_k^{\mathrm{d}}(\mu_{k-\ell}, d_k)$ is considered to be of the form

$$L_k^{\mathrm{d}}\left(\mu_{k-\ell}, d_k\right) = L^{\mathrm{d1}}\left(\mu_{k-\ell}, d_k\right) + L^{\mathrm{d2}}\ell, \tag{34}$$

where L^{d2} is a constant cost of deferring the decision by one time step and $L^{\mathrm{d1}}(\mu_{k-\ell}, d_k)$ is a cost function penalizing wrong decisions, that is chosen to be

$$L^{\mathrm{d1}}\left(\mu_{k-\ell}, d_k\right) = \begin{cases} 0 \text{ if } \mu_{k-\ell} = d_k, \\ 1 \text{ if } \mu_{k-\ell} \neq d_k. \end{cases} \tag{35}$$

It can easily be shown that regardless of the cost L^{d2} and the lag ℓ, the chosen form of the cost functions $L_k^{\mathrm{d}}(\mu_{k-\ell}, d_k)$ and $L^{\mathrm{d1}}(\mu_{k-\ell}, d_k)$ leads to the passive change detector of which the decisions are given as

$$d_k^* = \arg\max_{d_k} P\left(\mu_{k-\ell} = d_k|\mathbf{z}_0^k\right). \tag{36}$$

In the first scenario, a comparison between the smoothing and filtering decisions for particular parameters is provided. The cost of deferring the decision by one step is chosen to be $L^{\mathrm{d2}} = 0.01$. The depth of merging and the lag are chosen the same $h = \ell = 3$. A typical simulation run demonstrating the performance of the smoothing and filtering decisions is depicted in Fig. 2. The upper graph shows the model sequence, and the filtering and smoothing decisions. It can be seen that the passive detector based on the smoothing estimates generally detects the changes more reliably and thus produces fewer wrong decisions. From the bottom graph, which shows the filtered and smoothed probabilities, it can be seen that the smoothing decisions are also more likely to be correct because the probabilities converge closer to the limit values zero and one.

In the second scenario, the dependence of the criterion value J on the lag ℓ for various costs L^{d2} is examined. Since the improvement is significant only for short lags, the maximum lag is chosen to be $\ell_{\max} = 5$. To ensure comparability of the results for all considered lags, the criterion is evaluated only on a shorten horizon 0 to $F - \ell_{\max}$ and the depth for merging is set to maximum lag, i.e.

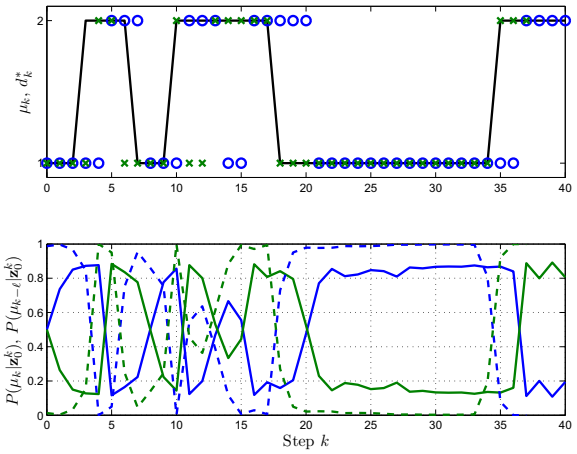


Fig. 2. A typical simulation run. Upper graph: Model sequence – black solid line, Filtering decisions ($\ell = 0$) – blue o-markers, Smoothing decisions ($\ell = 3$) – green x-markers. Bottom graph: Filtering probabilities – solid lines, Smoothing probabilities – dashed lines; Model 1 – blue lines, Model 2 – green lines.

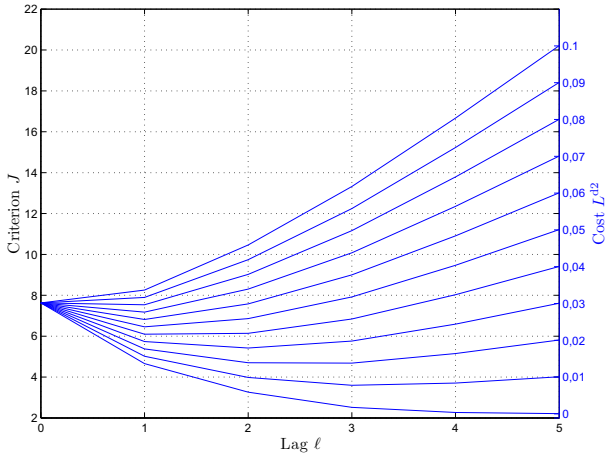


Fig. 3. The relation between criterion J and lag ℓ for various costs L^{d2}.

$h = \ell_{max}$. The results of 1000 Monte Carlo simulations are presented in Fig 3. The bottom line shows the dependence of the criterion value on the lag when the cost of deferring a decision is zero. Since the criterion value monotonically decreases to a limit value for the increasing lag, the optimal value of the lag can be determined based on computational complexity which increases with the increasing lag or more precisely with the increasing merging depth h. For a nonzero cost L^{d2} there is an optimal value of the lag as it can directly be seen from the figure. It is obvious that when the cost L^{d2} is too high, it does not pay off to defer the decision at all.

6. CONCLUSION

The paper was concentrated on employing the smoothing estimates in change detection and demonstrating the possible improvement of the detection quality. As the smoothing estimates can be employed only when the decisions are deferred, a compromise between the detection quality and the speed of detection was searched for. Although the cost function considered in the paper was quite simple because of demonstration purposes, a more complex cost function based on the real costs evaluated for a particular application could be readily used. The future work would be aimed at the evaluation of approximate smoothing algorithms in the context of change detection.

REFERENCES

Ackerson, G.A. and Fu, K.S. (1970). On state estimation in switching environments. *IEEE Transactions on Automatic Control*, 15(1), 10–17.

Anderson, B.D.O. and Moore, J.B. (1979). *Optimal Filtering*. Prentice Hall, Englewood Cliffs, NJ, USA.

Athans, M., Fekri, S., and Pascoal, A. (2006). Issues on robust adaptive feedback control. In *Proceedings of the* 16^{th} *IFAC World Congress*. Oxford, UK.

Bar-Shalom, Y., Li, X.R., and Kirubarajan, T. (2001). *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, New York, NY, USA.

Bar-Shalom, Y. and Tse, E. (1974). Dual effects, certainty equivalence and separation in stochastic control. *IEEE Transactions on Automatic Control*, 19, 494–500.

Basseville, M. and Nikiforov, I.V. (1993). *Detection of abrupt changes – Theory and application*. Prentice Hall, Englewood Cliffs, NJ, USA.

Blackmore, L. and Williams, B.C. (2006). Finite horizon control design for optimal dicrimination between several models. In *Proceedings of the* 45^{th} *IEEE Conference on Decision and Control*, 1147–1152. San Diego, USA.

Blom, H.A.P. and Bar-Shalom, Y. (1988). The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8), 780–783.

Boers, Y. and Driessen, H. (2005). A multiple model multiple hypothesis filter for Markovian switching systems. *Automatica*, 41(4), 709–716.

Fraser, D. and Potter, J. (1969). The optimum linear smoother as a combination of two optimum linear filters. *IEEE Transactions On Automatic Control*, 14(4), 387–390.

Gustafsson, F. (2000). *Adaptive Filtering and Change Detection*. John Wiley & Sons, Chichester, WSX, UK.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods – a survey. *Automatica*, 20(4), 387–404.

Jones, H.L. (1973). *Failure detection in linear systems*. Ph.D. thesis, Department of Aeronautics and Astronautics, M.I.T., Cambridge, Massachusetts.

Lewis, F.L. (1986). *Optimal Estimation*. John Wiley & Sons, New York.

Söderström, T. (1974). Discrete-time stochastic systems. *Inf. Sci.*, 7, 253–270.

Šimandl, M. and Královec, J. (2000). Filtering, prediction and smoothing with gaussian sum representation. In *Proceedings of the* 12^{th} *IFAC Symposium on System Identification*. Santa Barbara, USA.

Šimandl, M. and Punčochář, I. (2009). Active fault detection and control: Unified formulation and optimal design. *Automatica*, 45(9), 2052–2059.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Predictive fault-tolerant control of Takagi-Sugeno fuzzy systems ⋆

Łukasz Dziekan * **Marcin Witczak** *

* *Institute of Control and Computation Engineering, University of Zielona Góra, ul. Podgórna 50, 65–246 Zielona Góra, Poland, e-mail:* {L.Dziekan, M.Witczak} @issi.uz.zgora.pl.

AbstractIn this paper, an active FTC strategy is presented. After short introduction to Takagi-Sugeno fuzzy systems, it is developed for such systems with low conservatism in Lyapunov stability derivation. Fault identification is based on the use of an observer and integrated with the FTC controller, implemented as a model predictive controller.

Keywords: fuzzy, fault tolerant control, model predictive control, fault diagnosis

1. INTRODUCTION

Fault Tolerant Control (FTC, see Blanke et al. (2003)) system allows to control plant in such a way that it fulfils desired objectives (perhaps with a possible performance degradation) in the presence of non-critical faults in components of the system (actuators, measurement devices and/or plant). In general, FTC systems are classified into two distinct classes Zhang and Jiang (2003): passive and active. In passive FTC controllers are designed to be robust against a set of presumed faults .And active FTC systems, in contrast to passive ones, react to system components faults actively by reconfiguring control actions, and by doing so the system stability and acceptable performances are maintained. To achieve that, the control system relies on the detection and isolation (FDI) Korbicz et al. (2004); Witczak (2006, 2007) and accommodation technique.There is also a need for designing an integrated fault identification and fault-tolerant control strategy for both linear and non-linear systems, because of the fact that perfect fault identification is impossible to attain in practical applications. Dealing with faults means also dealing with constraints on control input, but standard state feedback is often incapable of dealing in an efficient way with such situations, so the use of model predictive control (MPC) is also considered here Maciejowski (2002). Additionally in normally defined MPC there is a problem in guaranteeing stability, the survey paper on the subject can be found here Mayne et al. (2000).
However due to computational complexity of the general MPC, it could be impossible to use for fast enough systems, therefore there is a need to optimize the method. One way to achieve it, is to use an adapted version of fast MPC Wang and Boyd (2010) for Takagi-Sugeno (T-S) fuzzy systems. Problem stated above with state dimension n, input dimension m and control horizon T_c takes $O(T_c^3(n+m)^3$ operations per step in an interior-point method, but if special structure of the problem is exploited then computational complexity is $O(T_c(n+m)^3$. Since interior-point methods require only a constant (and modest) number of steps, it follows that complexity of MPC is therefore linear, instead than cubic in control horizon. Also if weight matrices $\boldsymbol{Q}_R$ and $\boldsymbol{R}_R$ are diagonal, problem is state control separable and there are box constraints (as considered in this paper), then the overall complexity is of order $O(T_c(n^3+n^2m)$, which grows linearly in both T_c and m. The further optimizations stems from using warm start method, in which calculations are initialized using predictions made in the previous step, which with an appropriate choice of interior-point method can cut number of steps required by a factor of five or more. The last optimization is early termination of an appropriate interior-point method, which even after a few iterations can provide quite good control action. Although it is recommended that, if possible, for the very few first control steps the number of iterations were higher than for the rest of the optimization.
This paper is organised as follows. Sec. 2 presents background information about T-S fuzzy systems. In Sec. 3 an improved design technique for an integrated FTC and fault identification strategy for T-S fuzzy systems is proposed that allows to include input constraints into FTC system. Subsequently the input constraints for the FTC are considered, which are followed by the regulator problem for T-S fuzzy system and description of the MPC adapted for the proposed approach. The final part of the paper presents a numerical example which shows the performance of the proposed approach.

2. ELEMENTARY BACKGROUND ON T-S FUZZY SYSTEMS

A non-linear dynamic system can be described in a simple way by a Takagi-Sugeno fuzzy model, which uses series of locally linearised non-linear models (see, e.g. Takagi and Sugeno (1985); Korbicz et al. (2004)). A T-S model is described by fuzzy IF-THEN rules which represent local linear I/O relations of the non-linear system. It has a rule base of M rules, each having p antecedents, where ith rule is expressed as

⋆ The work was financed as a research project with the science funds for years 2007-2010.

$$R^i: \text{ IF } w_k^1 \text{ is } F_1^i \text{ and } \ldots \text{ and } w_k^p \text{ is } F_p^i, \quad \text{THEN } \begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{A}^i \boldsymbol{x}_k + \boldsymbol{B}^i \boldsymbol{u}_k, \\ \boldsymbol{y}_k = \boldsymbol{C}^i \boldsymbol{x}_k, \end{cases} \tag{1}$$

in which $\boldsymbol{x}_k \in \mathbb{R}^n$ stands for the reference state, $\boldsymbol{y}_k \in \mathbb{R}^m$ is the reference output, and $\boldsymbol{u}_k \in \mathbb{R}^r$ denotes the nominal control input, also $i = 1, \ldots, M$, F_j^i $(j = 1, \ldots, p)$ are fuzzy sets and $\boldsymbol{w}_k = [w_k^1, w_k^2, \ldots, w_k^p]$ is a known vector of premise variables Takagi and Sugeno (1985).
Given a pair of $(\boldsymbol{w}_k, \boldsymbol{u}_k)$ and a product inference engine, normalized rule firing strengths $h_i(\boldsymbol{w}_k)$ are defined as

$$h_i(\boldsymbol{w}_k) = \frac{\mathcal{T}_{j=1}^p \mu_{F_j^i}(w_k^j)}{\sum_{i=1}^M (\mathcal{T}_{j=1}^p \mu_{F_j^i}(w_k^j))} \tag{2}$$

and $\mathcal{T}$ denotes a t-norm (e.g., product). The term $\mu_{F_j^i}(w_k^j)$ is the grade of membership of the premise variable w_k^j. Moreover, the rule firing strengths $h_i(\boldsymbol{w}_k)$ $(i = 1, \ldots, M)$ satisfy the following constraints

$$\sum_{i=1}^M h_i(\boldsymbol{w}_k) = 1, \ 0 \leqslant h_i(\boldsymbol{w}_k) \leqslant 1, \quad \forall i = 1, \ldots, M. \tag{3}$$

3. FTC STRATEGY FOR T-S FUZZY SYSTEMS

Let us consider the following T-S reference model:

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{B}_k \boldsymbol{u}_k, \tag{4}$$
$$\boldsymbol{y}_{k+1} = \boldsymbol{C}_{k+1} \boldsymbol{x}_{k+1}, \tag{5}$$

with $\boldsymbol{A}_k = \sum_{i=1}^M h_i(\boldsymbol{w}_k)\boldsymbol{A}^i$, $\boldsymbol{B}_k = \sum_{i=1}^M h_i(\boldsymbol{w}_k)\boldsymbol{B}^i$, $\boldsymbol{C}_{k+1} = \sum_{i=1}^M h_i(\boldsymbol{w}_{k+1})\boldsymbol{C}^i$ for $i = 1, \ldots, M$.

Let us also consider a possibly faulty T-S system described by the following equations:

$$\boldsymbol{x}_{f,k+1} = \boldsymbol{A}_k \boldsymbol{x}_{f,k} + \boldsymbol{B}_k \boldsymbol{u}_{f,k} + \boldsymbol{L}_k \boldsymbol{f}_k \tag{6}$$
$$\boldsymbol{y}_{f,k+1} = \boldsymbol{C}_{k+1} \boldsymbol{x}_{f,k+1}, \tag{7}$$

with $\boldsymbol{L}_k = \sum_{i=1}^M h_i(\boldsymbol{w}_k)\boldsymbol{L}^i$. Where $\boldsymbol{x}_{f,k} \in \mathbb{R}^n$ stands for the system state, $\boldsymbol{y}_{f,k} \in \mathbb{R}^m$ is the system output, $\boldsymbol{u}_{f,k} \in \mathbb{R}^r$ denotes the system input, $\boldsymbol{f}_k \in \mathbb{R}^s$, $(s \leq m)$ is the fault vector, and $\boldsymbol{L}^i$ stands for its distribution matrices which are assumed to be known.
The main objective of this paper is to propose a control strategy which can be used for determining the system input $\boldsymbol{u}_{f,k}$ such that:

- the control loop for the system (6)–(7) is stable,
- $\boldsymbol{x}_{f,k+1}$ converges asymptotically to $\boldsymbol{x}_{k+1}$ irrespective of the presence of the fault $\boldsymbol{f}_k$.

The crucial idea is to use the following control strategy:

$$\boldsymbol{u}_{f,k} = -\boldsymbol{S}_k \hat{\boldsymbol{f}}_k + \boldsymbol{K}_{1,k}(\boldsymbol{x}_k - \boldsymbol{x}_{f,k}) + \boldsymbol{u}_k \tag{8}$$

where $\hat{\boldsymbol{f}}_k$ is the fault estimate. Note that, it is not assumed that $\boldsymbol{x}_{f,k}$ is available, i.e. an estimate $\hat{\boldsymbol{x}}_{f,k}$ can be used instead. Thus, the following problems arise:

- to determine $\hat{\boldsymbol{f}}_k$,
- to design $\boldsymbol{K}_{1,k}$ in such a way that the control loop is stable, i.e. the stabilisation problem. The control law in such a form is called the PDC (parallel distributed compensation, Wang et al. (1996)).

Due to space constraints in the next subsections, only basis needed to implement FTC technique will be presented, for greater details refer here Dziekan et al. (2009).

3.1 Fault identification

Let us assume that the following rank condition is satisfied at any given moment [1] $\operatorname{rank}(\boldsymbol{C}_{k+1}\boldsymbol{L}_k) = \operatorname{rank}(\boldsymbol{L}_k) = s$. This implies that it is possible to calculate pseudo-inverse $\boldsymbol{H}_{k+1} = (\boldsymbol{C}_{k+1}\boldsymbol{L}_k)^+$. Thus, the fault estimate is given as

$$\hat{\boldsymbol{f}}_k = \boldsymbol{H}_{k+1}(\boldsymbol{y}_{f,k+1} - \boldsymbol{C}_{k+1}\boldsymbol{A}_k\hat{\boldsymbol{x}}_{f,k} - \boldsymbol{C}_{k+1}\boldsymbol{B}_k\boldsymbol{u}_{f,k}). \tag{9}$$

Unfortunately, the crucial problem with practical implementation of (9) is that it requires $\boldsymbol{y}_{f,k+1}$ and $\boldsymbol{u}_{f,k}$ to calculate $\hat{\boldsymbol{f}}_k$ and hence it cannot be directly used to obtain (8). Therefore it is necessary to use a prediction of a current fault estimate $\mathring{\boldsymbol{f}}_k$ based in some way on the previous fault estimates. To settle this problem, it is assumed that there exists a diagonal matrix $\boldsymbol{\alpha}_k$ such that $\hat{\boldsymbol{f}}_k \cong \mathring{\boldsymbol{f}}_k = \boldsymbol{\alpha}_k \hat{\boldsymbol{f}}_{k-1}$ and hence the practical form of (8) boils down to

$$\boldsymbol{u}_{f,k} = -\boldsymbol{S}_k \mathring{\boldsymbol{f}}_k + \boldsymbol{K}_{1,k}(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{f,k}) + \boldsymbol{u}_k. \tag{10}$$

In most cases matrix $\boldsymbol{\alpha}_k$ should be equivalent to an identity matrix, i.e. it would simply mean an one time–step delay, which should have negligible effect on the outcome. In cases where the fault behaviour is a linear one, it is possible to design the matrix $\boldsymbol{\alpha}_k$ based on the previous changes of faults. In cases where faults changes in a nonlinear fashion and one time–step delay is unacceptable, one could try to predict the nature of the faults by using for example neural networks.

3.2 Observer design

As was already mentioned, the fault estimate (9) is obtained based on the state estimate $\hat{\boldsymbol{x}}_{f,k}$. This raises the necessity for an observer design. Let us assume that $\boldsymbol{S}_k$ at any moment satisfies the following equality $\boldsymbol{B}_k\boldsymbol{S}_k = \boldsymbol{L}_k$, e.g. for actuator faults $\boldsymbol{S}_i = -\boldsymbol{I}$ for all $i = 1, \ldots, M$. And

$$\bar{\boldsymbol{A}}_k = (\boldsymbol{I} - \boldsymbol{L}_k\boldsymbol{H}_{k+1}\boldsymbol{C}_{k+1})\boldsymbol{A}_k,$$
$$\bar{\boldsymbol{B}}_k = (\boldsymbol{I} - \boldsymbol{L}_k\boldsymbol{H}_{k+1}\boldsymbol{C}_{k+1})\boldsymbol{B}_k, \quad \bar{\boldsymbol{L}}_k = \boldsymbol{L}_k\boldsymbol{H}_{k+1}.$$

Then the observer structure, which can be perceived as an unknown input observer (see, e.g. Hui and Zak (2005); Witczak (2007)), is proposed to use a modified version of the celebrated Kalman filter, which can be described as follows:

$$\hat{\boldsymbol{x}}_{f,k+1/k} = \bar{\boldsymbol{A}}_k\hat{\boldsymbol{x}}_{f,k} + \bar{\boldsymbol{B}}_k\boldsymbol{u}_{f,k} + \bar{\boldsymbol{L}}_k\boldsymbol{y}_{f,k+1},$$
$$\boldsymbol{P}_{k+1/k} = \bar{\boldsymbol{A}}_k\boldsymbol{P}_k{\bar{\boldsymbol{A}}_k}^T + \boldsymbol{U}_k,$$
$$\boldsymbol{K}_{2,k+1} = \boldsymbol{P}_{k+1/k}\boldsymbol{C}_{k+1}^T\left(\boldsymbol{C}_{k+1}\boldsymbol{P}_{k+1/k}\boldsymbol{C}_{k+1}^T + \boldsymbol{V}_{k+1}\right)^{-1},$$
$$\hat{\boldsymbol{x}}_{f,k+1} = \hat{\boldsymbol{x}}_{f,k+1/k} + \boldsymbol{K}_{2,k+1}(\boldsymbol{y}_{f,k+1} - \boldsymbol{C}_{k+1}\hat{\boldsymbol{x}}_{f,k+1/k}),$$
$$\boldsymbol{P}_{k+1} = [\boldsymbol{I} - \boldsymbol{K}_{2,k+1}\boldsymbol{C}_{k+1}]\,\boldsymbol{P}_{k+1/k},$$

where $\boldsymbol{U}_k = \delta_1\boldsymbol{I}$ and $\boldsymbol{V}_k = \delta_2\boldsymbol{I}$ with δ_1 and δ_2 sufficiently small positive numbers.

It is important to note that the Kalman filter is applied here for state estimation of a deterministic system (6)–(7) and hence $\boldsymbol{U}_k$ and $\boldsymbol{V}_k$ play the role of instrumental matrices only (see Witczak (2007) and the references therein for more details).

[1] It is not easy to guarantee that, unless matrices $\boldsymbol{C}_k$ and $\boldsymbol{L}_k$ are time invariant, i.e. $\boldsymbol{C}^1 = \boldsymbol{C}^i$ and $\boldsymbol{L}^1 = \boldsymbol{L}^i$ for all $i = 1, \ldots, M$. However in real life cases, checking if rank condition is satisfied for every pair of matrices, i.e. $\operatorname{rank}(\boldsymbol{C}^i\boldsymbol{L}^i) = \operatorname{rank}(\boldsymbol{L}^i) = s$ for all $i = 1, \ldots, M$ is usually sufficient.

3.3 Integrated design procedure

First, let us start with two crucial assumptions:

- the pair $(\bar{\boldsymbol{A}}_k, \boldsymbol{C}_{k+1})$ is detectable,
- the pair $(\boldsymbol{A}_k, \boldsymbol{B}_k)$ is stabilisable.

Under these assumptions, it is possible to design the matrices $\boldsymbol{K}_{1,k}$ and $\boldsymbol{K}_{2,k}$ in such a way that the extended error

$$\bar{\boldsymbol{e}}_{k+1} = \begin{bmatrix} \boldsymbol{A}_k - \boldsymbol{B}_k\boldsymbol{K}_{1,k} & \boldsymbol{L}_k\boldsymbol{H}_{k+1}\boldsymbol{C}_{k+1}\boldsymbol{A}_k \\ \boldsymbol{0} & \bar{\boldsymbol{A}}_k - \boldsymbol{K}_{2,k+1}\boldsymbol{C}_{k+1}\bar{\boldsymbol{A}}_k \end{bmatrix} \bar{\boldsymbol{e}}_k. \tag{11}$$

converges asymptotically to zero.
It can be observed from the structure of (11) that the eigenvalues of the matrix are the union of those of $\boldsymbol{A}_k - \boldsymbol{B}_k\boldsymbol{K}_{1,k}$ and $\bar{\boldsymbol{A}}_k - \boldsymbol{K}_{2,k+1}\boldsymbol{C}_{k+1}\bar{\boldsymbol{A}}_k$. This clearly indicates that the design of the state feedback and the observer can be carried out independently. So, let us start with the controller design with the corresponding tracking error defined by

$$\boldsymbol{e}_{k+1} = [\boldsymbol{A}_k - \boldsymbol{B}_k\boldsymbol{K}_{1,k}]\boldsymbol{e}_k = \boldsymbol{A}_0(\boldsymbol{h}(\boldsymbol{w}_k))\boldsymbol{e}_k, \tag{12}$$

where $\boldsymbol{K}_{1,k} = \sum_{i=1}^{M} h_i(\boldsymbol{w}_k)\boldsymbol{K}_1^i$ and the matrix $\boldsymbol{A}_0(\boldsymbol{h}(\boldsymbol{w}_k))$ belongs to a convex polytopic set defined as

$$\begin{aligned} \mathbb{A}_0 = \Bigg\{ \boldsymbol{A}_0(\boldsymbol{h}(\boldsymbol{w}_k)) : \sum_{i=1}^{M} h_i(\boldsymbol{w}_k) = 1,\ 0 \leqslant h_i(\boldsymbol{w}_k) \leqslant 1 \\ \boldsymbol{A}_0(\boldsymbol{h}(\boldsymbol{w}_k)) = \sum_{i=1}^{M}\sum_{j=1}^{M} h_i(\boldsymbol{w}_k)h_j(\boldsymbol{w}_k)\boldsymbol{A}_{0,i,j}, \\ \boldsymbol{A}_{0,i,j} = \frac{1}{2}(\boldsymbol{A}^i - \boldsymbol{B}^i\boldsymbol{K}_1^j + \boldsymbol{A}^j - \boldsymbol{B}^j\boldsymbol{K}_1^i) \Bigg\} \end{aligned} \tag{13}$$

By adapting the general results of the work of Rong and Irwin (2003), the following definition is introduced:

Definition 1. The tracking error described by (12) is robustly convergent to zero in the uncertainty domain (13) iff all eigenvalues of $\boldsymbol{A}_0(\boldsymbol{h}(\boldsymbol{w}_k))$ have magnitude less than one for all values of $\boldsymbol{h}(\boldsymbol{w}_k)$ such that $\boldsymbol{A}_0(\boldsymbol{h}(\boldsymbol{w}_k)) \in \mathbb{A}_0$.

Theorem 2. The tracking error described by (12) is robustly convergent to zero in the uncertainty domain (13) if there exist matrices $\boldsymbol{Q}_{i,j} \succ \boldsymbol{0}$, $\boldsymbol{G}_1$, $\boldsymbol{W}_j$ such that

$$\begin{bmatrix} \boldsymbol{G}_1 + \boldsymbol{G}_1^T - \boldsymbol{Q}_{i,j} & * \\ \boldsymbol{N}_{0,i,j} & \boldsymbol{Q}_{m,n} \end{bmatrix} \succ \boldsymbol{0}, \tag{14}$$

for all $i, m = 1, \ldots, M$ and $j \geqslant i$, $n \geqslant m$, where $\boldsymbol{N}_{0,i,j} = \frac{1}{2}[(\boldsymbol{A}^i + \boldsymbol{A}^j)\boldsymbol{G}_1 - \boldsymbol{B}^i\boldsymbol{W}_j - \boldsymbol{B}^j\boldsymbol{W}_i]$.

Proof. see Rong and Irwin (2003).

Finally, the design procedure boils down to solving the set of $[\frac{1}{2}M(1+M)]^2$ LMIs (14) and then determining $\boldsymbol{K}_1^i = \boldsymbol{W}_i\boldsymbol{G}_1^{-1}$.

3.4 Constraints on the control input

When the initial tracking error is known (i.e., the deviation of a faulty system state from a nominal system state), an upper bound on the norm of the control input $\mathring{\boldsymbol{u}}_{f,k} = \boldsymbol{K}_1(\boldsymbol{x}_k - \boldsymbol{x}_{f,k})$ can be found as follows Boyd et al. (1994). Let us assume that initial tracking error $\boldsymbol{e}_0$ lies in an ellipsoid of diameter γ, i.e., $\|\boldsymbol{e}_0\| \leqslant \gamma$, then the constraint on a control input described as follows $\|\mathring{\boldsymbol{u}}_{f,k}\|_{\max} \triangleq \max_l|\mathring{\boldsymbol{u}}_{f,k}^l| \leqslant \lambda$ is enforced at all times if the LMIs

$$\begin{aligned} \begin{bmatrix} \boldsymbol{X} & * \\ 0.5(\boldsymbol{W}_i^T + \boldsymbol{W}_j^T) & \boldsymbol{G}_1 + \boldsymbol{G}_1^T - \boldsymbol{Q}_{i,j} \end{bmatrix} \succeq \boldsymbol{0}, \\ \boldsymbol{Q}_{i,j} \succeq \gamma^2\boldsymbol{I}, \quad \mathrm{diag}(\boldsymbol{X}) \preceq \lambda^2\boldsymbol{I}, \end{aligned} \tag{15}$$

hold, where $\boldsymbol{W}_i$, $\boldsymbol{W}_j$ satisfy conditions given by (14) for for all $i = 1, \ldots, M$ and $j \geqslant i$.

3.5 Regulator problem

In order to solve the regulator problem it is needed to find a PDC controller such that the following objective function is minimized,

$$J_\infty = \sum_0^\infty \left(\boldsymbol{y}_k^T\boldsymbol{Q}_R\boldsymbol{y}_k + \mathring{\boldsymbol{u}}_{f,k}^T\boldsymbol{R}_R\mathring{\boldsymbol{u}}_{f,k} \right) \tag{16}$$

where $\boldsymbol{Q}_R \succeq 0$ and $\boldsymbol{R}_R \succ 0$ are suitable weight matrices. However system described in this paper is uncertain and thus only the upper bound of the objective function can be minimized. Therefore the following Theorem 3 only gives a sub-optimal solution for the the regulator problem Rong and Irwin (2003).

Theorem 3. The upper bound for the objective function (16) for initial tracking error $\boldsymbol{e}_0$ lying in an ellipsoid of diameter γ can be obtained by solving the following LMI optimization problem of η scalar

$$\min_{\boldsymbol{Q}_{i,j}, \boldsymbol{G}_1, \boldsymbol{W}_i} \eta$$

subject to

$$\begin{aligned} \begin{bmatrix} \boldsymbol{G}_1 + \boldsymbol{G}_1^T - \boldsymbol{Q}_{i,j} & * & * & * \\ \boldsymbol{N}_{0,i,j} & \boldsymbol{Q}_{m,n} & * & * \\ 0.5\boldsymbol{Q}_R^{1/2}(\boldsymbol{C}^i + \boldsymbol{C}^j)\boldsymbol{G}_1 & 0 & \eta\boldsymbol{I} & * \\ 0.5\boldsymbol{R}_R^{1/2}(\boldsymbol{W}_i + \boldsymbol{W}_j) & 0 & 0 & \eta\boldsymbol{I} \end{bmatrix} \succ \boldsymbol{0}, \\ \boldsymbol{Q}_{i,j} \succeq \gamma^2\boldsymbol{I} \end{aligned} \tag{17}$$

for all $i, m = 1, \ldots, M$ and $j \geqslant i$, $n \geqslant m$, where $\boldsymbol{N}_{0,i,j} = \frac{1}{2}[(\boldsymbol{A}^i + \boldsymbol{A}^j)\boldsymbol{G}_1 - \boldsymbol{B}^i\boldsymbol{W}_j - \boldsymbol{B}^j\boldsymbol{W}_i]$ and local feedbacks gains are $\boldsymbol{K}_1^i = \boldsymbol{W}_i\boldsymbol{G}_1^{-1}$.
Proof. see Rong and Irwin (2003).

3.6 Model predictive control

Even though PDC regulator can be designed with constraints on the control input, the performance of such a control is suboptimal in an control areas close to saturation, this is due to linear behaviour of PDC at any given time point. One way to alleviate this problem is to use MPC with can deal with constraints in a very efficient way. Approach presented below achieve stability by implementing the terminal cost function $\boldsymbol{Q}_f$. One can find a terminal cost function by solving regulator problem (17) with constraints (14), and afterwards finding minimal $\boldsymbol{Q}_{i,j}$ which allows to compute final cost as $\boldsymbol{Q}_f = \eta(\boldsymbol{Q}_{i,j,min})^{-1}$. Taking this into account we can state the model predictive control as a Quadratic programming problem (QP), which is to solve at each iteration k, in a following way:
minimize

$$\mathring{\boldsymbol{x}}_{f,k+T_c}^T\boldsymbol{Q}_f\mathring{\boldsymbol{x}}_{f,k+T_c} + \sum_{\tau=k}^{k+T_c-1} \mathring{\boldsymbol{x}}_{f,\tau}^T\boldsymbol{Q}_r\mathring{\boldsymbol{x}}_{f,\tau} + \mathring{\boldsymbol{u}}_{f,\tau}^T\boldsymbol{R}_R\mathring{\boldsymbol{u}}_{f,\tau}$$

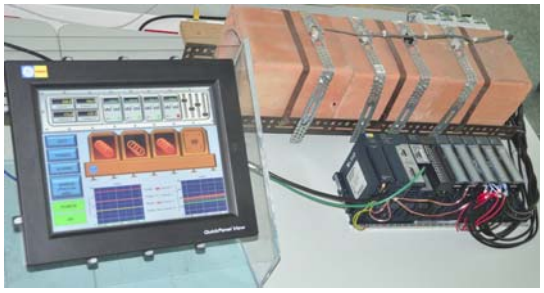


Figure 1. Laboratory model of a tunnel furnace – hardware setup

subject to

$$\mathring{\boldsymbol{x}}_{min} \leqslant \mathring{\boldsymbol{x}}_{\tau} \leqslant \mathring{\boldsymbol{x}}_{max}, \quad \tau = k+1, \ldots, k+T_c$$
$$\boldsymbol{u}_{min} + S\mathring{\boldsymbol{f}}_k - \boldsymbol{u}_{\tau} \leqslant \mathring{\boldsymbol{u}}_{\tau} \leqslant \boldsymbol{u}_{max} + S\mathring{\boldsymbol{f}}_k - \boldsymbol{u}_{\tau}, \quad \tau = k+1, \ldots, k+T_c$$
$$\mathring{\boldsymbol{x}}_{f,\tau+1} = \boldsymbol{A}_{\tau}\mathring{\boldsymbol{x}}_{f,\tau+1} + \boldsymbol{B}_{\tau}\mathring{\boldsymbol{u}}_{\tau}, \ \tau = k+1, \ldots, k+T_c \quad (18)$$

with variables $\mathring{\boldsymbol{x}}_{f,k+1}, \ldots, \mathring{\boldsymbol{x}}_{f,k+T}$ and $\mathring{\boldsymbol{u}}_{f,k}, \ldots, \mathring{\boldsymbol{u}}_{f,k+T_c-1}$. Here, T_c is a control (planning) horizon, $\boldsymbol{Q}_R \succeq 0$ and $\boldsymbol{R}_R \succ 0$ are suitable weight matrices with terminal cost function $\boldsymbol{Q}_f$. The problem is a convex QP, with problem data: starting point of optimization is $\mathring{\boldsymbol{x}}_{f,k} = \hat{\boldsymbol{x}}_{f,k} - \boldsymbol{x}_{f,k}$, and matrices $\boldsymbol{A}_{\tau}$ and $\boldsymbol{B}_{\tau}$ are defuzzified along the previously computed trajectory path (but not including $-S * \mathring{\boldsymbol{f}}_k$ factor). Let $\mathring{\boldsymbol{x}}^*_{f,k+1}, \ldots, \mathring{\boldsymbol{x}}^*_{f,k+T}, \mathring{\boldsymbol{u}}^*_{f,k}, \ldots, \mathring{\boldsymbol{u}}^*_{f,k+T_c-1}$ be optimal for (18). The MPC policy takes the first control action in this plan, as our control $\boldsymbol{u}_{f,k} = \mathring{\boldsymbol{u}}^*_{f,k} - \boldsymbol{S}_k\mathring{\boldsymbol{f}}_k + \boldsymbol{u}_k$. For the next control step, for defuzzification of matrices $\boldsymbol{A}_{\tau}$ and $\boldsymbol{B}_{\tau}$ it is assumed that last control takes a form of PDC control $\boldsymbol{u}_{f,f,k+T_c} = \boldsymbol{K}_{1,k+T_c}\mathring{\boldsymbol{x}}_{f,k+T_c} - \boldsymbol{S}_k\mathring{\boldsymbol{f}}_k + \boldsymbol{u}_{k+T_c}$, found by solving regulator problem (17) with constraints (14).

4. ILLUSTRATIVE EXAMPLE

The selected non-linear system results from a laboratory model of tunnel furnace. The considered tunnel furnace is designed for the simulation in the laboratory conditions of the real industrial tunnel furnaces, which can be applied in the food industry or production of ceramic among others. The furnace is equipped in three electric heaters and four temperature sensors. The required temperature of the furnace can be kept by the controlling of the heaters work. This task can be achieved by the group regulation of the voltage with the application of controller PACSystems RX3i manufactured by GE Fanuc Intelligent Platforms and semiconductor relays RP6 produced by LUMEL, providing impulse control with a variable impulse frequency, $f_{max} = 1$Hz. The temperature of the furnace is measured via IC695ALG600 module from Pt100 resistive thermal devices (RTDs). The visualisation of work of the tunnel furnace is made by Quickpanel CE device from GE Fanuc Intelligent Platforms. Its hardware setup can be seen on Fig. 1.

The tunnel furnace can be considered as a three-input and four-output system. Based on an experimental data, with a sampling time of 1s, a normalized T-S model, which approximates the non-linear behaviour of the tunnel furnace, is obtained by linearising system around five operating points. The matrices $\boldsymbol{A}^i$, $\boldsymbol{B}^i$ are acquired with the premise variable $\boldsymbol{w}_k = \boldsymbol{y}_{k,2}$ and triangular membership functions. For the subsequent simulation it was assumed that the third sensor is unavailable. The following numerical values were used:

$$\boldsymbol{A}^1 = \begin{bmatrix} 1.0021 & -0.0040 & -0.0230 & 0.0259 \\ 0.0023 & 0.9960 & -0.0083 & 0.0099 \\ 0.0024 & -0.0028 & 0.9907 & 0.0099 \\ 0.0009 & -0.0005 & -0.0059 & 1.0051 \end{bmatrix},$$

$$\boldsymbol{A}^2 = \begin{bmatrix} 0.9995 & -0.0048 & 0.0010 & 0.0038 \\ 0.0003 & 0.9956 & 0.0008 & 0.0028 \\ 0.0001 & -0.0014 & 0.9994 & 0.0011 \\ 0.0002 & -0.0023 & 0.0005 & 1.0011 \end{bmatrix},$$

$$\boldsymbol{A}^3 = \begin{bmatrix} 1.0013 & -0.0034 & 0.0024 & -0.0009 \\ 0.0021 & 0.9970 & 0.0002 & 0.0004 \\ 0.0011 & -0.0002 & 0.9976 & 0.0013 \\ 0.0006 & 0.0001 & -0.0006 & 0.9993 \end{bmatrix},$$

$$\boldsymbol{A}^4 = \begin{bmatrix} 0.9993 & -0.0022 & 0 & 0.0029 \\ 0.0002 & 0.9967 & -0.0005 & 0.0034 \\ -0.0003 & 0.0001 & 0.9987 & 0.0006 \\ 0.0004 & -0.0018 & 0.0016 & 0.9989 \end{bmatrix},$$

$$\boldsymbol{A}^5 = \begin{bmatrix} 0.9977 & -0.0054 & 0.0065 & 0.0002 \\ 0.0003 & 0.9925 & 0.0063 & 0 \\ -0.0030 & -0.0025 & 1.0071 & -0.0035 \\ -0.0046 & -0.0011 & 0.0093 & 0.9944 \end{bmatrix},$$

$$\boldsymbol{B}^1 = \begin{bmatrix} 0.4565 & 0.7132 & -0.3372 \\ 0.2529 & 0.5025 & -0.1768 \\ -0.0991 & 0.2829 & 0.2101 \\ 0.0196 & 0.1951 & 0.1526 \end{bmatrix},$$

$$\boldsymbol{B}^2 = \begin{bmatrix} 0.0590 & -0.0169 & 0.5376 \\ 0.1541 & 0.1590 & 0.2107 \\ 0.0760 & 0.0264 & 0.3059 \\ -0.1752 & 0.1992 & 0.3504 \end{bmatrix},$$

$$\boldsymbol{B}^3 = \begin{bmatrix} 0.1647 & 0.1054 & -0.0362 \\ 0.0853 & 0.1502 & -0.0291 \\ 0.0295 & 0.1020 & 0.0665 \\ 0.0184 & 0.0513 & 0.1089 \end{bmatrix},$$

$$\boldsymbol{B}^4 = \begin{bmatrix} 0.0090 & 0.3650 & -0.0617 \\ 0.2099 & 0.4394 & -0.2910 \\ 0.2848 & 0.3058 & -0.2542 \\ 0.2192 & 0.2708 & -0.2352 \end{bmatrix}, \boldsymbol{L}^i = -\boldsymbol{B}^i, \ \forall_{i \in \{1,\ldots,5\}},$$

$$\boldsymbol{B}^5 = \begin{bmatrix} 0.0383 & -0.0081 & 0.3631 \\ 0.3795 & -0.2730 & 0.2910 \\ 0.4411 & -0.4493 & 0.3601 \\ 0.7984 & -0.4391 & -0.1311 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The reference input is defined by

$$\boldsymbol{u}_{k,1} = \boldsymbol{u}_{k,2} = \boldsymbol{u}_{k,3} = \begin{cases} 0.4 & k < 3600 \\ 0.8, & k \geqslant 3600 \end{cases}$$

The actuators fault scenario, i.e. a multiplicative decrease of the performance of the heaters, is described as follows

$$\boldsymbol{f}_k = \mathrm{diag}(r_{k,1}, r_{k,2}, r_{k,3})\boldsymbol{u}_{f,k}$$

where

$$r_{k,1} = \begin{cases} 0, & k < 1800 \\ 0.2, & k \geqslant 1800 \end{cases}$$

$$r_{k,2} = \begin{cases} 0, & k < 1200 \\ 0.15 + 0.13 * sin(k/150), & k \geqslant 1200 \end{cases}$$

$$r_{k,3} = \begin{cases} 0, & k < 1500 \\ \frac{1}{15}(2 + cos(k/320) + 0.5 * sin(k/160) & \\ +0.5 * cos(k/640) + 0.3 * sin(k/60)), & k \geqslant 1500 \end{cases}$$

Also saturated inputs are considered with values in the range of $[0, 1]$. Therefore regulator for FTC system was designed with additional constraint condition with

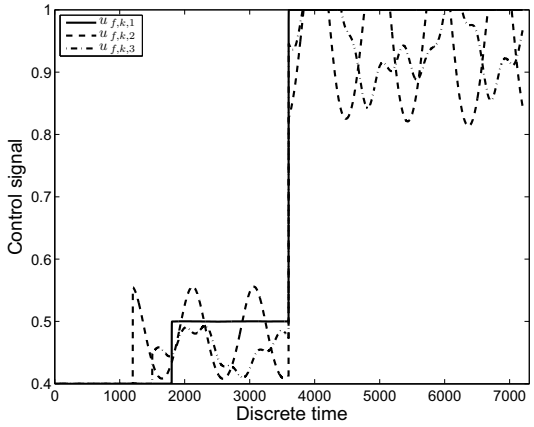


Figure 2. PDC control law – trajectory of $\boldsymbol{u}_{f,k}$

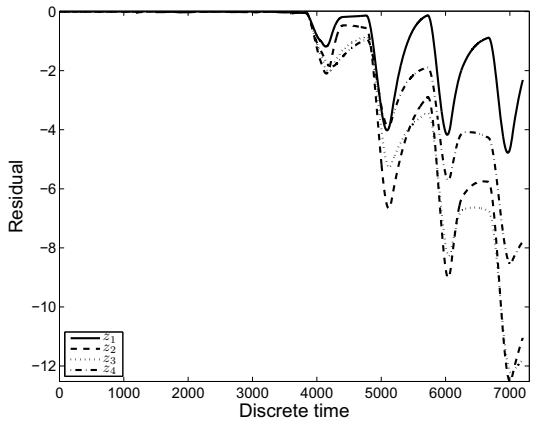


Figure 3. PDC control law – residuals

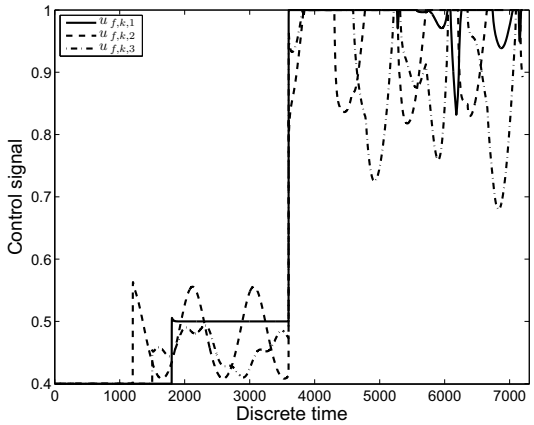


Figure 4. Exact MPC control law – trajectory of $\boldsymbol{u}_{f,k}$

parameters $\gamma = 20$ (initial tracking error $\boldsymbol{e}_0$) and $\lambda = 1$ (input constraints) and following weighting matrices, $\boldsymbol{Q}_R = 0.04\boldsymbol{I}$ and $\boldsymbol{R}_R = \boldsymbol{I}$. For MPC control horizon $T_c = 30$ is used and for fast MPC the number of iterations chosen were equal 5 with a barrier parameter $\kappa = 0.01$. Figs. 2 and 3 present the results achieved for the proposed FTC strategy (with $\boldsymbol{\alpha}_k = \boldsymbol{I}$) based on PDC control law. Similarly the results for exact MPC are presented on Fig. 4 and Fig. 5, whereas fast MPC strategy is shown on Fig. 6 and Fig. 7. The PDC control law and exact MPC were implemented using the generic optimization solver SeDuMi

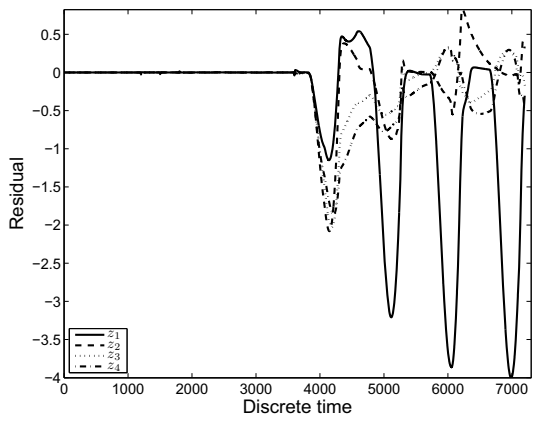


Figure 5. Exact MPC control law – residuals

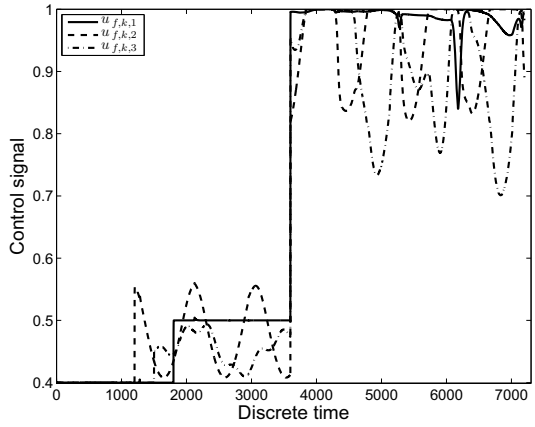


Figure 6. Fast MPC control law – trajectory of $\boldsymbol{u}_{f,k}$

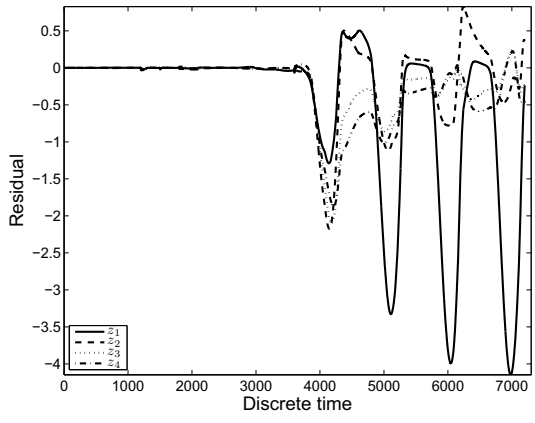


Figure 7. Fast MPC control law – residuals

Sturm (1999), called by YALMIP. SeDuMi is a state-of-the-art primal-dual interior-point solver that exploits sparsity. The simulations were performed in MATLAB on a 2GHz Intel Core 2 Quad Q9000 processor running Windows 7 x64 version. Given a problem at hand, an exact MPC required on average 283.6ms (with a maximum time 428.0ms) to solve a problem. (The reported times however include only SeDuMi CPU time, no YALMIP overhead, etc.) Whereas fast MPC took on average 1.8ms (with a maximum time 7.0ms) to solve a problem. It is clearly seen that even for small number of state dimensions it's one to two order of magnitudes faster. The exact MPC at best

allows for control rates of few hertz, while on the contrary fast MPC allows for control rate higher than 100Hz.
The comparisons of residuals $\boldsymbol{z}_k = \boldsymbol{x}_{f,k} - \boldsymbol{x}_k$ on Figs. 3, 5, 7 clearly shows that a sub-optimal PDC control law, that takes into account worst case scenario, is clearly worse than either MPC method. It can bee seen that for small order of fault magnitudes and control signal it works fine, but when control inputs goes into saturated state due to the faults, then the residual quickly rises, which is understandable. If there is no clear redundancy of actuators system, then the system will undoubtedly diverge from nominal state at some time point. But it allows almost 12 degree drop in temperature compared to nominal state and although it allows for return of the faulty system to its nominal state it does it very slowly and the constant overall decline of the system can be seen. Yet it should still be able to allow some time for system to persist in this degraded performance state, and thus give more time for eventual repairs. On the other hand both MPC policies (the exact one being slightly better in overall performance, but it hardly can be seen on the figures), allows for much more confidence in FTC system. We can see that in a worst scenario they allow only 4 degree drop in temperature, which is three times better than control achieved by only PDC control. It allows quite high rates of return to nominal state and even allows for a short periods of almost nominal (free of degraded performance) system state, when the PDC control was clearly in a degraded state.
As for a control trajectories seen on Figs. 2, 4, 6 it can be observed that inputs are often in a saturated state. Where there were no constraints involved all FTC strategies fared well. But in a saturated case the PDC control is much less active compared to either of MPC strategies. This is due to a fact that PDC control law is added to nominal input and fault correcting term, which could not be involved in its design. Whereas both MPC methods take into account current fault estimate and nominal input, which allows them to use the system matrices to devise an optimal control, permitting for correction of nominal input. It can be clearly seen that some inputs fall below nominal control input level. The difference in control strategies between exact and fast MPC is hardly noticeable, though it can be said that the fast MPC is somewhat less aggressive (compare $\boldsymbol{u}_{f,k,1}$), but it should be noted also that faults in this example where depending on current control action so the exact realisation of faults scenario acting on a system, where slightly different — the more aggressive control action, then greater absolute value of fault. Nevertheless as was mentioned above, both MPC strategies performed with a very similar performance, but fast MPC allows for much greater control rates and thus implementation of FTC strategies for a greater number of systems.

5. CONCLUSIONS

In this paper, an active FTC strategy has been proposed. This approach has been developed in the context of to T-S fuzzy systems.The key contribution of the proposed approach is an integrated FTC design procedure of the fault identification and fault-tolerant control schemes. Design procedure also allows to include input constraints into FTC system. Fault identification is based on the use of an observer. Once the fault have been identified, the FTC controller is implemented as a state feedback controller. This controller is designed such that it can stabilize the faulty plant using Lyapunov theory and LMIs and allows to minimize quadratic objective function similar. Lastly there is provided a way of predictive fault tolerant control of T-S fuzzy systems. Illustrative example for non-linear system described by T-S fuzzy models is provided that show the effectiveness of the proposed FTC approach.

REFERENCES

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. (2003). *Diagnosis and Fault-Tolerant Control.* Springer-Verlag, New York.

Boyd, S., Ghaoui, L.E., Feron, E., and Balakrishnan, V. (1994). *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics.* SIAM.

Dziekan, L., Witczak, M., and Korbicz, J. (2009). Active fault-tolerant control design for takagi-sugeno fuzzy systems. In *Proc. 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, 450–455. Barcelona, Spain.

Hui, S. and Zak, S. (2005). Observer design for systems with unknown input. *International Journal of Applied Mathematics and Computer Science*, 15(4), 431–446.

Korbicz, J., Kościelny, J., Kowalczuk, Z., and Cholewa (Eds.), W. (2004). *Fault diagnosis. Models, Artificial Intelligence, Applications.* Springer-Verlag, Berlin.

Maciejowski, J. (2002). *Predictive Control with Constraints.* Prentice Hall.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., and Scokaert, P.O.M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789–814.

Rong, Q. and Irwin, G. (2003). LMI-Based controller design for discrete polytopic LPV systems. In *Proc. European Control Conf.*

Sturm, J. (1999). Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12, 625–653. Version 1.3 available from `http://sedumi.ie.lehigh.edu/`.

Takagi, T. and Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Systems, Man and Cybernetics*, 15(1), 116–132.

Wang, H., Tanaka, K., and Griffin, M. (1996). An approach to fuzzy control of nonlinear systems: stability and design issues. *IEEE Transaction on Fuzzy Systems*, 4(1), 13–23.

Wang, Y. and Boyd, S. (2010). Fast model predictive control using online optimizations. *IEEE Transactions on Control Systems Technology*, 18(2), 267–278.

Witczak, M. (2006). Advances in model-based fault diagnosis with evolutionary algorithms and neural networks. *International Journal of Applied Mathematics and Computer Science*, 16(1), 85–99.

Witczak, M. (2007). *Modelling and Estimation Strategies for Fault Diagnosis of Non-linear Systems.* Springer-Verlag, Berlin.

Zhang, Y. and Jiang, J. (2003). Bibliographical review on reconfigurable fault-tolerant control systems. In *Proc. IFAC Safeprocess, Washington, D.C.*, 265–276.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Communication Gains Design in a Consensus Based Distributed Change Detection Algorithm

Nemanja Ilić, * **Srdjan S. Stanković** *

* *Faculty of Electrical Engineering, University of Belgrade, Belgrade, Serbia (e-mail: in085049p@student.etf.rs, stankovic@etf.rs).*

Abstract: In this paper a consensus based distributed recursive algorithm for real time change detection using sensor networks is considered. It is shown how the consensus gains can be designed by using linear programming. Cases of constant and random consensus gains are distinguished. Convergence of the algorithm to the optimal centralized solution defined by a weighted sum of the results of local signal processing is discussed. Simulation results illustrate characteristic properties of the algorithm.

Keywords: Sensor networks, distributed detection, recursive algorithms, linear programming.

1. INTRODUCTION

Signal processing using *distributed sensors* is a field of growing interest, having in mind the low cost and increased computational capabilities of sensors, as well as the availability of high speed *networks* connecting the sensors (*e.g.*, Vishwanathan and Varshney (1997)). One of the typical tasks of *sensor networks* is *distributed detection*. In the case of detection of changes in the monitored environment, it is often desirable to have a possibility to test the decision variables in real time at any node in the network, and not only at predefined fusion nodes.

There have been some recent attempts to apply *consensus techniques* to the distributed detection problem. However, the underlying assumption is that the dynamic agreement process starts *after all data have been collected*. In Stanković et al. (2009, 2007) algorithms for distributed state and parameter estimation have been proposed by combining local overlapping decentralized estimation schemes with a dynamic consensus algorithm. Decentralized fault detection and isolation observer based on a combination of local observers and a consensus strategy has been proposed in Ilić et al. (2010). In Stanković et al. (2010) a distributed *real time* change detection algorithm based on a consensus has been introduced. It can be aimed at distributed fault detection based on the obtained residuals from the aforementioned decentralized observer but its design allows general application in the field of distributed change detection via sensor networks. Analogous algorithms for distributed detection based on the "running consensus" methodology have been proposed and discussed in Braca et al. (2008).

In this paper a consensus based algorithm for *distributed change detection* while monitoring the environment through a wireless sensor network is discussed (Stanković et al. (2010)). Specific requirements regarding consensus matrices that should be satisfied in order to achieve convergence of the algorithm to the optimal centralized solution are analyzed. It is shown how the communication gains used by the algorithm can be obtained by linear programming, starting from the selected weights of local decision variables, for both constant and random consensus gains. In the case of simple measurement models, defined in the form of a sum of a parameter and the measurement noise, the weights can be derived on the basis of the *a priori* known local parameter jumps and the measurement uncertainty.

The outline of the paper is as follows. In Section 2 a distributed consensus based change detection scheme is described. Subsection 3.1 is devoted to the convergence analysis assuming constant consensus gains, while in subsection 3.2 the convergence results are extended to the case of random consensus gains. Communication gains design and simulation results are discussed in Section 4.

2. DISTRIBUTED CHANGE DETECTION ALGORITHM

Consider a sensor network containing n nodes, where each node collects locally available measurements and generates at each discrete time instant t a scalar quantity $x_i(t)$, $i = 1, \ldots, n$, directly, or as a result of local signal processing. We shall consider in the sequel $\{x_i(t)\}$ as mutually independent stationary random sequences with means $E\{x_i(t)\} = m_i$ and covariances $r_i(\tau) = E\{(x_i(t) - m_i)(x_i(t+\tau) - m_i)\}$. We shall assume that the network is aimed at change detection purposes.

The simple model usual for the domain of hypotheses testing is assumed: $x = m + \epsilon$, where $x = [x_1, \ldots, x_n]^T$, $m = [m_1, \ldots, m_n]^T$ and $\epsilon = [\epsilon_1, \ldots, \epsilon_n]^T$, with $\epsilon \sim \mathcal{N}(0, \Sigma)$, where $\Sigma = \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_n^2\}$. Assuming that $m = \theta^0 = 0$ in the case of no change and that $m = \theta^1 = [\theta_1^1, \ldots, \theta_n^1]^T$, where $\theta_i^1 > 0$ for some i in the case of change, we can calculate the log likelihood ratio for the data set containing $x(t)$, $t = 1, \ldots, N$, and obtain

$$L(N) = \sum_{t=1}^{N} \log \frac{p_{\theta^1}(x(t))}{p_{\theta^0}(x(t))} = \theta^{1T} \Sigma^{-1} \sum_{t=1}^{N} (x(t) - \frac{1}{2}\theta^1) \quad (1)$$

(Ding (2008); Basseville and Nikiforov (1993)). Starting from (1), one can apply the general methodology for constructing *on-line change detection* algorithms belonging to the *geometric moving average control charts* (Basseville and Nikiforov (1993)) and obtain the global decision function for the whole network, generated recursively by

$$s_c(t+1) = \alpha s_c(t) + (1-\alpha)\sum_{i=1}^{n} w_i x_i(t+1), \qquad s_c(0) = 0 \quad (2)$$

where $0 < \alpha < 1$ is the *forgetting factor* and $w_i = k\theta_i^1 \sigma_i^{-2}$, with $k = (\sum_{i=1}^{n} \theta_i^1 \sigma_i^{-2})^{-1}$, are the components of the vector $w^T = k\theta^{1T}\Sigma^{-1}$. It is important to notice that the algorithm (2) requires a *fusion center*.

The considered consensus based *distributed change detection* algorithm (Stanković et al. (2010)) *does not require a fusion center* and output of any preselected node can be used as a representative of the whole network and be tested w.r.t. a pre-specified *common threshold*. The basic assumption for this algorithm is that the nodes of the network are connected in accordance with an $n \times n$ time varying matrix $C(t) = [c_{ij}(t)]$ satisfying $c_{ij}(t) \geq 0$, $i \neq j$ and $c_{ii}(t) > 0$, $i, j = 1, \ldots, n$, which formally represents the weighted adjacency matrix for the underlying time varying graph representing the network, and that $C(t)$ is *row stochastic* for all t. The algorithm generates the *vector decision function* of the network, denoted as $s(t) = [s_1(t), \ldots, s_n(t)]^T$:

$$s(t+1) = \alpha C(t)s(t) + (1-\alpha)C(t)x(t+1), \qquad s(0) = 0. \quad (3)$$

Notice that the consensus matrix $C(t)$ performs "convexification" of the neighboring states for each node and enforces in such a way (under appropriate conditions) consensus between all the nodes. In such a way, after achieving the condition that $s_i(t) \approx s_j(t)$, $i, j = 1, \ldots, n$, change detection can be done by testing $s_i(t)$ for any preselected i with respect to a given common threshold λ_c, provided (3) gives a good approximation of $s_c(t)$ generated by (2).

3. CONVERGENCE ANALYSIS

3.1 Constant Consensus Gains

We start from the following assumptions:
A1) C has the eigenvalue 1 with algebraic multiplicity 1;
A2) $\lim_{i\to\infty} C^i = \mathbf{1}w^T$.
Under A1), C^i converges when i tends to infinity to a nonnegative row stochastic matrix with equal rows, *e.g.* Olfati-Saber et al. (2007). Knowing w from the general problem setting based on the centralized detection strategy, we can construct C satisfying A2) by solving for C the linear equation known from the theory of stationary Markov chains

$$w^T C = w^T, \quad (4)$$

under the constraints that: 1) preselected elements of C are equal to zero (indication that there can be no communication between the corresponding nodes) and 2) matrix C is row stochastic, satisfying the given assumptions.

The error vector between the vector $s(t)$ and the state of the optimal centralized scheme will be defined as $e(t) = s(t) - \mathbf{1}s_c(t)$, where $\mathbf{1} = [1 \cdots 1]^T$. Iterating (3) and (2) back to the zero initial conditions, we obtain that $e(t) = (1-\alpha)\sum_{i=0}^{t-1} \alpha^i [\varphi(t-1, t-i-1) - \mathbf{1}w^T]x(t-i)$, where $\varphi(i,j) = C(i)\cdots C(j)$, $i \geq j$. The focus of the analysis is placed on the error covariance matrix, defined as

$$Q(t) = E\{e(t)e(t)^T\} - E\{e(t)\}E\{e(t)\}^T \quad (5)$$

and the following theorem (Stanković et al. (2010)) provides an insight into its asymptotic properties.

Theorem 1. Let assumptions A1) and A2) hold, together with:

A3) $\max_i \sum_{\tau=0}^{t} |r_i(\tau)| \leq K; \quad 0 < K < \infty$.

Then,

$$\max_{i,j} Q_{ij}(t) \leq O((1-\alpha)^2),$$

where $Q_{ij}(t)$ are the elements of $Q(t)$ in (5).

3.2 Random Consensus Gains

The results from the previous section related to the case of time invariant consensus matrices will be generalized here to time varying random consensus matrices, case of substantial importance from the point of view of applications of the proposed algorithms in real sensor networks.

We shall assume that the sequence $\{C(t)\}$ is a sequence of mutually independent identically distributed random matrices independent from the sequence $\{x(t)\}$, such that matrix $C(t)$ is realized at each discrete time instant t as matrix $C^{(k)}$ with probability p_k, $k = 1, \ldots, N$, $N < \infty$, $\sum_{k=1}^{N} p_k = 1$; the realization matrices $C^{(k)} = [c_{ij}^{(k)}]$, $i, j = 1, \ldots, n$, are constant nonnegative row stochastic matrices, satisfying $c_{ii}^{(k)} > 0$. The mathematical expectation of $C(t)$ is given by $\bar{C} = E\{C(t)\} = \sum_{k=1}^{N} C^{(k)} p_k$.

We shall design the algorithm and analyze its convergence using the following assumptions:
B1) $\bar{C}$ has the eigenvalue 1 with algebraic multiplicity 1;
B2) $\lim_{i\to\infty} \bar{C}^i = \mathbf{1}w^T$.
The design of the detection algorithm incorporates definition of the realization matrices $C^{(k)}$, together with the corresponding probabilities p_k. According to assumption B2), we have to solve solve for p_k and the elements of $C^{(k)}$, $k = 1, \ldots, N$, the following nonlinear equation

$$w^T \bar{C} = w^T \sum_{k=1}^{N} (I + C^{[i,j]}) p_k = w^T, \quad (6)$$

where $w \geq 0$ is a given weighting vector resulting from the centralized strategy (2).

Error covariance matrix central for the convergence analysis in the case of random consensus gains is defined as

$$Q(t) = E_\varphi\{E_X\{e(t)e(t)^T\} - E_X\{e(t)\}E_X\{e(t)\}^T\} \quad (7)$$

and the following theorem concerned with it can be proved.

Theorem 2. Let assumptions B1), B2) and A3) hold. Then, in the case of random consensus matrices satisfying the above assumptions

$$\max_{i,j} Q_{ij}(t) \leq O((1-\alpha)),$$

where $Q_{ij}(t)$ are the elements of $Q(t)$ in (7).

It is important to notice at this point that the result of Theorem 2, when compared to the result of Theorem 1, shows that randomness of the network causes an increase of the error covariance, as it could be expected. Illustrations of the resulting detection efficiency, which is still very satisfactory, will be given in the next section.

4. COMMUNICATION GAINS DESIGN AND SIMULATION RESULTS

The starting point in the design of the communication gains, for both constant and random consensus gains, is the weight vector w. Based on the a priori known local parameter jumps (θ_i^1) and the measurement uncertainty (σ_i^2) its components are determined by $w_i = \frac{\theta_i^1 \sigma_i^{-2}}{\sum_{i=1}^n \theta_i^1 \sigma_i^{-2}}$.
In the case of constant C communication gains are obtained by solving the linear equation (4) under the following constraints: 1) elements of $C = [c_{ij}]$ are bounded, $0 < c_{ij} < 1$; 2) C is row stochastic and 3) some elements of C are zero (no communication between the corresponding nodes where nodes represent, *e.g.*, randomly spatially distributed agents within the square area, nodes are connected if their distance is less than some predetermined threshold, in this case half of the side of the square).
Problem of finding the appropriate sequence $\{C(t)\}$ in the case of random consensus gains is more complex than the one discussed above when $C(t) = C$. The setting of matrices $\{C(t)\}$ from previous section obviously encompasses the asynchronous asymmetric gossip algorithm with one message at a time; *e.g.*, if the node j communicates to the node i, the corresponding realization simply has the form $C^{(k)} = I + C^{[i,j]}$, where $C^{[i,j]} = [c^{[i,j]}]$ is an $n \times n$ matrix which contains zeros at all places except the (i,j)-th place where it contains γ_{ij} and the (i,i)-th place where it contains $-\gamma_{ij}$, $0 < \gamma_{ij} < 1$. Synchronous asymmetric gossip algorithms can also be treated by constructing the corresponding realizations in an appropriate way using realizations $C^{(k)}$ with more nonzero off-diagonal elements. Communication faults can be obviously modelled similarly, by forming realizations $C^{(k)}$ in accordance with the structure of faults (see, *e.g.*, Stanković et al. (2009)). Just to explain the main idea, let us consider a sensor network with $n = 3$ nodes where every node is connected to the other two. Assuming that we have the asymmetric asynchronous gossip algorithm with one communication at a time, there are $N = 6$ possible realization matrices $C^{(k)}$, $k = 1, \ldots, N$, so we have $C^{(1)} = I + C^{[1,2]}$, $C^{(2)} = I + C^{[1,3]}$, $C^{(3)} = I + C^{[2,1]}$, $C^{(4)} = I + C^{[2,3]}$, $C^{(5)} = I + C^{[3,1]}$, and $C^{(6)} = I + C^{[3,2]}$. Consequently, one obtains that $\bar{C} = \sum_{k=1}^N C^{(k)} p_k =$

$$\begin{bmatrix} 1-\gamma_{12}p_1-\gamma_{13}p_2 & \gamma_{12}p_1 & \gamma_{13}p_2 \\ \gamma_{21}p_3 & 1-\gamma_{21}p_3-\gamma_{23}p_4 & \gamma_{23}p_4 \\ \gamma_{31}p_5 & \gamma_{32}p_6 & 1-\gamma_{31}p_5-\gamma_{32}p_6 \end{bmatrix}.$$

Solving the equation (6), compared to the problem in (4), involves more degrees of freedom, and we have two options: a) to adopt values of the probabilities p_k - one can, *e.g.*, set $p_k = 1/N$, $k = 1, \ldots, N$. In this case the resulting constraint is that the non-diagonal elements of $\bar{C}$ are bounded within the interval $(0, 1/N)$ while diagonal elements are bounded within $(0, 1)$, the additional two are constraints 2) and 3) introduced above, holding for $\bar{C}$ instead of C;
b) to adopt values of the elements of $C^{(k)}$, *i.e.*, the set of parameters γ_{ij}. The first choice is to set all γ_{ij} to one value (the case when $\gamma_{ij} = 0.5$ corresponds to the randomized "gossip" algorithm, with the difference that there is only one communication at a time (instead of pairwise communication), leading to nonsymmetric consensus matrices (instead of symmetric)). Now the additional constraint is that the sum of all non-diagonal elements in $\bar{C}$ is equal to that common value of γ_{ij} (as can be seen from the example above), the other three constraints are those corresponding to the case when $C(t) = C$, holding for $\bar{C}$ instead of C.

From the point of view of application of the proposed algorithm in real sensor networks the case a) corresponds to the case when possible communications between the nodes occur with some predefined probabilities (*e.g.*, each possible communication occurs with the same probability) while the case b) encompasses situations where the nodes send their data with some predefined weights each time communication occurs (*e.g.*, whenever communication occurs nodes send their data with the same weights). More degrees of freedom in (6) allow addition of a constraint that all columns (or rows) in $\bar{C}$ have equal elements (excluding diagonal elements). In practical applications this means that all possible communications when a node sends (or receives) data happen with the same probabilities or, alternatively, with the same weights. As can be seen, different ways of design of a sequence $\{C(t)\}$ open up many new possibilities for the design of sensor networks in practise.

Sensor network with $n = 10$ nodes is considered in the simulation, where the means θ_i^1 are randomly taken from the interval $(0, 1]$, and variances σ_i^2 randomly taken from the interval $[0.5, 1.5]$ ($\theta_i^0 = 0$ in the case of no change), $i = 1, \ldots, n$. Communication gains in the case of constant C are obtained by solving the linear equation (4), as described above. For the random consensus gains case the sequence $\{C(t)\}$ was obtained by putting all γ_{ij} to be 0.5 (one can also, using the above described methodology, adopt values of p_k), with the addition that all columns in $\bar{C}$ have equal elements (excluding diagonal elements). We analyze properties of the proposed algorithm w.r.t. forgetting factor α and cases of constant and random consensus gains are distinguished. The moment of change is chosen to be $t = 200$. The algorithm effectively achieves very similar behavior of all of the nodes, with local decision functions getting closer to the global decision function as $\alpha \to 1$. In Fig. 1 one realization of the global decision function is given by magenta line, together with the decision function of one randomly selected node (blue line) for $\alpha = 0.9$ and $\alpha = 0.99$; case of constant consensus gains is shown on the left while case where consensus gains are random is shown on the right. In addition, in Fig. 2 mean $\pm$ one standard deviation of the global decision function is given by dashed lines, together with the decision function of one randomly selected node (solid line). In accordance with the convergence analysis in the previous section, the introduction of randomness into consensus matrices increases error covariance. Nevertheless, the resulting detection efficiency is still satisfactory (for α close to 1), as can be also seen from Fig. 3 where estimation of the distribution of detection times (the moment of change is $t = 500$) is given for: global decision function (up), case of constant C (middle) and random consensus gains case (down).

ACKNOWLEDGEMENTS

This work has been supported by the EU FP7 Project "Power Systems Robustification Based on Fault Detection and Isolation Algorithms - PRODI" and Serbian Ministry of Science and Technological Development.

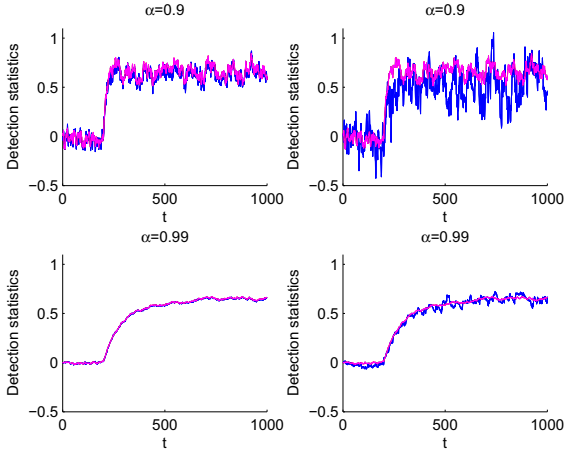


Fig. 1. Decision functions for one node (blue) and global (magenta) - one realization. Constant consensus matrix C - left; random $C(t)$ - right.

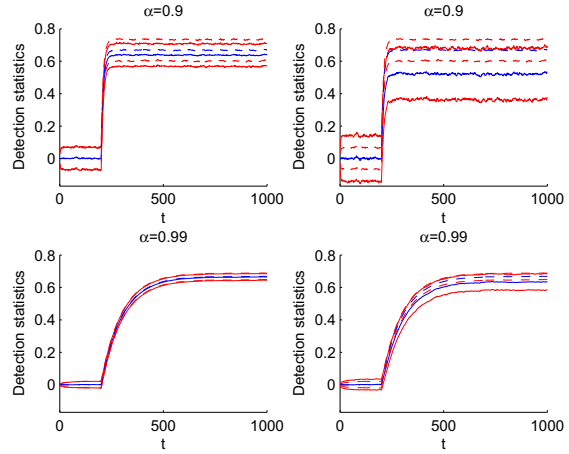


Fig. 2. Decision functions for one node (solid lines) and global (dashed) - mean ± standard deviation. Constant consensus matrix C - left; random $C(t)$ - right.

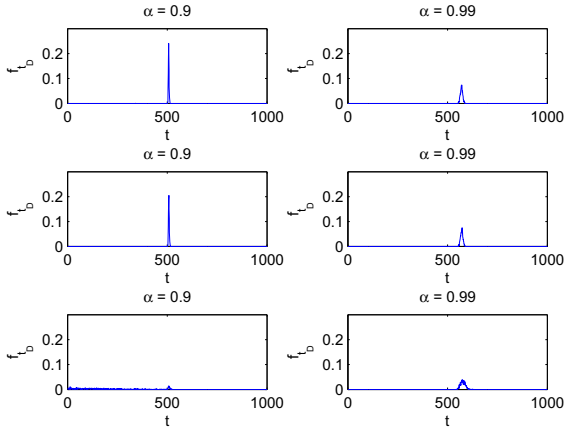


Fig. 3. Estimation of the distribution of detection times: global decision function (up); case of constant consensus matrix C (middle); random $C(t)$ (down).

REFERENCES

Basseville, M. and Nikiforov, L.V. (1993). *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall.

Braca, P., Marano, S., and Matta, V. (2008). Enforcing consensus while monitoring the environment in wireless sensor networks. *IEEE Trans. Signal Processing*, 56, 3375–3380.

Ding, S.X. (2008). *Model Based Fault Diagnosis Techniques - Design Schemes, Algorithms and Tools*. Springer Verlag.

Ilić, N., Stanković, M., and Stanković, S. (2010). Consensus based overlapping decentralized observer for fault detection and isolation. In *Proc. Melecon 2010 Conf.*

Olfati-Saber, R., Fax, A., and Murray, R. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95, 215–233.

Stanković, S., Ilić, N., Stanković, M., and Johansson, K.H. (2010). Distributed change detection based on a consensus algorithm. *Proc. 2nd IFAC Workshop on Distr. Estim. Contr. Networked Systems*.

Stanković, S.S., Stanković, M.S., and Stipanović, D.M. (2007). Decentralized parameter estimation by consensus based stochastic approximation. In *Proc. 46th IEEE Conference on Decision and Control*, 1535–1540.

Stanković, S.S., Stanković, M.S., and Stipanović, D.M. (2009). Consensus based overlapping decentralized estimation with missing observations and communication faults. *Automatica*, 45, 1397–1406.

Vishwanathan, R. and Varshney, P. (1997). Distributed detection with multiple sensors: Part i - fundamentals. *Proc. of the IEEE*, 85, 54–63.

Appendix A. CONSENSUS GAINS - NUMERICAL EXAMPLES

One realization of a sensor network from Section 4 gives the weights $w^T = [10.09\ 0.8\ 8.08\ 14.53\ 1.75\ 27.47\ 7.64\ 3.05\ 10.72\ 15.87] \cdot 10^{-2}$. The obtained consensus matrix $C(t) = C$ for one realization of spatial distribution of sensors is

$$\begin{bmatrix} 44.49 & 0 & 9.79 & 9.78 & 4.34 & 16.29 & 7.34 & 0 & 7.96 & 0 \\ 0 & 24.67 & 44.63 & 0 & 30.71 & 0 & 0 & 0 & 0 & 0 \\ 14.12 & 4.54 & 23.94 & 0 & 8.97 & 48.44 & 0 & 0 & 0 & 0 \\ 6.89 & 0 & 0 & 16.08 & 0 & 27.71 & 10.13 & 0 & 12.07 & 27.13 \\ 28.01 & 13.51 & 39.05 & 0 & 19.43 & 0 & 0 & 0 & 0 & 0 \\ 4.23 & 0 & 14.99 & 15.16 & 0 & 42.63 & 8.24 & 5.12 & 9.64 & 0 \\ 11.61 & 0 & 0 & 19.23 & 0 & 27.21 & 14.27 & 11.49 & 16.19 & 0 \\ 0 & 0 & 0 & 0 & 0 & 46.69 & 28.12 & 25.19 & 0 & 0 \\ 8.62 & 0 & 0 & 16.54 & 0 & 24.92 & 11.29 & 0 & 13.27 & 25.36 \\ 0 & 0 & 0 & 23.97 & 0 & 0 & 0 & 0 & 18.00 & 58.03 \end{bmatrix} \cdot 10^{-2}$$

It can be easily verified that $\lim_{i\to\infty} C^i = \mathbf{1}w^T$. In the case of random consensus gains the obtained mathematical expectation $\bar{C}$ of $C(t)$ is

$$\begin{bmatrix} 92.48 & 0 & 0.87 & 1.56 & 0.19 & 2.94 & 0.82 & 0 & 1.15 & 0 \\ 0 & 98.95 & 0.87 & 0 & 0.19 & 0 & 0 & 0 & 0 & 0 \\ 1.08 & 0.09 & 95.70 & 0 & 0.19 & 2.94 & 0 & 0 & 0 & 0 \\ 1.08 & 0 & 0 & 92.31 & 0 & 2.94 & 0.82 & 0 & 1.15 & 1.70 \\ 1.08 & 0.09 & 0.87 & 0 & 97.97 & 0 & 0 & 0 & 0 & 0 \\ 1.08 & 0 & 0.87 & 1.56 & 0 & 94.20 & 0.82 & 0.33 & 1.15 & 0 \\ 1.08 & 0 & 0 & 1.56 & 0 & 2.94 & 92.94 & 0.33 & 1.15 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.94 & 0.82 & 96.24 & 0 & 0 \\ 1.08 & 0 & 0 & 1.56 & 0 & 2.94 & 0.82 & 0 & 91.90 & 1.70 \\ 0 & 0 & 0 & 1.56 & 0 & 0 & 0 & 0 & 1.15 & 97.30 \end{bmatrix} \cdot 10^{-2}$$

One can easily verify that $\lim_{i\to\infty} \bar{C}^i = \mathbf{1}w^T$, and that sum of all non-diagonal elements is equal to 0.5.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Control of Independent Mobile Robots by Means of Advanced Monitoring, Diagnosis and Prediction.

L. Seybold*, J. Krokowicz**,
K. Patan***, R. Stetter****, A. Paczynski****.**

** RAFI GmbH & Co. KG, 88276 Berg, Germany*
(e-mail:lothar.seybold@rafi.de).
****University of Zielona Gora, 65-246 Zielona Gora, Poland.*
(e-mail:k.patan@issi.uz.zgora.pl).
*** Hochschule Ravensburg-Weingarten, 88241 Weingarten, Germany*
(e-mail: stetter;paczynski; j.krokowicz@hs-weingarten.de)

Abstract: The paper reports an innovative extension of the possibilities concerning monitoring, planning, control and diagnosis for mobile robots in production systems. This extension is realized through the integration of monitoring, planning, control and diagnosis in a well-considered concept and through vertical connection with supervising industrial levels and direct control system of the mobile robots. The developed exemplary application is based on a kinematic-dynamic prediction model which makes use of data collected in a common database and sends request for decisions to higher levels when it is necessary. The levels mentioned above form a hierarchical IT-structure which main purpose is increasing system efficiency and reliability by ensuring the best possible data flow as well as data processing. In this paper ways of increasing the performance of the control system operation of mobile robots, methods of information collecting and methods of sharing them in the entire structure are described. Furthermore ways of cooperation for the particular applications are considered. Finally, an example of using data in the robot prediction model, which was collected in the system during task planning and testing, is presented.

Keywords: mobile robot, torque/position control, predictive control, diagnosis, monitoring.

1. INTRODUCTION

1.1 Paper content

The presented research is based on the integration of the production system IT-infrastructure based on mobile robots as they are frequently used in production companies. The main emphasis is placed on ensuring the information flow as appropriate as possible. This flow is realized among all infrastructure components in order to enable high optimization possibilities of the system operation. The proposed sensible structure of the system is described in Section 1.2. Section 2 presents possible extensions of the system performance compared against the state of the art. Section 3 describes one exemplary system operation cycle in order to elucidate the main ideas behind the advanced control and diagnosis concept. The scenario is concentrated on the complex energy consumption optimization of multiple robots while the trajectories and driving behaviours are being predicted.

1.2 Industrial IT infrastructure – description

The highest position in the IT-structure (Fig.1) of production companies is usually the process and business Management level which consists of programs supporting resource management, customer and supplier relations and product management. The main task of this level is creating analyses, supporting business decisions, parameters tuning and activities planning. On this level system **Enterprise Resource Planning** (ERP) is often applied, its main task is centralization of the data in one place in order to effectively use them. ERP connects data come from various applications working in different functional areas such as finance, marketing and sales, human resource and manufacturing applications in order to ensure proper communication in the Business Management domain. The next lower level includes Process Control Systems which connect production and business management areas. Those systems are responsible for performing production plans. That level consists of various applications of supervision systems and production visualization. The main tasks of those systems are supervision and management of technological processes. MES (**Manufacturing Execution Systems**) use systems, technologies, applications, electronic devices and automation elements. This enables collecting data directly form production and transfer to the Business Management field. In order to enable communication between all levels the use of an uniformed and unified communication standard is common. OPC (OLE - Object Linking and Embedding - for process control) is an open communication standard used in

industrial automation and in information systems for Process and Business Management. OPC allows using uniformed access methods and data descriptions (interface) for technological processes. Those methods are independent of type and data source. For many packages applications server OPC supplies data in a uniformed way from devices controlling and supervising the technological process. The OPC mission is the definition of the common interface which once made can be used by any business client, SCADA applications, HMI or any other application packages. The use of that standard eliminates necessity of special driver implementations in order to allow access to process data for any other software. If a server OPC will be created for a specific device, it can be also used again by any other application which is an OPC client.

The specific concept presented in this publication can be found on the third level. Here a system is located which integrates **Monitoring, Planning, Control and Diagnosis** (MPCDS). The main task of this system is the supervisory control of a highly flexible manufacturing system based on mobile robots. This system must have the possibility to easily connect to the existing structure. There is no necessity to create separate systems collecting data from individual devices, Business Management department etc. Those data can be taken using the unified standard OPC which is widely used in industry. Basing on the computations results the MPCD system supplies necessary parameters for the movement planning level as well as reports for Process and Business Management and Process Control Systems.

On the three lower levels the **task and trajectory planning** for mobile robot systems and individual robots can be found, as well as the individual mobile robots including their own distributed intelligence and a simplified mathematical (kinematic and dynamic) model of their behavior and the single device which also may dispose of their own local intelligence in the sense of ubiquitous computing.

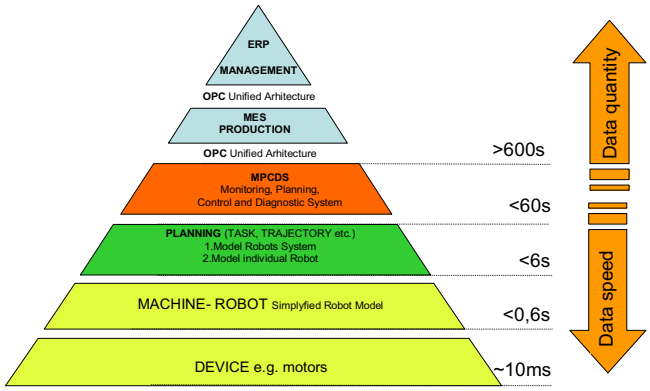


Fig. 1. Hierarchical industrial IT infrastructure.

2. STATE OF THE ART

2.1 Autonomous driving

Autonomous vehicles (also called AGV (autonomous guided vehicles)) are vehicles driving mostly on wheels which can perform complex tasks in real environments and work for an extended period of time without human guidance. For the sake of simplicity in this publication such autonomous vehicles will be referred to as ”mobile robots”. Different mobile robots can be fully or partially autonomous. The required degree of autonomous operation is defined by the place of operating; the highest degree of independency is desired in inaccessible environments where human assistance is particularly inconvenient (small spaces, harmful environments) or even impossible e. g. space exploration missions. Mobile robots working autonomously must also avoid situations when they could be dangerous especially while interacting with humans, but also have to avoid collisions with surrounding objects, other robots and also have to avoid damaging themselves. Nowadays a number of intensive researches in the domain of autonomous driving for wheeled mobile vehicles are carried out. Very extensive and interdisciplinary knowledge was complied in these works. An overall view of the complex functional structure of the design and development of intelligent mobile robots is presented by e. g. Yavuz (2007) and Ziemniak et al. (2009a). Usually researchers are focused on one selected issue, considering it in various aspects and applying various methods to solve problems in different applications regions. In order to describe the latest directions in this field we can divide this domain into a few main branches like perception, localization, planning and navigation. The following part of this section is structured according to these branches.

The notion **Perception** summarizes that mobile robots take information about the environment by means of measurements using various kinds of sensors and methods of measuring. Crucial to proper estimation of the distance are methods and algorithms to extract information from those measurements. It is necessary to pay attention for difficult noisy data filtering given by sensors Holland (2003). A very good overview can be found in Sigewart et al.(2004) and more detailed information about many kinds of the sensors which have application in mobile vehicles in Borenstein et al.(2004). In this field we can find applications for a wide range of knowledge such as signal analysis and specialized bodies of knowledge such as computer vision to properly employ a multitude of sensor technologies.

Under the notion **Localization** the fact that mobile robots need to exactly determine their position in the environment in order to ensure performing the task is considered. That problem has received high research attention and, as a result, considerable advances have been made on this field. It is necessary to involve computer algorithms, information theory, artificial intelligence and probability theory in order to find the position of the vehicle effectively.

Many approaches have been proposed for solving **Planning and Navigation** problems for single mobile robots as well as multiple mobile robots which interact as multi agent systems Tian-Tian et al. (2008). Two main branches dealing with robot movement can be distinguished – the first focuses on path planning problems and the other one on obstacle avoidance.

Path/motion planning is a strategic problem-solving competence. An application implemented in the robot's control system has to calculate trajectories to enable to reach the goal position as efficiently and reliably as possible. In order to drive the vehicle must follow the previously planned path which could be optimized considering different aspects (e.g. by means of costs functions) such as minimum time, energy consumption, etc. (Posedlowski et al. 2001), (Hahour (2008). Important are also physical limitations – acceleration limits caused by limited friction force between the ground and wheels (Lopetic et al.2003) or limitations of the acceleration and centrifugal forces which help saving energy and also help lowering uncertainties related to the robot movement and secure transported materials (Ziemniak et al. 2009b). The second competence is obstacle avoidance because a vehicle must, on the basis of information given by sensor readings, change/modulate its trajectory in real time in order to avoid collisions. Often a separation of planning and navigations problems seems not to be suitable because they are strongly connected to each other (Sigewart et al. 2004).

2.2 Monitoring, planning, diagnosis and control

Monitoring is a process consisting of the data collection from the entire industrial production system. It includes continuous measurements as well as reading and processing of data which is received from the mobile robot's sensors (and additional sensors i.e. cameras mounted on the walls). Moreover data is collected also by cooperating with the devices of the mobile robots. Considerable attention is paid on taking advantage of existing sensors measurements instead of fitting additional ones. One can extend monitoring if all actuators are used as sensors and are combined with advanced model based methods. In the case of many devices monitoring is a function of its own. For this reason it is very often necessary to collect the information and than send it to an appropriate place. The data is recorded for subsequent decision process in diagnostic expert system (Pieczynski 2003). The Diagnostic application consists of appropriate computational tools for detecting information from the stream of the data received form the monitoring. That level also supervises actual state of the task completion. The monitoring in case of mobile robots provides data to enable the optimization of the entire system operation and in general for the sake of increasing safety, availability and reducing energy consumption. Information which provides monitoring is used on each level of industrial IT infrastructure.

Planning is the theoretical (pre-)processing of future activities and is probably the essential management function. Managing departments need to plan all resources such as raw materials, external parts, personnel and all equipment including mobile robots in order to produce effectively and efficiently.

Diagnosis is usually understood as the process of estimating the object condition. Diagnosis is carried out by the estimation of important parameters and the determination what should be done in case when faults occur. It is possible to increase the meaning of the notion "diagnosis" if this level has both communication with the higher levels and access to all data which an advanced monitoring system is able to provide. Ensuring a suitable vertical information flow (see Fig. 1) this level is able to make more complex analyses. The diagnosis application has functions/tools to extract relevant information from entire industrial system and then to make appropriate decision or send reports to the right places. Moreover, the goal of an advanced diagnostic application is to check continuously how the system works instead of just checking if and what is wrong. This application is then able to supervise a system during normal work conditions. In that way the data collected before is used to verify if changed values of the measured parameters are caused by occurrence of the fault or normal gradually device wear. The advanced functions of diagnosis provide also information about incorrect behaviour of devices cooperating with mobile robots e. g. inaccurate load positioning on the mobile robot platform. In some cases that level needs to communicate with higher levels to ask which decision proposal should be chosen or ask for new task parameters for the mobile robots. In this case diagnosis is a supervising and decision process. The one of main tasks in the more conventional approach to diagnosis is **fault detection**. Applications of autonomous mobile robots have high requirements for reliability and safety. Fault diagnosis can be a critical task for mobile robots, especially when the vehicles operate in a hazardous and difficult to reach environments such as planetary rovers or while interacting with the human both in industrial tasks and traffic in the urban environment. It is possible to divide diagnosis into fault detection systems and fault tolerant systems. A **fault tolerant control** (FTC) system is a control system which enables to continue the operation in case of the expected occurrence of a fault. It is possible to reduce performance (drive abilities) and then the system is able to continue the task instead of failing completely, while some part of control system or hardware fails. Many FTC methods have been recently developed and many are based on active (analytical) systems called model based fault detection and isolation (FDI) scheme and are based on analytical redundancy. Those systems called high level protection systems are based on precise diagnostic information about the state of the system. Underlying is a online process and a control reconfiguration mechanism. Advantages of that FTC system are introduced e. g. by Kościelny et al. (2006). For purposes of robot's diagnosis advanced methods are used. Algorithms are based on analytical-mathematical model, heuristic (expert's knowledge or learning machines methods), artificial intelligence (neural networks, fuzzy logic) as well as on hybrid models which contain all mentioned methods (Merzouki et al. 2010). For our further investigations we use appropriate methods for robust and efficient real-time diagnosis which are proposed by Freitas et al. (2004). The analytical methods to create a diagnostic system, such as Kalman filtering and particle filters which are used widely in robotics are proposed by our research group in Zając et al. (2008). An artificial intelligence method using neural networks is also introduced in Zajac et al. (2009).

The notion **Control** is used in the sense of the regulation. It relies on adjusting input signals sent to the controlled object/structure in order to achieve the desired level of the output signal. Input values are estimated in the theoretical model

of the object. Control can be performed in open loop where input signals are dependent only on the predicting method. However, in engineering often systems working with feed back are used. Those systems are equipped with sensors which measure outputs signals and allow respecting those measurements in predicting process. On the base of the data collected by the sensors controller regulate a variable at a setpoint or reference value. In case of an autonomous guided vehicle motor outputs are modulated. It is done in order to improve the efficiency of the usage of energy, the robustness, the drive abilities and finally in order to achieve the desired trajectory (Lamon et al. 2004).

2.3 Safety of human beings

This section discusses the safety of human beings as this the most crucial concern for diagnosis and control of mobile robots. While AGV are increasingly becoming a part of usual industrial environments it is important to ensure the safety of human beings which have to interact with them. The safety of human beings has to be taken into consideration on each development level. On the stage of mechanical designing the design has to be optimized in order to decrease the risk of injury in the event of accident. Then vehicle's control system has to meet high requirements. In case the safety depends on the correct operation of control subsystems such as Electrical, Electronic or Programmable Electronic (E/E/PE) then the related system has to fulfill functional safety. The norm IEC (The International Electrotechnical Commission) 61508 specifies 4 levels (Safety Integrity Level SIL) of safety performance for a safety function.

It is one of the most prominent goals of diagnosis to find a fault and to aid the control system to make a decision how to perform further parts of a given task also regarding safety. Additionally it is possible to provide information for higher levels, what will make it possible to prevent the occurred dangerous situation in the future. It is necessary to take into account the human presence while paths for mobile robots are planned. In many cases "Double Safety" in the control systems is used. The most often Double Safety is demanded when the probability of occurrence of significant damages is present if a failure of the device would happen. When double safety methods are required, usually it is necessary to equip the respective device with two independent control systems or at least it is necessary to double some control circuits. That method is also used for storing data. It means that all data is stored in two places simultaneously in order to avoid loss of important data. It is necessary to comply with the legal regulations. AGV have to fulfil the conditions which contains requirements to structure such as electrical or protective equipment i. e. to cover norms of International Standard Organization (ISO) e. g. safety of the motion (ISO 10218) (2006).

2.4 Actuators as sensors

Large expenditures for diagnosis can be saved if actuators are directly used as sensors. An actuator is typically the mechanical or electromechanical device which converts energy into motion or applies a force. A sensor is the physical tool that measures physical quantities. It sends the information in a form possible to read by the converter and next to the measuring device. The most popular sensors deliver information in the form one of the electric quantities voltage, resistance or intensity. There are approaches in the literature using direct measurements of applied signals for actuators. These approaches are made in order to get information about actuator's state of work. The mobile robot drive motors can be controlled by special electronic control units. Those units control torques of the particular motors by adjusting suitable current and voltage. For this reason those units must measure both mentioned parameters. Thus it is possible to read appropriate data from these control units instead of measuring it again. For this purpose it is not necessary to integrate additional sensors and electrical circuits but is enough to adapt an appropriate application controlling the mobile robot. The advantage of using actuators as sensors is the possibility to ensure delivering additional information to the entire control system without making the system more complex. Usually in that way actuator load is measured. One example of using this method is described in Washington et al. (2000).

3. EXAMPLE OF THE SYSTEM OPERATING CYCLE

In the previous sections, possibilities of enhancing the system efficiency as well as ways of improving the gathering and processing of data were described. This section describes one example of a task performed by a mobile robot. This description is focused on explaining the cooperation between the robot's control system and the IT-Infrastructure. Presented are the stages of the essential data received for computation and testing in the mathematical robot's model. In those steps the optimization of the trajectory and the energy consumption is the major task. Considered in this prediction process are safety issues (compare section 2.3) in case of receiving the warning signal from MPCDS level or disruption in communication between the robot and the supervising MPCDS level. The first part of the presented solution takes into the consideration especially the important role of the management systems ERP and MES. In particular this solution disposes of possibilities of long-term supervision, parameterization and robot's motion optimization respecting management missions. Additionally the safety of the diagnostic system is thoroughly considered in order to be able to store the data about robot's operation according to the industrial software engineering safety norms. The core of this system is called MPCDS (Monitoring Planning Control Diagnosis System). In the example, the respective task is given by the MPCDS using the following parameters: distance, time (production cycle) and also current values of the parameters: maximum acceleration change, acceleration, maximal velocity, friction coefficient between ground surface and the robot's wheels. These data are sent to the robot's mathematical model and tested before sending it to the direct control system of the particular drive unit. In the first step, the possibility of performing this task is tested (feasibility test) and then as the second and last step the real mobile robot performs the task. In this example energy consumption is optimized which is especially important for battery powered systems such as usual AGV.

3.1 Schematic operation of the robot control system

The application checks at the beginning whether the mobile robot is able to perform task received from MPCDS level (Fig.2).

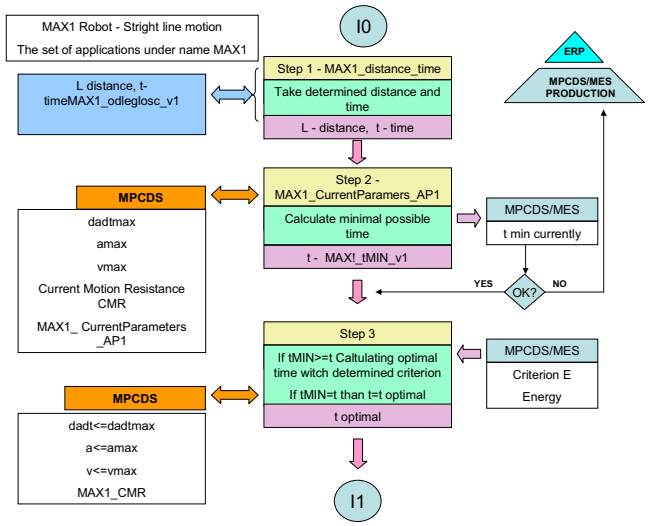


Fig.2. Testing of the task in the robot's model.

This check is possible by processing the received data in the robot's dynamic-kinematic model. The model allows generating expected behaviours of the real robot. When the given parameters have higher values than the required maximal values the application computes robot's motion parameters (Fig.3).

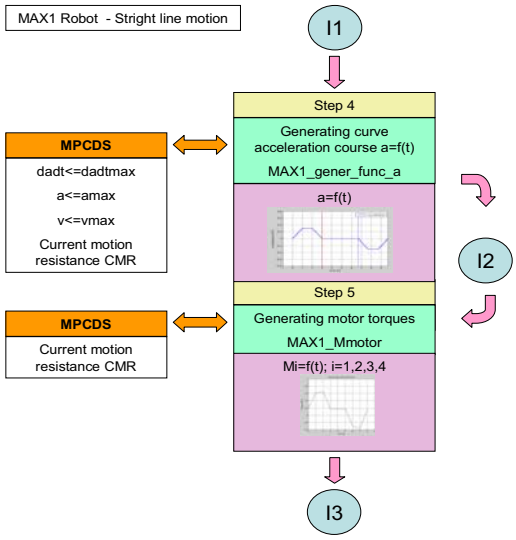


Fig. 3. Generating parameters of the motion.

After the preliminary task verification the next stage of the program is used for testing and performing highly dynamic behaviours of the mobile robot (Fig. 3). The predefined industrial definitions of the short-term motor states are used. The result of this stage is set of predicted motors' torques in the dynamic-kinematic robot's model (Fig. 4). In the subsequent step the parameters are sent to the real robot. This step may also, if it is necessary decide to interrupt the currently performed order and as a second activity ask the application MPCDS to change the task.

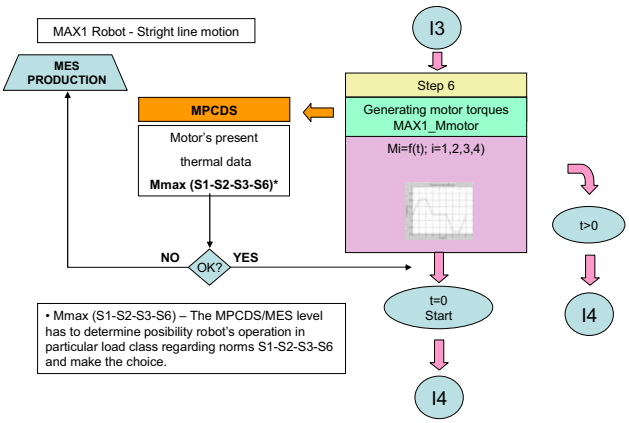


Fig. 4. Computing the final motor torques.

The next step (Fig.5) is used for modelling, analysing and testing robot's behaviours while the task is performed. The main part of the task consists of the correction of the robot's trajectory. It is done in the case of disturbance such as presence of other robots or human beings or slipping of the wheels. Parameters used in this step to define dynamic states tentatively taken from MPCDS are dynamically updated. This updating is done every 60s.

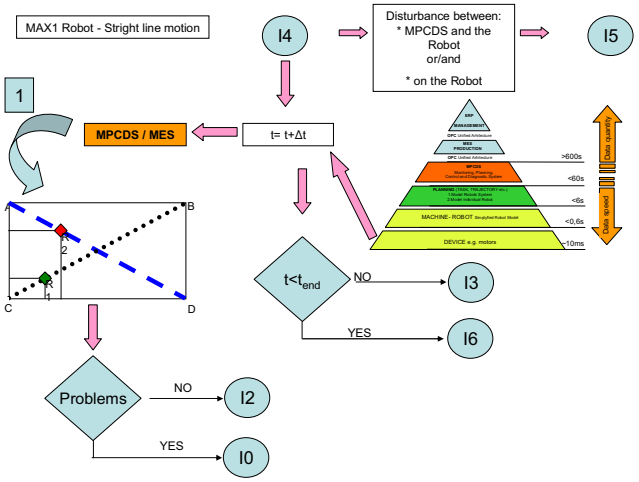


Fig 5. The modelling, analysis and test of robot's behaviours.

The next step (Fig.6) is used in order to ensure safety in the technological process. This steps offers additional protection in case of disruptions in the real robot's operation. The main emphasis is placed on the safety task division between the central intelligence (central dynamic-kinematic robot's model) and the local intelligence (the real drive unit).
In the last stage (Fig.7) the long-term verification of the assumed optimization process is carried out. The peculiarity of the proposed solution is visible in the holistic usages of safety software engineering by means of the MPCDS also integrating ERP systems and MES.

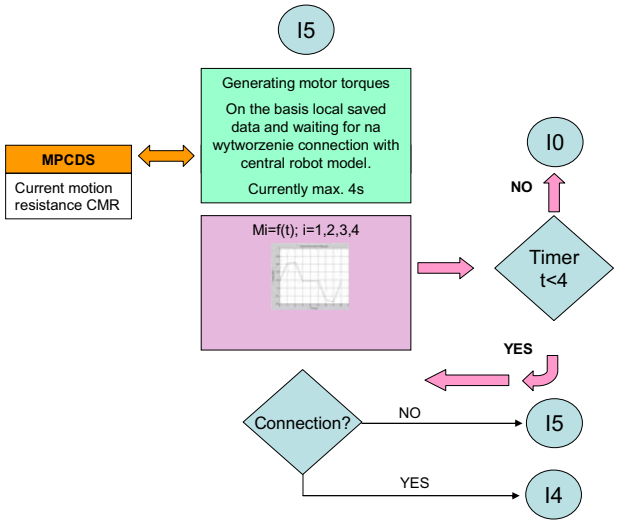


Fig 6. The safety ensuring of the technological process.

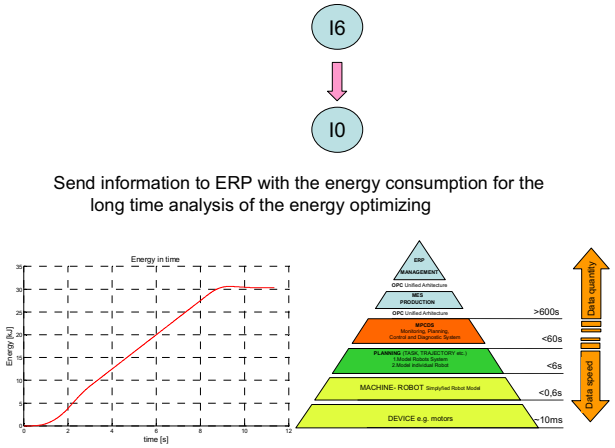


Fig.7. Long-term verifications.

4. SUMMARY

The core objective of the presented research activities is the realisation of extensions of the current possibilities concerning monitoring, planning, control and diagnosis of mobile robots in production systems as an example for complex industrial systems. The main result is a hierarchical concept and the exemplary realisation of a monitoring, planning, control and diagnosis system (MPCDS). This supervisory system is under development and requires further investigations in order to make an appropriate connection among particular applications as well as to ensure cooperation with other levels in the considered IT infrastructure.

REFERENCES

Borenstein, J., Everett, H. R., Feng, L., (1996). Where am I? Sensors and Methods for Mobile Robot Positioning. Ann Arbor, University of Michigan: 1996.

Freitas, N., Dearden, R., Hutter, F., Morales-Menendez, R., Mutch, J., Poole, D., (2004). Diagnosis by a Waiter and a Mars Explorer. Proceedings of the IEEE Special Issue on Sequential State Estimation.

Hachour, O., (2008).Path planning of Autonomous Mobile robot. International Journal of Systems Applications, Engineering & Development Issue 4, Volume 2.

Holland, J., (2004) Designing Autonomous Mobile Robots , Elsevier Inc.

ISO10218, (2006) "Robots for industrial environments - Safety requirements -Part 1: Robot,"

Kościelny, J.M., Leszczyński, M., Bartyś, M., (2006). Investigations of fault tolerant systems.

Lamon, P., Krebs, A., Lauria, Siegwart, M., Shooter, S., (2004). Wheel torque control for a rough terrain rover. IEEE International Conference on Robotics and Automation.

Lepetic, M., Klancar, G., Škrjanc, I., Matko, D., Potocnik, B. (2003). Time optimal path planning considering acceleration limits. Robotics and Autonomous Systems vol.45, p. 199–210.

Merzouki, R., et al., (2010). Hybrid fault diagnosis for telerobotics system. Mechatronics, 2010.

Pieczynski, A.,(2003). Knowledge representation in the diagnostic expert system.Lubuskie Scientific Society Zielona Góra, Poland (in Polish).

Podsedkowski, L., Nowakowski, J., Idzikowski, M., Vizvary, I., (2001). A new solution for path planning in partially known or unknown environment for nonholonomic mobile robots, Robotics and Autonomous Systems vol.34 p.145–152.

Siegwart R., Nourbakhsh, I. R., (2004). Introduction to Autonomous Mobile Robots, The MIT Press Cambridge, Massachusetts London, England.

Tian-Tian, Y., Zhi-Yuan, L., Hong, C., (2008). Formation Control and Obstacle Avoidance for Multiple Mobile Robots Acta automatica Sinca Vol.34, No.5

Washington, R., (2000). On-Board Real-Time State and Fault Identification for Rovers.Autonomy and Robotics Area NASA Ames Research Center, Proceedings of IEEE International Conference on Robotics and Automation (ICRA2000)

Yavuz, H., (2007). An integrated approach to the conceptual design and development of an intelligent autonomous mobile robot. Robotics and Autonomous Systems 55 (2007) 498–512

Zajac, M., Uciński, D., Stetter, R., (2008). Mobile robot Diagnosis with Bayesian Filters Mechatronic Systems and Materials - MSM 2008, 4th international conference. Bialystok, Poland.

Zając, M., Patan, K., (2009). Fault detection of the mobile robotusing dynamic neural networks, In :Fault detection, analysis and tolerating systems, Z. Kowalczuk (Ed.) Gdańsk, Pomeranian Science and Technology Publishers, PWNT (in Polish).

Ziemniak, P., Stania, M., Stetter, R., (2009a). Mechatronics engineering on the example of an innovative production vehicle, International conference on engineering design, ICED'09, 24 - 27 August 2009. Stanford University, Stanford, CA, USA.

Ziemniak, P., Paczynski, A., Uciński, D., Voos, H., (2009b). Motion Planning for Mobile Robots in Microproduction under Special Constraints. 7th Workshop on Advanced Control and Diagnosis, Zielona Góra, Poland.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Modelling of positive displacement pumps for monitoring, planning, control and diagnosis

Stefan Kleinmann*, M. Fairusz Abdul Jalal,
Ralf Stetter****

** Allweiler AG, 78315 Radolfzell, Germany (e-mail: S.Kleinmann@allweiler.de).*
***Hochschule Ravensburg-Weingarten, 88241 Weingarten, Germany (e-mail: stetter@hs-weingarten.de)*

Abstract: This paper describes the development of a model of three-spindle pumps. Such pumps are used in many application areas such as power generation. The model serves for the simulation of pump systems and will be applied in future steps in a simplified form in the control of pumps in order to allow advanced techniques of monitoring, planning, control and diagnosis. For this purpose a hierarchical concept for monitoring, control and diagnosis was developed in connected work (Kleinmann et al. 2010). The presented model is a core element of the future system. The results were found in a joint project of a leading pump system manufacturer and three universities in Germany, Poland and Switzerland.

Keyword: Monitoring, Planning, Control, Diagnosis, Pump Systems

1. INTRODUCTION

The functionality of the screw spindle pump did not change much since the beginning of its creation and it is widely used in industry. Several investigations in Europe have shown that most pump systems in industrial applications still do not dispose of any control system and that up to a third of the energy used could be saved by means of intelligent control systems. According to a study from the „Energy Information Administration“ from the year 2007 the global demand for electrical energy will be twice as high as today by the year 2030. About 4 % of the energy generated worldwide is used to transport fluids. In Germany electrical motors account for about 70% of the industrial electrical energy demand, about 30% of these are used to drive pumps and pump systems. However, still the efficiency plays a minor role in connection with pumps and pump systems (Friedl 2007).

Many pump systems have the potential of considerable energy conservation. Mainly systems are concerned with their ability to operate at several different work points (relation between volume/pressure and delivery volume). For this purpose certain work points must be given either by a machine control or by a leading vantage point or an adaptive control (i.e. a control which reacts independently on divergent operating conditions) has to be realized.

In particular, revolutionary changes require extensive control concepts. A special requirement in this case is the claim to develop a system without sensor, i.e. the available components - pump and electrical motor should operate as "sensors".

Today, a wide range of advanced fault detection and diagnosis systems are available, especially for pumps. A concept describing the sensible application of these systems is sought and the derivation of suitable feedback is aimed at which will lead to direct customer advantages in later stages. The researched system should be independent from the pump dimensions and, as far as possible from the pump design. The system should give information about how the pump “is”, i.e. how the different physical measuring dimensions must be gathered and be interpreted.

The screw spindle pumps are available as one spindle (eccentric pump), two spindle or multi spindle variants and their applications depend on the field or process purposes (Vetter 1987). In general, the screw spindle pump is widely used in diesel engine applications and burner industries as a transfer pump for transporting heavy and light oils, all kinds of lubricating fluids, waste oil, residual oil, grease, also little abrasive components or contaminates. Furthermore, the screw spindle pump is also extensively used in chemical and petrochemical industries as a transfer pump for all lubricating, non-lubricating fluids from low to high viscosity such as lube oil-, crude oil-, tar-, with contaminates, grease-, resin-, adhesive-and glycerin products.

From the design point of view, the multi spindle pump consists of two or more spindles which are enclosed by a housing casing. In the case of the three screw spindle pump, the rotating elements are only one active spindle and two idlers or passive spindles. The profile geometry of the active and passive spindles creates sealed and enclosed chambers. When they rotate, the driving spindle closely meshes with the passive spindles in the pump casing, which tightly surrounds the complete spindle set, creating series of cavities, trapping the liquid and moving it axially from suction to discharge side. Theoretically, this principle provides a continuous and pulsation-free flow without agitation of the fluid.

Figure 1 shows the cross sectional view of a three screw spindle pump and its important elements.

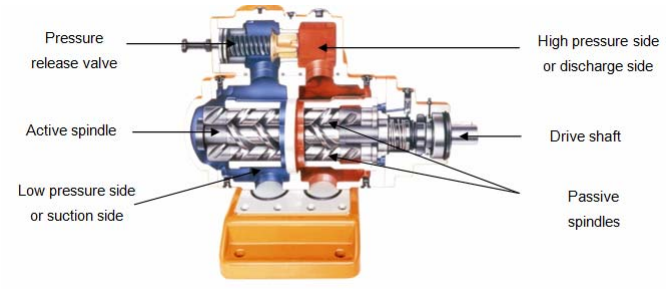


Fig. 1. Three screw spindles pump

2. STATE OF THE ART

The overview of the functionality of pump as well as its operating behavior is described by Faragallah & Surek (2004), Stiess (1966), Schulz (1977), Feindeisen & Findeisen (1994) and Krist (1991). Furthermore, Faragallah & Surek (2004) and Henkelmann & Sippel (1988) show the application of different types of pumps and also give examples of the empirical basic calculations of the three screw spindle pumps to determine flow rate, power consumption, efficiency and NPSH-values.

As stated in Wincek (1992) and Körner (1998), the mathematical model of the displacement pump and motor that have been developed by Schlösser (1961), describe the influence of media viscosity and density to the volumetric, mechanical-hydraulic and total efficiency of pump. The model is based on identification of the pump leakage loss. The leakage loss consideration will be divided into two cases. For the first case, the leakage loss is determined by the influence of media viscosity. For the second case, the leakage loss is considered to be dependent on the media density. The working point characteristic and the power consumption of the screw spindle pump are determined by introducing a loss factor.

However the result from the mathematical model by Schlösser (1961) is slightly different from the real behavior of the pump. One of the reason is that Schlösser considers that the pump behaves elastically and therefore the size of the leakage gap is calculated as a variable (not as a constant value) in the mathematical model of the pump. Other research work of Schlösser & Hilbrands (1963) is concentrated on determination of theoretical displaced volume, volumetric efficiency, mechanical-hydraulic efficiency and total efficiency of displacement pump.

Moreover, as stated in Körner (1998), Wilson (1950) develops a connected mathematical model for flow rate calculation of displacement pump and motor through observation of leakage loss with consideration of the dynamic viscosity of the media. Hammelberg illustrated in Wilson (1950) the characteristic of pressure distribution in the screw spindle pump, spindle- profile as well as the resulting spindle bending for the two screw spindle pump. The spindle profile which has been used is spur gear with helical thread. The common screw profiles are involute or cycloid screw shape. Other than that, Hammelberg formulated a procedure to calculate power consumption of pumps (Hammelberg 1968).

Wincek (1992), Körner (1998), Schmidt (1999), Rausch (2006) and Etzold (1993) described the mathematical model of two screws spindle pumps for multi phase application. For the first step, Wincek (1992) and Körner (1998) formulated the mathematical model by considering a one phase liquid and then extending the mathematical model to multi phase application. Furthermore, Wincek (1992), Körner (1998), Sabine Schmidt (1999), Rausch (2006) and Etzold (1993) illustrated the total leakage losses as the sum of losses through radial gap, circumferential gap and spindle rotation. The leakage losses model considers the flow from chamber to chamber under consideration of the back flow behavior.

The three screw spindle pump is not extensively being researched and the mathematical equations that describe the three screw spindle pump are merely a black box consideration without any details like back flow consideration from chamber to chamber as in two screw spindle pump model by Wincek (1992), Körner (1998), Schmidt (1999), Rausch (2006) and Etzold (1993).

3. DEVELOPMENT OF THE MODEL

The calculation of flow using a detailed cross section area between the spindles is briefly described by Faragallah & Surek (2004). The cross sectional areas are calculated by considering rectangular and cycloid screw thread forms. The mathematical model for three screw spindle pump is derived accordingly. The mathematical equation based on cycloid cross sectional areas will be described in the following subchapter.

3.1 Screw spindle pump with cycloid screw thread shape consideration

The cross section of the screw spindle pump with cycloid screw thread shape consideration is illustrated in Figure 2.

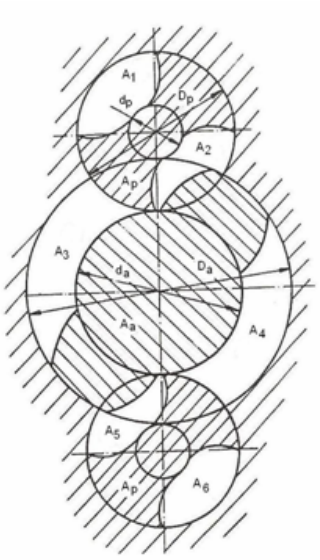


Fig. 2. Cross sectional area of the three screw spindle pump with cycloid screw shape consideration

The theoretical volume V_{th} can be determined through the free area A_c and the spindle pitch h_s:

$$V_{th} = A_c.h_s = \left(A_1 + A_2 + A_3 + A_4 + A_5 + A_6\right)h_s \quad (1)$$

The pitch h_s describes the displacement of the thread by one full spindle rotation. The relationship between pitch angle α and inner diameter of the active spindle d_a can be expressed as (Wincek 1992), (Schmidt 1999):

$$tan\alpha = \frac{h_s}{d_a.\pi} \quad (2)$$

The total cross sectional area of the screw thread A_c, is determined either graphically or mathematically. The further assumptions of cycloid screw shape design are according to Faragallah & Surek (2004). The outer diameter of passive spindle D_p is equal to the inner diameter of the active spindle d_a:

$$D_p = d_a \quad (3)$$

The relationship between the outer diameter of the active spindle D_a, the inner diameter of the active spindle d_a and the inner diameter of the passive spindle d_p, which has been described in Faragallah & Surek (2004), are:

$$\frac{D_a}{D_p} = \frac{D_a}{d_a} = \frac{5}{3} \quad (4)$$

$$\frac{D_p}{d_p} = 3 \quad (5)$$

The relationships in the equations (4) and (5) have also been confirmed to be the same which are currently used by the design teams of Allweiler AG for describing the screw spindle pump design parameters.

The cross sectional areas of the active spindle A_a, the passive spindle A_p and the pump housing A_h according to Faragallah & Surek (2004) are:

$$A_a = 1.26787D_p^{\,2} \quad (6)$$

$$A_p = 0.42832D_p^{\,2} \quad (7)$$

$$A_h = 3.36757D_p^{\,2} \quad (8)$$

Therefore the calculated free area A_c between the spindles and the pump housing is:

$$A_c = A_h - \left(A_a + 2A_p\right) = 1,24307D_p^{\,2} \quad (9)$$

Hence the theoretical volumetric flow V_{th} of three spindle screw pump with cycloid thread shape is:

$$V_{th} = 1.24307D_p^{\,2}.h_s = 4.14357D_p^{\,3} \quad (10)$$

And the rate of flow Q_{th} can be described as:

$$Q_{th} = V_{th}.n_2 \quad (11)$$

As already stated in equation (2), the screw pitch h_s can be reformulated in terms of the pitch angle α as:

$$h_s = d_a.\pi.tan\alpha = \frac{3}{5}.D_a.\pi.tan\alpha \quad (12)$$

For simplification, the relation between profile factors K_p due to different number of spindles, active spindle diameter D_a and screw pitch h_s in term of the pitch angle α, the equations of volumetric flow V_{th} and the rate of flow Q_{th} can be generalized as below:

$$V_{th} = \frac{D_a^{\,2}.h_s.K_p}{4.10^6} = \frac{(1.5).D_a^{\,3}.\pi.tan\alpha.K_p}{10^7} \quad (13)$$

$$Q_{th} = V_{th}.n_2 = \left(\frac{(1.5).D_a^{\,3}.\pi.tan\alpha.K_p}{10^7}\right).n_2 \quad (14)$$

The profile factor K_p can be found in Table 1.

Table 1. Profile factor K_p for different number of spindles (Faragallah & Surek 2004)

	Profile factor K_p
2 spindle	2.16
3 spindle	1.80
4 spindle	3.60
5 spindle	5.66

In the company Allweiler AG extensive studies with existing pumps have resulted in slightly different formulae using a new corrected profile factor K_p:

$$Q_{th} = \left(\frac{3.D_a^{\,3}.\pi.tan\alpha.K_p}{2.10^7}\right).n_2 \quad (15)$$

3.2 Leakage loss

The leakage loss, which has been described by Faragallah & Surek (2004) illustrates the gap leakage loss inside the screw pump. The gap leakage loss increases proportionally by the

increase of the differential pressure Δp, the decrease of viscosity vis_2 and the pitch h_s. The leakage loss Q_v of three spindle screw pump depends on the differential pressure Δp and the media viscosity vis_2:

$$Q_v = \sqrt{\frac{\Delta p}{20}} . \frac{D_a^{\ a}}{c} K . h_s^{\ b} \sqrt{\frac{20}{vis_2}} \tag{16}$$

By substitution of the screw pitch h_s in terms of pitch angle α (see equation (12)) in equation (16), the corresponding leakage loss Q_v can be reformulated as

$$Q_v = \sqrt{\frac{\Delta p}{20}} . \frac{D_a^{\ (a+1)}}{c} . \frac{3}{5} . K . \pi . tan\alpha^{b} . \sqrt{\frac{20}{vis_2}} \tag{17}$$

Whereas, the parameter values for the calculation of the leakage loss Q_v are simplified in Table 2.

Table 2. Parameter values for leakage loss calculation

	Value	
K	0.025-0.05	
D_a<75mm	a=0.5	c=1
D_a≥75mm	a=1.22	c=1.5
vis_2 <20mm²/s	b=4	
vis_2 ≥20mm²/s	b=2	

In the company Allweiler AG extensive studies with existing pumps have resulted in much lower (better) results for the leakage loss. Here empirical formulae were generated which are also used in the developed model. The structure and the input variables of these calculations are similar to the calculation possibilities listed above.

3.3 Power loss

The power loss P_r, which is described by Faragallah & Surek (2004), is divided into two cases depending on geometry of active spindle D_a.

For the first case, which is D_a < 75 mm, the power loss P_r is calculated as:

$$P_r = \frac{Q_{th}}{600} \left(0.062 + \frac{3}{D_a} + \frac{\Delta p}{1000} \right) \left(\frac{vis_2^{\ \frac{1}{2.5}} . n^{0.5}}{7.75} \right) \tag{18}$$

For the second case, which is D_a ≥ 75 mm, the power loss P_r can be described as

$$P_r = \frac{Q_{th}}{600} \left(0.008 . vis_2^{\ \frac{1}{2.8}} . n^{\frac{1}{1.75}} \right) \tag{19}$$

For both cases, if vis_2 ≤12 cSt, it will be rounded up to media viscosity vis_2 equal to 12 cSt.

In the company Allweiler AG extensive studies with existing pumps have resulted in slightly worse results for the power loss. Here empirical formulae for the power loss were generated which are also used in the developed model. The structure and the input variables of these calculations are similar to the calculation possibilities listed above.

3.4 Pump torque load

The behavior of the screw spindle pump varies depending on the moment of inertia of the rotating parts of the pump and electrical motor, friction and also system pressure which is connected to the pump (Faragallah & Surek 2004), (Wincek 1992). However, the moment of inertia of the rotating parts in the pump is smaller in comparison with the rotating parts in electrical motor. Therefore, the moment of inertia of the rotating parts in the pump is considered as negligible. The resistance from the media viscosity is also considered as negligible.

The torque at operating point of the pump M_{load} is the result of motor speed n_2 and the actual power consumption P_{act}, which is described in specific tables depending on media viscosity, pressure difference and Q_{act}. . The corresponding moment at operating point of the pump M_{load} is:

$$M_{load} = \frac{P_{act}}{\omega_2} \tag{20}$$

Furthermore, the resulting moment for the acceleration M_{acc} is considered as:

$$M_{acc} = M_{motor} - M_{load} \tag{21}$$

3.5 Efficiency

The efficiency of the pump η_{pump} is the product of volumetric efficiency η_{vol} with regards to flow rate and mechanical efficiency η_{pump} with regard to power.

The volumetric efficiency η_{vol} can be calculated according to Faragallah & Surek (2004):

$$\eta_{vol} = \frac{Q_{act}}{Q_{th}} \tag{22}$$

The mechanical efficiency η_{mech} can be calculated according to Faragallah & Surek (2004):

$$\eta_{mech} = \frac{P_{th}}{P_{act}} \quad (23)$$

The total pump efficiency η_{pump} can be calculated according to Faragallah & Surek (2004):

$$\eta_{pump} = \eta_{vol} . \eta_{mech} \quad (24).$$

4. SPF PUMP MODEL IN MATLAB SIMULINK

MATLAB is an interactive software tool for modeling, simulating and analyzing dynamic systems, which is showing optimal results in all fields of engineering. MATLAB also includes the Simulink graphical environment used for multi-domain simulation and model-based design. Simulink can be used for simple mathematical manipulations by means of already built in and even self built mathematical libraries or toolboxes.

Since the beginning of the modeling project, it has been decided that the system block will be built mostly in Simulink with no external programming code in MATLAB in the form of MAT files (*.m). The overall system model consists of motor and converter system, pump system and consumer system blocks. Each of the model blocks will be modeled separately and later on the model blocks will be assembled. The overall system model with input and output parameters of each of the blocks is illustrated in Figure 3.

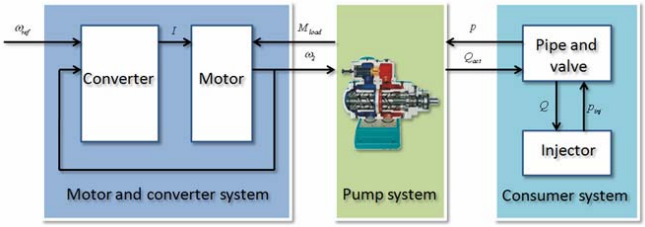


Fig. 3. Overall system

The overall system blocks are considered as parts of a modular model library which is advantageous for further model development and simulation purposes. Since the model blocks are being developed separately, they are upgradeable at any time during the model development process. The model alteration will not affect other system blocks as long as the inputs and outputs from each block are compatible with each other in terms of units and simulation parameter settings.

Furthermore, a user can create the desired system by drag-and-drop operations, with no programming necessary. The user just needs to know how to connect the input and output of each block and can run the simulation afterwards. However, this modular library model also has a disadvantage. There is no consideration of reciprocal influence between the components.

The overall system consists of three distinguished model blocks, which are motor and converter system model (orange block), pump system model (blue block) and consumer system model (green block) as in Figure 4.

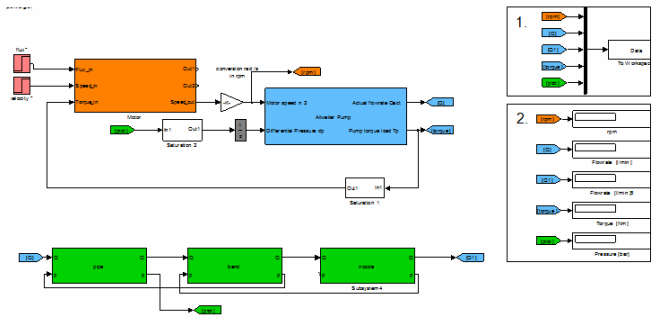


Fig. 4. Overall system in MATLAB Simulink

The consumer system library consists of simple pipe line, pipe bend, valve, Y-bend and nozzle. The user can drag and drop the self built toolboxes for the consumer system to visualise the model as in the workbench for the simulation purposes.

The motor speed n_2 from the motor and converter system block as well as differential pressure Δp from the consumer system block will be fed to the pump system block as an input. The actual flow rate Q_{act} from the pump system block will be connected as an input to the consumer system block and the pump torque load M_{load} will be an input to the motor and converter block.

The motor speed n_2, actual flow rate Q_{act}, pump torque load M_{load} and differential pressure Δp data are transferred to Workspace for further analysis when needed (see 1. in Figure 4). The result can also be seen directly from the display (see 2. in Figure 4). Figure 5 shows the pump model block with the explanation of its input and output parameters.

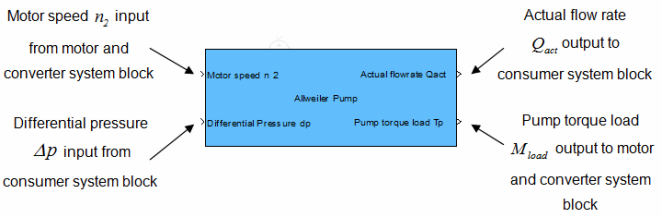


Fig. 5. Pump model

A user can choose the desired SPF pump by double clicking the pump model block. A pop-up window will then appear. The user may choose pump size with different pitch angle before running the simulation. To avoid confusion in pitch angle selection, the other unrelated pitch angle option will be disabled. For instance, as the user chooses pump size 10, the other pitch angle selection for SPF 20 and SPF 40 will be disabled. A user can also change the media viscosity vis_2 value in the pop-up window before executing the simulation.

The pump model is based on the considerations shown in section 3; the realization is not described in detail in this paper.

5. VALIDATION AND VERIFICATION

In order to validate the model test field data of pumps made by Allweiler were evaluated and compared with the calculated result using the developed model. The test data is the result from the test bench by using the same pumps with different speeds. There are inconsistencies in the relationship between calculated results and test field data at very low as well as at very high differential pressure Δp. The relationship between calculated results and test field data by excluding the inconsistencies at very low as well as at very high differential pressure Δp are illustrated in Figure 6.

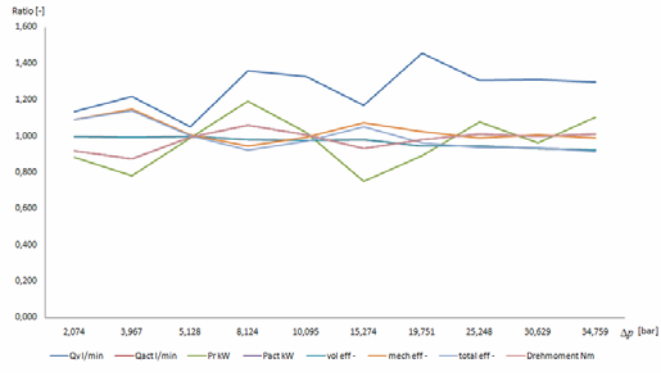


Fig. 6. Relationship between results of the model and test field data at a motor speed of 1450 [1/min]

By excluding the inconsistency area, the relationship between the ratios of actual flow rate Q_{act}, power consumption P_{act}, volumetric- η_{vol}, mechanical- η_{mech}, pump efficiency η_{pump} and pump load torque M_{load} is relatively close to 1, which means that the calculated result is comparatively same as the test field result. The leakage loss Q_v and power loss P_r have higher ratios compared to other parameters. The leakage loss Q_v has the ratio relatively consistent between 1.3 and 1.5. The calculated result of leakage loss Q_v using the model is slightly bigger than the test field data. The friction power loss P_r has a ratio, which fluctuates close to 1. However, the leakage loss Q_v is very small in comparison to the theoretical flow rate Q_{th}. Therefore, the volumetric efficiency η_{vol} is not very much affected by the big deviation of the calculated result and test field data. The friction power loss P_r influences the mechanical efficiency η_{mech} heavily in the lower differential pressure Δp range when the motor speed is constant at certain value. Therefore, the only range where frictional power loss P_r ratio effects the efficiency at a low differential pressure Δp which has little importance in real applications.

6. SUMMARY

In conclusion, the results from the test field shows that the pumps that have been tested in the test field behave accordingly to the developed model with a certain degree of tolerance, which has been confirmed as acceptable by the involved experts. Further work will concern the dynamic behavior and the integration in the IT-structure.

AKNOWLEGEMENTS

The project is funded by the European Union in the scope of the European Fond for regional development in the Interreg IV program “Alpenrhein-Bodensee-Hochrhein“ together with the “Schweizer Bund” and the “Fürstentum Lichtenstein” as well as the “Internationale Bodensee-Hochschule”.

REFERENCES

Etzold, S.: Verlustanalyse von Schraubenspindelpumpen bei Mehrphasenförderung. Hannover : VDI Verlag, 1993.

Faragallah, W. H. and Surek, D.: Rotierende Verdränger-maschinen. Sulzbach : Verlag und Bildarchiv, 2004.

Findeisen, D. and Findeisen, F.: Ölhydraulik. Berlin: Springer Verlag, 1994.

Friedl, C.: Energie-Optimierung von Pumpen rechnet sich. In: Maschinenmarkt. November 2007.

Hammelberg, F. W.: Untersuchungen an Pumpen-Läuferprofile, Läuferkräfte und Leistungen von Schraubenpumpen. s.l. : VDI-Forschungsheft 527, 1968.

Hammelberg, F. W.: Die Läuferkräfte bei Schraubenpumpen. TH Hannover: s.n., 1966.

Henkelmann, N. and Sippel, F.: Schraubenspindlepumpen-Auswahl für spezielle Anwendungsfälle, Erfahrungen aus Praxis. Karlsruhe : Pumpentagung Karlsruhe, 1988.

Körner, H.: Zum Förderverhalten von Schrauben-spindelpumpen für zweiphasengemische hohen Gasgehalts, Dr.-Arb., Uni Erlangen-N., 1998.

Kleinmann, S., Koscielny, J.M, Koller-Hodac, A., Paczynski, A., Stetter, R.: Concept of an advanced monitoring, control and diagnosis system for positive displacement pumps. Proceedings of SysTol 2010, Nice.

Krist, T.: Hydraulik Fluidtechnik. Darmstadt : Vogel Fachbuch, 1991.

Rausch, T.: Thermofluiddynamik zweiphasiger Strömungrn in Schraubenspindelpumpen. Hannover : Cuvillier Verlag Göttingen, 2006.

Schulz, H.: Die Pumpen-Arbeitweise, Berechnung, Konstruction. Berlin : Springer-Verlag, 1977.

Schlösser, W. M. J.: Das theoretische Hubvolumen von Verdrängerpumpen Ölhydraulik und Pneumatik 7. 1961.

Schlösser, W. M. J. and Hilbrands, J. W.: Das theoretische Hubvolumen von Verdrängerpumpen Ölhydraulik und Pneumatic. 1963.

Schmidt, S.: Verschleiß von Schraubenspindelpumpen beim Betrieb mit abrasiven Fluid, Dr.-Arb., Universität Erlangen-Nürnberg, 1999.

Stiess, W.: Pumpen-Atlas Teil 1. Ludwigsburg : A.G.T.-Verlag Georg Thum, 1966.

Vetter, G.: Einführung und Einblick zur Pumpentechnik Jahrbuch Pumpen. 1987.

Wilson, W. E.: Positive displacement pumps and fluid motors. New York : Pitman Publishing Corp., 1950.

Wincek, M.: Zur Berechnung des Förderverhaltens von Schraubenspindelpumpe bei der Förderung von Flüssigkeiten / Gas-Gemischen, Uni Erlangen-N., 1992.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Concept of an advanced monitoring, planning, control and diagnosis system for autonomous vehicles

Lothar Seybold*, Andrzej Pieczyński**
Andreas Paczynski*, Ralf Stetter*****

** RAFI GmbH & Co. KG, 88276 Berg, Germany (e-mail: lothar.seybold@rafi.de).*
*** Uniwersytet Zielonogórski, 65-417 Zielona Góra, Poland (A.Pieczynski@issi.uz.zgora.pl)*
****Hochschule Ravensburg-Weingarten, 88241 Weingarten, Germany (e-mail: stetter@hs-weingarten.de)*

Abstract: This paper describes an innovative concept for an advanced monitoring, control and diagnosis system for autonomous vehicles. Such vehicles are used in production environments and in large public institutions such as hospitals. The potential for such vehicles is huge but the growth rates of this market have not been as tremendous as it was expected before. One major cause for this rather small application ratio may be the complicated control systems. Findings of the authors in other application areas can be summarized to the insight that only an integrated control and diagnosis will be able to overcome the opposition towards complicated systems in industry (Kleinmann et al. 2009). Such integrated control and diagnosis systems can assure to operate complicated systems at the best efficiency point (operating on demand), to prevent failures and break downs (fault protection) and to indicate maintaining actions required from the users (maintenance on demand). Furthermore, in other industries today usually the operation data of all systems are being monitored for several reasons, e. g. for better planning of the company's resources. Such monitoring of autonomous vehicles is currently also only realized up to a certain degree but could be greatly expanded and could be an additional function of a control and diagnosis systems. Furthermore the planning of future actions, trajectories and driving behavior of autonomous vehicles requires this monitoring information, has to take into account the control activities and is mutually dependent on diagnosis activities. It is therefore a logical step to include planning in a holistic concept for autonomous vehicles. The application of a system for monitoring, planning, control and diagnosis systems means a shift of paradigm and has therefore be consciously planned and be based on a well-considered concept. A concept for such systems is proposed in this paper. The considerations and proposals are based on a collaboration of two Universities and a world-leading production company for electrical components for production systems and vehicles and eMobility.

Keyword: Monitoring, Planning, Control, Diagnosis

1. INTRODUCTION

Autonomous vehicles are currently used in many industry branches. They are frequently used for in-company logistics within the warehouse or production facilities. The first concepts go back to the early Seventies of the last century. In these and the following times a tremendous growth potential was assigned to the application of such vehicles. However, after considerable market growth a phase of disillusionment could be observed. The produced and installed systems were not as flexible, efficient, reliable and safe as expected. All the aspects were gradually improved and currently a rising demand for autonomous vehicles can be observed. Future production scenarios rely strongly on these vehicles not only for the transport of goods but also of production machinery. The vision is a production hall where any production step could be flexibly realised at any position in this production hall. In the core of this production vision are extremely flexible autonomous vehicles which dispose of multiple manoeuvring possibilities and the capacity to change the level and inclination of a transportation platform. An example of such vehicles is shown in Figure 1.

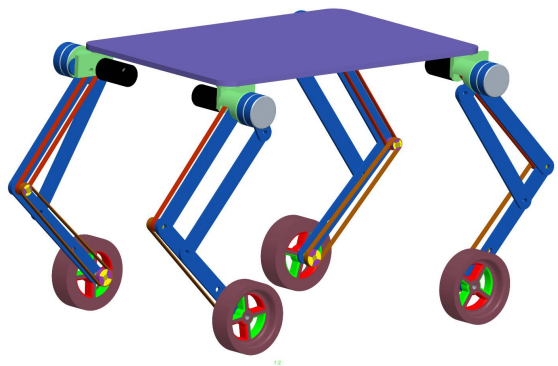


Fig. 1. Autonomous vehicle (example – CAD data)

One major building block for the further success of such autonomous vehicles is the information technology in general. Four main aspects need to be considered:

monitoring, planning, control and diagnosis. The terms planning, control and diagnosis are widely known and will only be shortly explained in the following part. An important extension to diagnosis – predictive diagnosis –and the rather unknown notion "monitoring" will be elucidated in more detail in section two and three.

Planning is essentially the process of predefining activities in the future. In the area of autonomous vehicles planning can apply to the assignments of tasks to an autonomous vehicle (e. g. "transport part x from station y to station z"), the planning of paths and trajectories (e. g. "drive along a certain line") and the planning of the driving behaviour (e. g. "drive with velocity x and acceleration y and maximum acceleration change z"). Obviously these planning activities concern different levels of detail and the required performance in terms of data amount and data speed differs greatly. This fact makes a holistic approach desirable.

The term "**control**" names activities which serve to manage, command, direct or regulate the behaviour of devices or systems and has been the core of extensive research for many decades. In recent years the techniques of predictive control have found rising attention (compare e.g. Camacho&Bordons 2004, Wang&Boyd 2008). Predictive control usually relies on dynamic models of the process, most often linear empirical models obtained by system identification. In the area of autonomous vehicles predictive control can pursue three different objectives:

- smoothing changes of system states,
- better coordination of multiple autonomous vehicles and
- evaluating decision alternatives.

Over the last three decades, the growing demand for safety, reliability, and maintainability in technical systems has drawn significant research in the field of **diagnosis**. Such efforts have led to the development of many techniques; see for example the most recent survey works (Blanke et al. 2006, Isermann 2005, Witczak 2007, Zhang and Jiang 2008, Korbicz et al. 2004). The application of a collection of these techniques gathered in one system (DIASTER - Koscielny et al. 2006) was analyzed by Dabrowska und Kleinmann (2009). For fault compensation in general fault tolerant control methods are proposed which can be classified into two types, i.e. Passive Fault Tolerant Control Scheme (PFTCS) and Active Fault Tolerant Control Scheme (AFTCS) (Blanke et al. 2006, Zhang and Jiang 2008).

2. PREDICTIVE DIAGNOSIS

The term "predictive diagnosis" is in contrast to "predictive control" rarely used in the technical domain (it is widely used e. g. in medicine). One example for the usage of this term is the presentation of an automated system for fault diagnosis based on vibration data recorded from an main power transmission (Diwakar 1998). For autonomous vehicles predictive diagnosis (essentially in the meaning of failure detection and identification before these failures even occur) presents a promising field of research. Five main problems in the operation of autonomous vehicles bear the possibility to be identified early:

- reduced pressure in air tires leadings to increase power demand for similar operations and later to destruction of the tires
- wear of bearings and gear systems leading firstly to increased power demand for similar operations or increased vibrations and finally to system failure;
- wear of electrical motors (e. g. brushes) leading firstly to increased power demand for similar operations or increased vibrations and finally to system failure;
- wear or staining of sensors leading to imprecise or contradicting sensor readings;
- wet and slippery floors leading to imprecise movement and to danger for goods, machinery and persons.

Autonomous vehicles are usually part of larger systems. A failure of the larger system which is caused by a failure of a vehicle usually leads to enormous consequences in terms of cost of idleness (e. g. of a whole production segment). Therefore preventive maintenance is desirable for industrial autonomous vehicles; however such preventive maintenance today is aggravated by the fact the upcoming failures can usually not be detected. The only preventive maintenance systems possible are time based but not state based. A predictive diagnosis system would allow scheduling maintenance and service in dependence of the current state of a vehicle (wear of bearings, gear systems and electrical motors) and the state of the sensor (wear and staining). The authors believe that predictive diagnosis will only by possible as an extension of an elaborate advanced diagnosis system and include this extremely important aspect in the following considerations.

3. MONITORING

The notion monitoring summarizes all kinds of systematic observation, surveillance or recording of an activity or a process by any technical means. In the area of autonomous vehicles monitoring could be understood as a systematic collection of data concerning the state of certain physical characteristics such as distance, speed, acceleration, temperature, vibrations, torque, currents, voltage, power consumption, current gradient, velocity gradient, etc. In leading industries such as computer chip production or car manufacturing today usually nearly all operation data of the productions systems are being monitored for the three main reasons safety, efficiency and planability:

- The safety of production systems can be enhanced because a reliable safety system with a fast reaction can be realized on the basis of a real-time monitoring system. The role of coincidence for detecting possibly dangerous faults is diminished if a continuous monitoring is in place.
- The efficiency of production systems can be enhanced because any kind of waste (of energy, time and production goods) will be detected and can subsequently be prevented or reduced.

- The planning possibilities and planning quality can be enhanced if accurate data from a real-time continuous monitoring system are available as realistic prognosis is enabled by such data.

Such monitoring of autonomous vehicles is currently only realized to a limited degree but could be an additional function of a control and diagnosis system.

Discussion with leading customers of autonomous vehicles made clear that control and diagnosis systems will only be adapted in future if the are an integral part of the production system information infrastructure, namely the Enterprise Resource Planning (ERP) system and Manufacturing Execution System (MES). Enterprise resource planning (ERP) is an integrated computer-based system used to manage internal and external resources including tangible assets, financial resources, materials, and human resources. It is a software architecture which purpose is to facilitate the flow of information between all business functions inside the boundaries of the organization and manage the connections to outside stakeholders. Built on a centralized database and normally utilizing a common computing platform, ERP systems consolidate all business operations into a uniform and enterprise wide system environment (Bidgoli 2004). In all kinds of companies ERP systems usually present the top level of control within production. On the next level below are Manufacturing Execution Systems (MES). Boldly speaking, an ERP system defines what is to be produced within a given time period and the execution level (MES) takes this planning output and executes this plan on a near real-time/on-line basis (McCellan 1997). This control loop from ERP over MES to the real production operations is usually not closed today. An advanced monitoring would contribute to offer the major advantages usually connected with closed loop control (Ward 2007):

- disturbance compensation,
- guaranteed performance even with model uncertainties, when the model structure does not match perfectly the real process and the model parameters are not exact,
- stabilisation possibilities for unstable processes,
- reduced sensitivity to parameter variations and
- improved reference tracking performance.

In the following section a hierarchical concept is presented which is intended to combine the functionalities of monitoring, planning, control (including predictive control), diagnosis (including predictive diagnosis) into a sensible system structure for a holistic operation of industrial systems including autonomous vehicles.

4. DISTRIBUTED AND HIERARCHICAL CONCEPT

The insights from numerous discussions with representatives of leading companies indicate that holistic monitoring, planning, control and diagnosis systems for autonomous vehicles in industrial applications need to be integrated in the information system infrastructure of the respective industrial company. In these applications like in most other industry a hierarchical system can be observed. Figure 2 shows a proposal of a sensible hierarchy of the levels of these information systems in form of a pyramid.

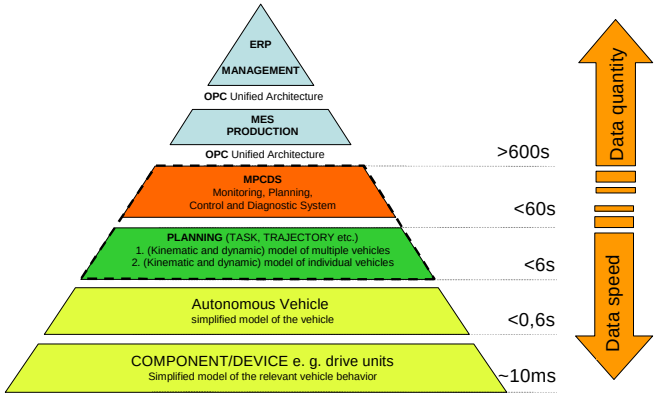


Fig. 2. Distributed and hierarchical concept

On the highest level the Enterprise resource planning (ERP) system can be found. It is not present in all kind of enterprises and is sometimes referred to with different names. The main function is always the same: this level concerns the planning of the entities to be produced on a not time-critical level. On the next lower level the production of these entities is executed by a plant control system which fulfils similar tasks than a Manufacturing Execution System (MES) Sometimes the two highest levels are realized in only one system.

On the next level is the first core of the proposed concept – the monitoring, planning, control and diagnosis system (MPCDS) for a section of a system which usually is including a number of autonomous vehicles. The MPCDS level with the planning level is described in detail in the respective subsection. The next lower level is the single autonomous vehicle. The lowest level contains components of a vehicle with an own local intelligence. Such a concept of distributed intelligence (compare Seybold et al. 2009) is very frequent in the upcoming age of ubiquitous computing and offers many advantages such as flexibility and reliability. On this level the real-time control has to take place and the most important safety functions should be realized on this level for the sake of a quick reaction. However, a number of aspects have to be considered for a sensible structure on this level. These aspects are discussed in the respective subsection.

It is one central hypothesis of this paper and a decisive building block of the presented concept that certain sets of data and procedures have to be available on all levels (in different amount and granularity). Four distinct sets of data and procedures should be present already on this level:

- Information concerning the configuration of the respective entity (meta-data) such as dimensions, typical efficiency or typical vibrations. These data can be static (not changing during the life-time of the entity) or dynamic.
- The current sensor readings (e. g. voltage, current and velocity) and a selected history of sensor readings allowing the application of certain modelling techniques or certain filters.

- The objectives, tasks and orders, i. e. the information what output the entity should achieve.
- The possibility to simulate the behaviour of the entity for several reasons concerning advanced monitoring, control and diagnosis by means of any mathematical model. This model can be accompanied by PID-controllers for certain functions and rules and procedures.

The four most important sets of data and procedures are summarised in Figure 3.

Sensor Data Current and past sensor readings	**Objectives, Tasks, orders** Local and global set
Meta Data, Configuration Robot settings	**Dynamic, Kinematic Model** Procedures, rules

Fig. 3. Four important sets of data and procedures.

4.1 Component/device level

The lowest level of the distributed and hierarchical concept is the component/device level. In this instance the focus are components and devices with own intelligence (processing unit with memory). In the age of ubiquitous computing with decreasing costs for intelligence the share of such devices and components is increasing for the sake of flexibility and reliability. Usually on the level just a limited amount of sensor information is available. However, usually in each component or device actuators are presents which could also be used as sensors. Concepts taking into consideration the characteristics for instance of motors (current, voltage, time constants) may be a future answer to many challenge in monitoring, diagnosis and control as sensor information is a necessary basis for such endeavours. However, in order to integrate this approach of virtual sensors (compare e. g. Koscielny et al. 2006) in a monitoring, planning, control and diagnosis system a certain structure should be realized. This structure is sketched in Figure 4.

Ubiquitous computing is one of the current megatrends (compare e. g. Greenfield (2006)). It is therefore no speculation to assume that it will be very easy, very appropriate and very cheap to equip components and devices in the near future with a decentral intelligence. Only relatively small improvements of efficiency or reliability will compensate these additional costs. However, in order to be able to cooperate in an integrated monitoring, planning control and diagnosis system also this decentral intelligence should dispose of a certain structure (shown in Figure 4).

The four distinct sets of data and procedures which were mentioned above should be present already on this level:

- Information concerning the configuration of the respective component or device (meta-data) such as dimensions.
- The current sensor readings (e. g. voltage, current) and optionally a small history of sensor readings allowing the application of certain modelling techniques or certain filters.
- The objectives, tasks and orders, i. e. the information what output the component or device should achieve.
- The possibility to simulate the behaviour of the component or device on a very basic level such as simple linear mathematical relationships mainly for the reason of advanced control. This model is usually accompanied by PID-controllers for certain functions and rules and procedures.

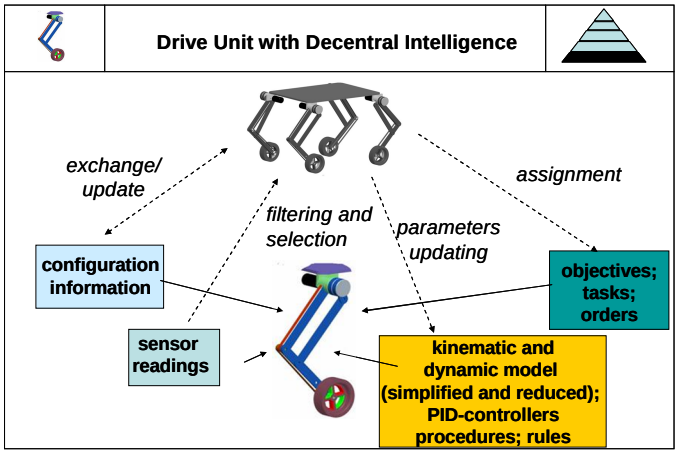


Fig. 4. Lowest level of the concept: component/device.

All these information will be communicated to the next higher level, to the autonomous vehicle. The configuration information are exchanged and updated in order to provide current meta-data on both levels. The sensor readings of the component or device are reduced to important and compressed sensor readings and are transferred to the vehicle in order to allow elaborate calculations and simulations on this level and in order to allow monitoring and closed-loop control. The vehicle assigns objectives, tasks and orders to the component or device and may also assign weights or priorities. Furthermore the vehicle may update parameters (or even models and procedures) concerning the calculation possibilities of the component or device.

4.2 Vehicle level

The vehicle level refers to the decentral intelligence located on the individual autonomous vehicle. Usually a small industrial PC will be the hardware basis for this level. The sensible structure on this level is very similar to the lower level “component/device” and is therefore not explained in detail in this publication. The second lowest level of the distributed and hierarchical concept is shown in Figure 5.

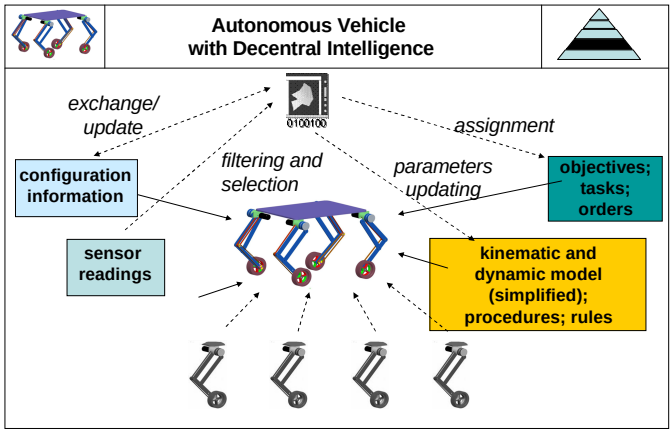


Fig. 5. Vehicle level.

4.3 MPCDS level

The core element of the concept is the monitoring, planning, control and diagnosis system (MPCDS) which is responsible for a certain sub-section of the production system usually containing a number of autonomous vehicles. The examples for the four partial objectives of this system are shown in Figure 6. The structure for this level is sketched in Figure 7.

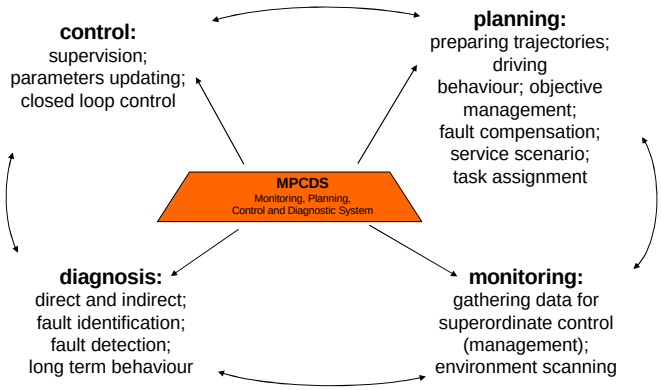


Fig. 6. Partial objectives of the MPCDS.

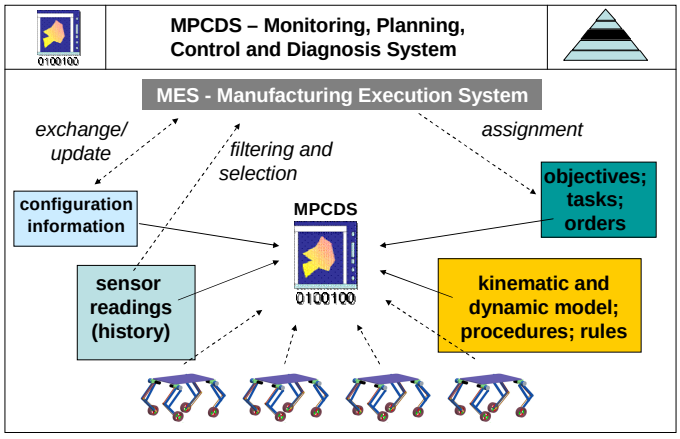


Fig. 7. MPCDS level.

This system is communicating with all autonomous vehicles using four different "channels" as described in the prior section. The structure is very similar also containing configuration information, sensor readings, objectives/tasks/orders and the dynamic model. However, on this level all the information are more condensed and the models are more elaborate. This is caused on the one hand by the fact that more computing power is available on this level and that the operations, calculations and simulations carried out are less time-critical.

All the information on this level will be communicated to the next higher level, to the Manufacturing Execution System MES. Again, the configuration information of all vehicles in this subsection of the production and about their surroundings and connections are exchanged and updated in order to provide current meta-data on both levels. The sensor readings of the vehicles are further reduced to important and compressed sensor readings and are transferred to the MES in order to allow monitoring and closed-loop control. The MES assigns objectives, tasks and orders to the MPCDS and may also assign weights or priorities.

4.4 MES level and ERP level

The information from the lower levels are "digested" on these levels and planning operations leadings to future objectives/tasks/orders for the MPCDS are carried out. However, these levels are beyond the scope of the described research and are therefore not discussed in detail.

5. APPLICATION SCENARIOS

In order to underline the potential of and the necessities leading to the presented concept a small number of application scenarios is elucidated in this section on the two core levels "component device" and "MPCDS".

5.1 Component/device level

A very straight forward example for the application of this concept might be the detection of a loss of pressure in a tire by the component itself. This application scenario is sketched in Figure 8.

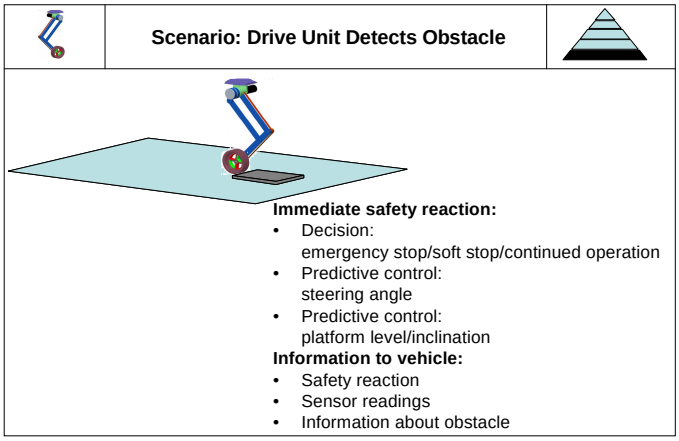


Fig. 8. Scenario on the component/device level.

The local intelligence in the component is continuously calculating a (simplified) model of the motors and can by means of comparing the results of this model with current sensor readings detect a pressure loss which will result in an increase of power demand. The quickest reaction will be carried out by the component itself (compare Figure 2). For instance by means of a rule-based procedure the intelligence on the component will decide the first reaction, e. g. emergency stop if the pressure loss is large and if a danger for the vehicle or human beings may be the result of the pressure loss. This decision will also be based on the hierarchical objective system, which informs even the component roughly about the amount of danger. Furthermore extremely high temperatures of the motors caused by the additional power demand can be avoided by means of predictive control. The component will also communicate with the vehicle in order to allow the necessary activities on the next higher level.

5.2 MPCDS level

The MPCDS level is responsible for the monitoring, planning, control and diagnosis of a sub-section of a production system usually containing a number of vehicles. Possible application scenarios on the MCDS level can rather be found in the area of combined monitoring. Cameras installed in the production sub-section may detect obstacles which may hinder the movement of the autonomous vehicles (e. g. building material left by inexperienced service personnel). In this scenario also the sensor readings form multiple autonomous vehicles can be collected and compared allowing higher precision, plausibility checks and advanced knowledge about the state of the respective sub-section of the production system and the obstacle. This kind of scenario is shown in Figure 9.

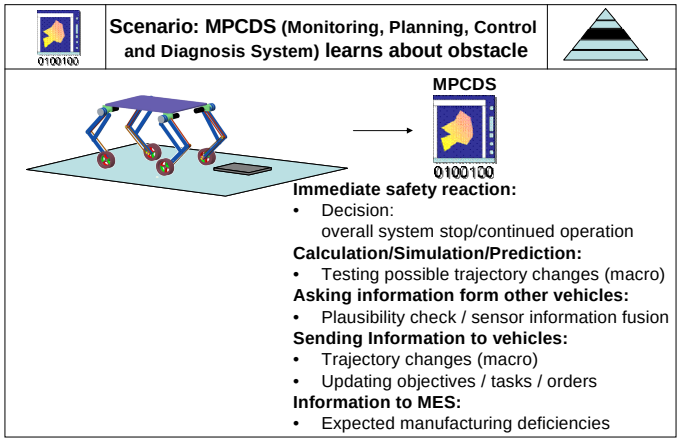


Fig. 9. Scenario on the MPCDS level.

6. SUMMARY

This paper presented a distributed and hierarchical concept for monitoring, planning, control and diagnosis. Elements on all lower levels of this concept are currently realised at the Hochschule Ravensburg-Weingarten. Further application areas which have been identified are pump systems and electrical cars (eMobility).

REFERENCES

Bidgoli, H.: The Internet Encyclopedia, Volume 1, New York: John Wiley & Sons, 2004.

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M.: Diagnosis and Fault–Tolerant Control. Berlin: Springer, 2006.

Camacho, E., Bordons, C.: Model Predictive Control. Berlin: Springer, 2004.

Dabrowska, A., Kleinmann, S.: Analysis of Possible Applications of the Advanced Modelling and Diagnosis System AMandD in German Industry. In: Proceedings of the 7th Workshop on Advanced Control and Diagnosis, ACD'2009, 19. und 20. November 2009, Zielona Góra, Poland.

Diwakar, S. Essawy, M.A. Sabatto, S.Z.: An intelligent system for integrated predictive diagnosis. In: Proceedings of the Thirtieth Southeastern Symposium on System Theory, 1998.

Greenfield, A: Everyware - The dawning age of ubiquitous computing. Peachpit Press, 2006.

Isermann, R.: Mechatronic Systems: Fundamentals. Berlin: Springer, 2005.

Isermann, R.: Fault–Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance. Berlin: Springer 2005.

Kleinmann, S.; Dabrowska, A.; Koller-Hodac, A.; Leonardo, D.: Model of a combined pump and drive system for advanced control and diagnosis. In: Proceedings "Mechatronic Systems and Materials" (MSM 2010).

Korbicz, J., J.M. Kościelny, Z. Kowalczuk and W. Cholewa: Fault Diagnosis: Models, artificial intelligence methods, applications. Springer: Berlin, 2004.

Koscielny J..M., Syfert M., Wnuk P.: Advanced monitoring and diagnosis system "AmandD", In: Proceedings of SafeProcess, Beijing, 2006.

McCellan, M.: Applying Manufacturing Execution Systems. Boca Raton: CRC Press, 1997.

Stania, M., Stetter, R.: "Mechatronics Engineering on the Example of a Multipurpose Mobil Robot". In: Solid State Phenomena Vols. 147-149 (2009) pp 61-66.

Voos, H.; Stetter, R.: "Design and Control of a Mobile Exploration Robot". In: „Proceedings of Mechatronics 2006. 4th IFAC-Symposium on Mechatronic Systems". Heidelberg, 2006.

Wang, Y., Boyd, S.: Fast Model Predictive Control using Online Optimization. In: Proceedings of the 17th World Congress. The International Federation of Automatic Control. Seoul, Korea, July 6-11, 2008.

Ward, S.: Electrical Engineering. Global Media: 2007.

Witczak, M.: Modelling and Estimation Strategies for Fault Diagnosis of Non–Linear Systems: From Analytical to Soft Computing Approaches. Lecture Notes in Control & Information Sciences. Berlin: Springer 2007.

Zhang, Y., Jiang, J.: Bibliographical review on reconfigurable fault–tolerant control systems. Annual Reviews in Control, 32, 229–252, 2008.

Ziemniak P.; Ucinski D.; Paczynski A.: "Robust Control of an All-Terrain Mobile Robot". In "Solid State Phenomena" Vols. 147-149 pp 43-48, Switzerland, 2009.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Reliability Assessment of Technical Devices Based on Degradation Data and Stochastic Equations

Ryszard Kopka

Institute of Control and Computer Engineering, Opole University of Technology,
45-272 Opole, Poland (e-mail: r.kopka@po.opole.pl)

Abstract: Degradation processes taking place in technical elements and objects are more and more frequently used to assess a technical state and to modify mathematical models of devices and their controlling procedures. Widely used simple linear models may be replaced by Levy's complex stochastic processes or Markov's diffusive processes. These models inherently may take into consideration both the changeability of the degradation process of a single element and changes differing from unit to unit. Results of using Markov's stochastic processes to describe the progressing degradation processes and to estimate the reliability function, defining the probability of achieving the adopted limit level by this process are presented in this paper.

Keywords: reliability, degradation processes, stochastic equations.

1. INTRODUCTION

Continuous technological progress is followed by benefits resulting from limiting the use of energy or emission of flue gas but becomes a potential source of dangers connected with the increasing complexity of the controlling process or the use of specific methods or substances. The increasing reliability requirements made the contemporary technological objects cause the necessity of using more and more advanced methods and control systems. It also concerns controlling systems (Korbicz et al., 2009). The possibility of continuous control of some degradation processes allows not only to predict the time of changing such an element or making its overhaul but also may be used in order to improve the quality of controlling the process itself. A huge increase of possibilities of modern diagnostic elements and systems in the scope of measuring and communication possibilities allows to control continuously many parameters of the process and at the same time send them and then process and archive in computer systems. The increase of the computing possibilities allows to use more and more complex analyses realized by using neural networks, fuzzy logic, evolutionary algorithms or expert systems.

An assessment of the technical state does not concern only the estimation of its reliability. Progressive degradation processes may in a considerable way influence the properties of the element or system causing an influence of these changes on remaining elements of the system (Korbicz et al., 2009). It forces the necessity to modify mathematical models of the objects and algorithms of them controlling.

A traditional way of assessing the reliability is connected with an observation of times since the immediate damages causing a complete end of the realization of utility functions were appeared. However, in the case of more and more reliable elements or equipment, sudden damages occur more and more rarely. It causes that an assessment of reliability is inaccurate or even impossible to make because of a small number of these data or even the lack of data. However, additional, valuable information may be delivered by data describing degradation processes which take place (Chandrupatla, 2009; Hamada et al., 2008). Such processes are a typical reaction of an element or a system on conditions in which it is used and they may supply us with information on the technical state without the necessity of waiting till the moment of an occurrence of a catastrophic damage. Additionally, an observation of continuous, gradual changes of the properties of elements allows us to modify the description of the controlling object occurring in the form of its model.

In the last years we observe an important increase in interests in the possibility of using degradation data in order to infer about the technical state in different domains of technique (Pham, 2003). They concern electronics (Fukuda, 1991; Kopka, 2009), mechanics (Sobczyk 1991, 2000) and controlling techniques (Korbicz et al., 2009). It turns out that they are followed by significant information on the reliability properties of the elements and necessary changes which are to be made in the models and the controlling signals.

Measuring the progressing degradation process may be made in a direct way by measuring the controlled value or by an observation of certain, other parameters, changes of which are the result of occurring processes.

Different ways of describing degradation processes are known. Most authors assume general degradation path models (Meeker and Escobar, 1998; Yang, 2009). Moreover, other (Bogdanof and Kozin, 1985) propose to use linear models but after the introduction of some transformations of variables. Park and Padgett (Park and Padgett, 2005) proposed to use the Wiener's and gamma processes. In a similar way, Lehmann (Lehmann, 2006) proposes connecting these processes with the Poison's double stochastic process. Zhao and others (Zhao and Elsayed, 2004) propose to connect the Wienere's process describing parametric

damages with the Weibul's process for hard faults in order to model the LED diodes degradation process. Sobczyk (Sobczyk, 2000) proposes stochastic dynamic models to describe degradation stiffness changes. Bae (Bae, 2007) proposes to draw conclusions on distribution of time to failure based on the degradation process with taking into consideration additive and multiplicative models. Moreover, in order to make degradation tests in the work, he proposes nonlinear models with random coefficients. Kościelny and others (Korbicz et al., 2000) propose using the linear model in order to describe the sedimentation process of the regulation valve.

Because of the fact that degradation processes are the result of many random factors, stochastic models are the best way to describe them. The following models belong to them: models of stationary and independent growth, more complex models using Markov's processes, diffusion Markov's processes or processes with marked points.

Recently, it has been possible to observe a significant growth of the interest in the possibility of using stochastic processes of diffusion to describe the technical state of objects. It results mainly from more and more calculation possibilities of current computers making it possible to introduce very complex and time-consuming calculations in a relatively short time.

2. STOCHASTIC MODELS

Reliability properties of elements or devices are connected with the possibility of realizing use functions ascribed to them. Damages which can occur during the operation can cause a complete end to possibilities of realization of these functions or can cause only worsening of certain properties influencing the use properties. However, exceeding a given limit value must also be treated as a state of damage and can cause its exclusion. Damages of this type are an effect of certain progressing changes taking place inside the operated elements and the time of their occurrence depends on very many factors resulting from conditions of their operation and from acquired features at the stage of projection and production.

A lot of factors influence the progressing degradation process and a wide range of their changeability cause that it is possible to use stochastic differential equations for their description.

The stochastic process Y_t in the general form can be described as a sum of two factors

$$\mathrm{d}Y_t = \alpha(Y_t,t)\mathrm{d}t + \sigma(Y_t,t)\mathrm{d}Z_t, \tag{1}$$

where $\alpha(Y_t,t)$ is called a drift, $\sigma(Y_t,t)$ the diffusion of the process and $\mathrm{d}Z_t$ is a given stochastic noise. The drift shows a certain tendency connected with the occurring process, whereas the diffusion represents certain disruptions influencing the degradation process. The noise may be modeled by different stochastic processes. It may be described by the Poison's process. Then, it shows the occurrence of rare disruptions but of relatively big amplitudes. It may be used in the Wiener's process, which in turn presents the occurrence of changes which are frequent but are characterized by small amplitudes. It may also be the gamma process which is different from the Wiener's process because it is strictly monotonous.

Degradation processes occurring inside elements or devices may take place in different ways. Additionally, other accessible values connected with the degradation process are frequently observed and registered in the measuring systems. Thus, different models can be used for its description so as to present the analyzed processes in the most accurate way as it is possible.

Several basic stochastic models are defined in the literature. They result from considering certain assumptions concerning the drift module and the diffusion. The first model describes the linear dependency of the degradation process in the time function

$$\mathrm{d}Y_t = \left(\alpha - \sigma^2/2\right)\mathrm{d}t + \sigma\mathrm{d}W_t, \quad Y_0 = y_0. \tag{2}$$

The second describes the processes taking place in the concave way

$$\mathrm{d}Y_t = (\alpha Y_t + \beta)\mathrm{d}t + \sigma\mathrm{d}W_t, \quad Y_0 = y_0, \tag{3}$$

and the third describes processes taking place in the convex way

$$\mathrm{d}Y_t = \alpha Y_t\mathrm{d}t + \sigma Y_t\mathrm{d}W_t, \quad Y_0 = y_0. \tag{4}$$

Model (4) is so called geometric Brownian motion. It is necessary to take into consideration the fact that the model form also depends on the value of coefficients. Exemplary models of degradation processes in the time function are presented in Fig. 1.

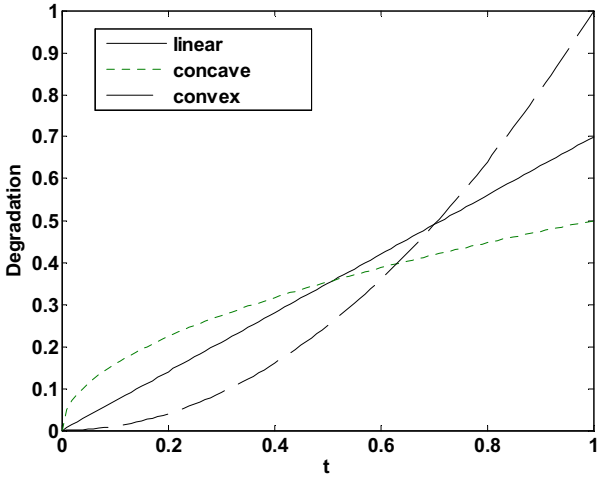


Fig. 1. Different models of degradation processes.

Apart from the one-dimension models it is also possible to use multi-dimension

$$\mathrm{d}\mathbf{Y_t} = \boldsymbol{\beta}(\boldsymbol{\alpha} - \mathbf{Y_t})\mathrm{d}t + \mathbf{g}\mathrm{d}\mathbf{W_t} \text{ and } \mathbf{Y_0} = \mathbf{y_0}, \tag{5}$$

where

$$\mathbf{Y_t} = \left(Y_t^1, Y_t^2, ..., Y_t^n\right)^{\mathrm{T}}, \quad \boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)^{\mathrm{T}}, \tag{6}$$

$$\mathbf{y_0} = \left(y_0^1, y_0^2, ..., y_0^n\right)^{\mathrm{T}}$$

and

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{pmatrix}. \tag{7}$$

The solution of the stochastic process

$$\mathrm{d}Y_t = f(Y_t, t; \theta)\mathrm{d}t + g(Y_t, t; \theta)\mathrm{d}W_t \tag{8}$$

is connected with finding the vector of parameters θ. It is assumed that the forms of the function $f(\cdot)$ and $g(\cdot)$ are known. They depend only on the unknown vector of parameters θ. Knowing the transition probability density function $p_y(y_t; y_s, \theta)$ of the process Y_t for $s < t$, it is possible to define a vector θ using the maximum likelihood method (MLE). Assuming that the initial value y_0 is deterministic and that $y_0, y_1, ..., y_n$ is a sequence of historical observation from the diffusion process sampled at discrete time $t_0 < t_1 < \cdots < t_n$, the log-likelihood function of θ is given by

$$l_n(\theta) = \sum_{i=1}^{n} \log p_y(y_i; y_{i-1}, \theta) \tag{9}$$

and the maximum likelihood estimator of $\hat{\theta}$ can be found by maximizing (9) with respect to θ. However, for many stochastic differential equations, the closed form of the transition density function is not known. The possible solution is the numeric estimation of the density function through applying appropriate algorithms of discretization. Several approximation methods are applied: the method of indirect variables, Hermit's polynomials or solutions of Fokker's, Planck's and Kolmogorov's equations (Kostrzewski, 2006; Picchini, 2008).

Using the observation of the degradation process to assess the reliability properties, the moment of the element damaging is defined as the time in which the degradation process exceeds a given, defined level. Assuming that the given process Y_t and its set S form the subset $\Re$ (so $S \subseteq \Re$), for which it is assumed that the element remains in form. Taking into consideration the fact that the limit level is equal D_f, where D_f is the boundary of S, and $Y(t_0) = y_0 \in S$ for the time t_0, the probability that the element or the object remains in the functional zone S in a given moment t is specified as (Sun, 2008)

$$\begin{aligned} R_S(t, t_0, y_0) &= \Pr\{t < T \cap Y(t) \in S \mid Y(t_0) = y_0\} \\ &= \int_S p_y(y, t \mid y_0, t_0)dy, \end{aligned} \tag{10}$$

where T is the first time when the process Y_t crosses the boundary D_f. For the adopted assumptions, the probability density function for the variable T, assigned as $p_T(t \mid y_0, t_0)$ is described as

$$p_T(t \mid y_0, t_0) = -\frac{\partial R_S(t, t_0, y_0)}{\partial t_0}. \tag{11}$$

Thus, using stochastic equations to describe the degradation processes, it is necessary to possess the knowledge of their transition density function because estimation of the probability density function p_T and the reliability function R_S based on it.

In general, for a large number of the observed degradation processes, the reliability function distribution may be estimated on the basis of the number of processes which reached the given critical level D_f, for a given moment t, that is as

$$R(t) \approx 1 - \frac{\text{Number of Processes First Crossing Times} \le t}{N}. \tag{12}$$

The precision of the reliability estimation is directly connected with the number of the observed processes and the standard deviation is estimated as

$$\sqrt{R(t)(1 - R(t))/N}, \tag{13}$$

where N is the number of simulated processes. The exactness of the stochastic process parameters estimation is possible only on the basis of determining the range limited by quantile lines. For the consider confidence level $\alpha \in (0,1)$, the α-quantile line is called the deterministic function q_α meeting the condition (Kostrzewski, 2006)

$$\Pr\{Y_t \le q_\alpha(t)\} = \alpha. \tag{14}$$

Calculating the value q_α corresponds to the determination of the quantile line α of the random variable Y_t for a given moment t. The procedure of estimating the quantile lines is based on putting the generated R-the element of the Y_t process sample in the increasing order and then, determining the index, that is the function $k(\alpha)$, in accordance with the dependency

$$k(\alpha) = \begin{cases} \alpha R & if \ \ \alpha R \in N, \\ [\alpha R] + 1 & f \ \ \alpha R \notin N. \end{cases} \tag{15}$$

It is necessary to take $k(\alpha)$-the element of the sample ordered in such a way, for the quantile of the range α.

The rate of the degradation processes depends on many factors. Having a possibility to control them during a laboratory analysis, it is possible to significantly accelerate the processes occurring in the elements. For semiconductors one of the accelerating factor is the temperature. Its influence is defined by the Arrhenius' law

$$AF_T = \exp\left(\frac{\Delta H}{k}\left(\frac{1}{273 + T_N} - \frac{1}{273 + T_U}\right)\right), \tag{16}$$

where T_N and T_U is the temperature of the normal work and the temperature of making tests respectively, ΔH is the activation energy for a given semiconductor element and k is the Boltzmann's constant.

In the case of the description of the probability density function of the time to be met by the degradation process of the adopted limit level by the inverse Gaussian distribution, described by the equation (13), the value of the acceleration factor of the ageing process can be used as

$$X \sim N(\mu,\sigma) \Rightarrow AF_T X \sim N(AF_T\mu, AF_T\sigma). \qquad (17)$$

3. APPLICATION EXAMPLE

The possibility of using the Markov's diffusion models to assess the state of technical objects was made based on the laser diode. The example demonstrates the real degradation process that can take place in semiconductor device. The stabilization of the optical power was realized by the photodiode built in the common casing. The photodiode current in the feedback circuit controls the efficiency of the current source supplying the laser diode. The scheme of the measuring set up is shown in Fig. 2. The computer role is both gathering the data and generating control signal to change the intensity of current source to keep the optical power constant. The signal is generating based on the photodiode current, measured by A/D converter.

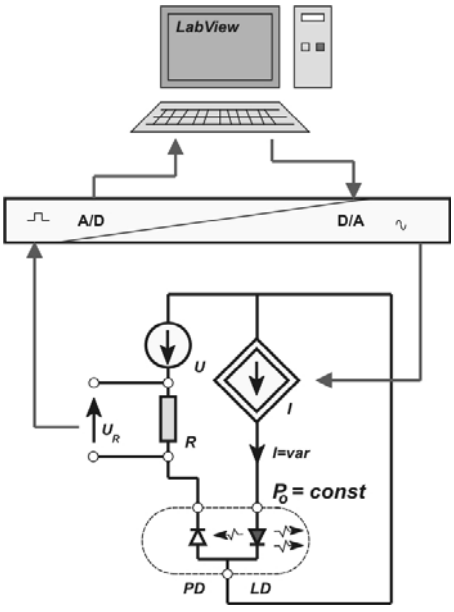


Fig. 2. Schematic diagram of the measuring set up for the laser diode degradation process.

The measuring of the degradation process was made at a temperature of 54°C. The initial value of the laser diode

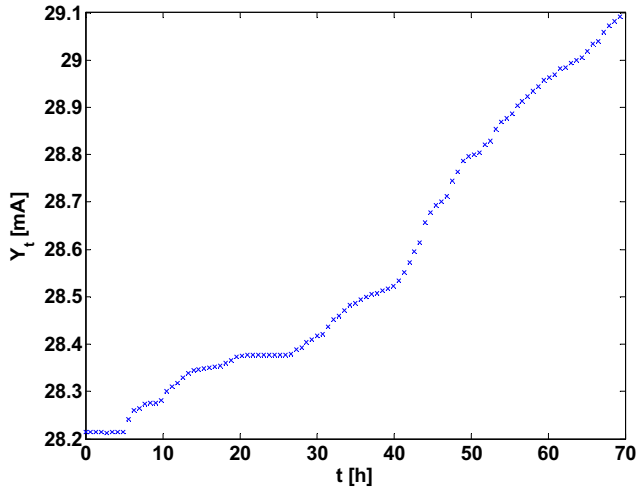


Fig. 3. Measured laser diode degradation process (Y_t – laser diode current, generating the same optical power).

current was set to its nominal value ($I_N = 28\,mA$). Supply current changes generating the same optical power during tests are shown in Fig. 3. Based on these value the parameters of the stochastic model were determined. A convex model named a geometric Brownian's motion (4) was adopted. Used SDE Matlab Toolbox (Picchini, 2008), the following values were calculated: $\hat{\alpha} = 0.00046$ and $\hat{\sigma} = -0.00038$. Simulated degradation processes based on calculated values of the parameters are shown in Fig. 4.

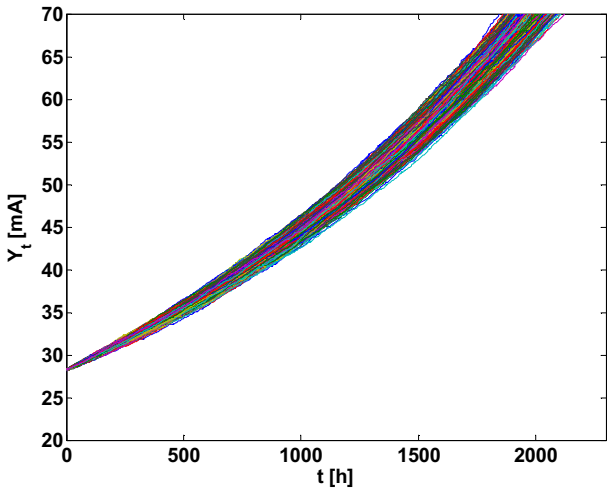


Fig. 4. Simulated laser diode degradation processes for parameters received based on the measurements.

The only way of determining the accuracy of the values of the received parameters is drawing quantile lines. A range limited by 95% of quantile lines, determined for the derived parameters is presented in Fig. 5.

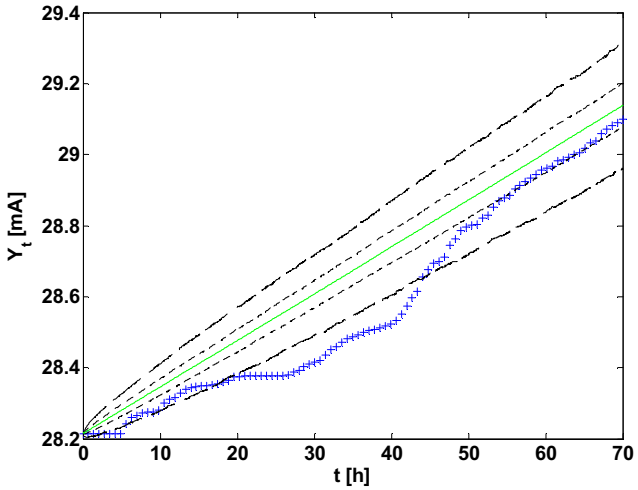


Fig. 5. Quantile lines defining the accuracy of the determined parameters of the stochastic model.

Taking into consideration that the limit level for the laser diode current is equal to 150% of the nominal value, based on the generated processes, times for which the value of the degradation process will reach the limit level were determined (Fig. 6).

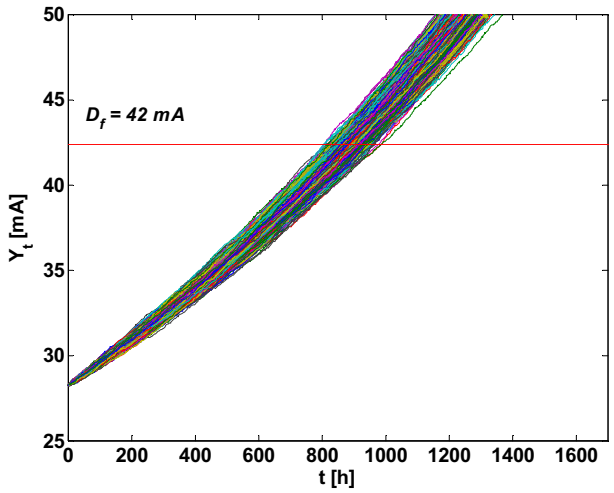


Fig. 6. Simulated processes of degradation of a laser diode with the marked limit level specifying 50% of the diode current increase.

Based on these times the parameters of the probability density function of normal distribution were determined. The mean value $\hat{\mu}_U = 887{,}5$ and the standard deviation $\hat{\sigma}_U = 25.2$ were calculated. Using equations (17) parameters of probability density function corresponding to normal conditions (25°C) could be estimated. For the tested laser diode $\Delta H = 0.7eV$ and $T_N = 25^{\circ}\mathrm{C}$ were adopted together with the temperature of the test equal to $T_U = 54^{\circ}\mathrm{C}$ and $k = 8.62 \times 10^{-5} eV/K$. Using these data based on equation (16), $AF_T \approx 11.2$ was received and using (17) $\hat{\mu}_N = 9948.4$ and $\hat{\sigma}_N = 282.7$ were determined. The probability density function for both, normal and ageing conditions, assuming 50% increase of the laser diode current, are presented in Fig. 7.

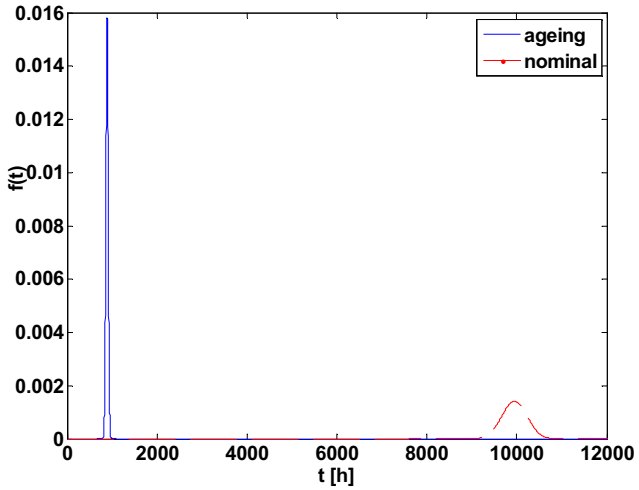


Fig. 7. Probability density function for the adopted 50% of the diode current increase for the accelerated ageing process (solid line) and for normal work conditions (dashed line).

The reliability function in the case of the accelerated reliability tests and for nominal conditions is presented in Fig. 8.

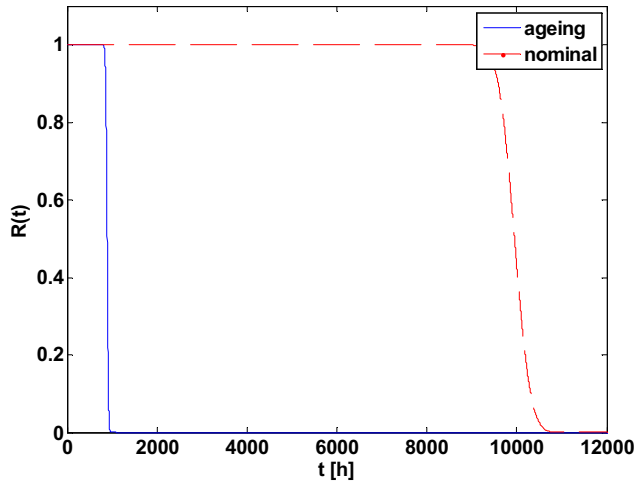


Fig. 8. Reliability function received for the adopted 50% increase of the diode current, for an accelerated ageing process (solid line) and for conditions of a normal work (dashed line).

It gives a probability that a laser diode does not reach the current greater than 150% of the nominal value for a given value, for conditions of an accelerated ageing process and for normal working.

4. CONCLUSION

Degradation processes comprise important information describing not only the current reliability level of an element or a device but also its properties from the point of view of the controlling process. The possibility of their measuring and analyzing can specify the necessary time of turning the device off and can have an influence on the procedures of its regulation. A large number of factors influencing on the process and a wide range of their variation prefer to use stochastic models for their description. The use of complex diffusion models of unknown transition probability density functions is possible due to more and more fast computer making it possible to simulate the process repeatedly. The parameters values of the model can be estimating only using numeric procedures. A significant problem is accuracy assessment. It is possible based only on quantile lines received by multiple simulations of the process.

REFERENCES

Bae, J.S., Kuo W., Kvam H.P. Degradation models and implied life time distributions. *Reliability Engineering and System Safety*, vol. 92, pp. 601-608, 2007.

Bogdanoff, J.L., Kozin, F. *Probabilistic Models of Cumulative Damage*. John Wiley & Sons, New York, 1985.

Chandrupatla T.R. *Quality and Reliability in Engineering*. Cambridge University Press, New York, 2009.

Fukuda, M. *Reliability and Degradation of Semiconductor Lasers and LEDs*. Artech House, Boston, 1991.

Hamada, M.S., Wilson, A.G., Shane Reese, C., Martz, H.F. *Bayesian Reliability*. Springer Science+Business Media, New York, 2008.

Jain, R.K. Inverse Gaussian Distribution and its application to reliability. *Microelectronic Reliability*, vol. 36, no. 10, pp. 1323-1335, 1996.

Kopka, R. *Diagnostics of optical devices based on degradation processes*. 7th Workshop on Advanced Control and Diagnosis, ACD 2009, 19-20 November 2009, Zielona Góra. Available: http://www.issi.uz.zgora.pl/ACD_2009/program/Papers/87_ACD_2009.pdf

Kostrzewski, M. *Bayesian analyses of financial time series modeled by diffusion processes*. Uczelniane Wydawnictwa Naukowo-Dydaktyczne, Kraków, 2006 (In Polish).

Korbicz, J., Kościelny (Eds.), J. M. *Modeling, Diagnostics and Process Control. Implementation in the DiaSter System*. Warszawa: WNT, 2009 (in Polish).

Lehmann, A. Degradation-threshold-shock models. *Probability, Statistics and Modelling in Public Health*, New York: Springer, 2006, pp. 286–298.

Meeker, Q. W., Escobar, A. L. *Statistical Methods for Reliability Data*. John Wiley & Sons, Inc., New York, 1998.

Park, Ch., Padgett, J.W. Accelerated Degradation Models for Failure Based on Geometric Brownian Motion and Gamma Processes. *Lifetime Data Analysis*, vol. 11, pp. 511-527, 2005.

Pham, H. *Handbook of reliability engineering*. Springer, London, 2003.

Picchini, U. *SDE Toolbox: An introduction to the Simulation and the Numerical Solution of Stochastic Differential Equations with Matlab*. Available: http://sdetoolbox.sourceforge.net (05.01.2008).

Sobczyk, K. *Stochastic Differential Equations: With Applications to Physics and Engineering*. Kluwer Academic Publishers B.V., Dordrecht, 1991.

Sobczyk, K., Trebicki, J. Stochastic dynamics with fatigue-induced stiffness degradation. *Probabilistic Engineering Mechanics*, vol. 15 issue 1, pp. 91-99, 2000.

Sun, J.Q. *Stochastic Dynamics and Control*. Elsevier, Amsterdam, 2006.

Zhao, W., Elsayed, E.A. An Accelerated Life Testing Model Involving Performance Degradation, Reliability and Maintainability, Annual Symposium-RAMS, Los Angeles CA, 2004, pp. 324-329.

Yang, G., Reliability Demonstration Through Degradation Bogey Testing, *IEEE Trans. Reliability*, vol. 58, no. 4, pp. 604-610, 2009.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Intelligent techniques for faults diagnosis and prognosis of CHP plant with gas turbine engine

L. Miozza * **A. Monteriù** * **A. Freddi** * **S. Longhi** *

* *Università Politecnica delle Marche, Ancona, Italy*
(email: luimiozza@libero.it, a.monteriu@univpm.it, freddi@diiga.univpm.it, s.longhi@univpm.it)

Abstract: Prognostic and Health Management (PHM) systems are used for the diagnosis and the prognosis of faults that affect the lifetime of dynamic processes. They can detect, identify and isolate fault mechanisms and predict their evolution in order to prevent them from causing system failures. When used in combination with Condition-Based Maintenance (CBM), they not only extend the system reliability, but significantly reduce operating costs and emergency response strategies through intelligent planning of maintenance. In this paper a CBM/PHM system for a cogeneration plant with gas turbine has been developed. Simulation trials have tested the performance and functionality of the system, showing how to produce useful information for health management and intelligent maintenance.

Keywords: Diagnosis, Prognosis, Condition-Based Mainentance, Prognostic Health Management, Intelligent Mainentance

1. INTRODUCTION

Condition-Based Maintenance (CBM) and Prognostics and Health Management (PHM) have emerged over recent years as significant technologies which are making an impact on both military and commercial maintenance practices. These technologies improve reliability and safety of machines, with a dramatic decrease in life-cycle costs. The goal is achieved by diagnosis and prognosis of faults before they result in the system failure. As noted by Vachtsevanos et al. [2006], the growth in the diagnostic capability of modern systems has naturally evolved into something more: the desire for prognosis. Since it is already possible to use existing data and data sources to diagnose faulty components, it would be of great interest to use the same information to predict future failures before they can actually affect the capability of the machine to work properly. In this paper the design of a CBM/PHM system for a Combined Heating and Power (CHP) plant with gas turbine is presented. The plant consists in a set of several linked subsystems and is described by a mathematical non-linear model. Plant measurements are made available to the health management system which elaborates them and extracts information useful to diagnose faults that affect the components lifetime. Then, the prognosis module forecasts the Remaining Useful Life (RUL) to system failure. The CBM module, finally, intelligently decides the maintenance tasks to perform.

The paper is organized as follows. Subsection 2.1 presents the hardware and the considered model. Subsection 2.2 describes the proposed CBM/PHM architecture following the ISO 13374 rules. Section 3 analyzes the failure modes that may affect the subsystems of the plant and how these modes can be exploited to design the CBM/PHM system, defining a special library. Sections 4, 5 and 6 contain the description of the functional modules of the CBM/PHM system. In Section 7 some significant simulation results are presented. Finally Section 8 concludes the paper.

2. CBM/PHM SYSTEM FOR CHP PLANT

2.1 CHP plant description

The process considered is the CHP plant T100 of Turbec S.p.A.. This plant generates electrical power (max 100 ± 3 kW) and thermal power (nom. 155 ± 5 kW) from combustion of methane. A plant scheme is shown in Fig. 1.

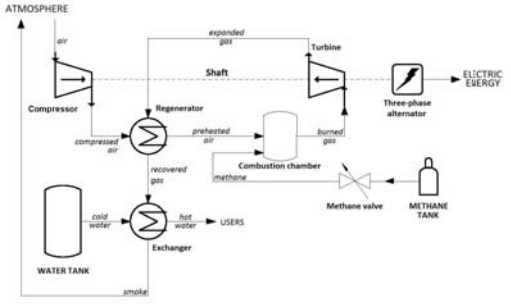


Fig. 1. Scheme of CHP plant.

Thermodynamic cycle of reference is Joule's or Brayton's cycle:

(1) The compressor increases pressure and temperature of the atmospheric air;
(2) The air is burned with methane at constant pressure;
(3) The burned gas expands in the turbine, thus activating the shaft linked both to compressor and generator;

(4) Finally, the hot gas transfers heat to cold water.

The Mathematical model (Feliciotti [2009]) is represented by differential equations of each subsystem. The model is implemented using Matlab® and Simulink® .

2.2 Architecture of CBM/PHM system for CHP plant

The architecture adopted in this work follows the conceptual framework provided by the International Standards Organization (ISO). Fig. 2 depicts the ISO 13374-1 processing model. The base of integrated architecture is the database which stores data from multiple and diverse sensor suites. Information are extracted in the form of features or condition indicators and used as input to diagnostic and prognostic routines. Integrated system architecture involves six main modules:

(1) a set of sensors and suitable sensing strategies collect process data of critical variables and parameters;
(2) a features extraction module selects useful information from raw data;
(3) an operating mode identification routine determines the current operational status of the system and correlates fault mode characteristics with operating conditions;
(4) the diagnostic module assesses through online measurements the current state of critical machine components;
(5) the prognostic module estimates the RUL of a failing component/subsystem;
(6) the final module of the integrated architecture (mainentance scheduler or advisory generation) schedules mainentance operations without affecting adversely the overall system functionalities.

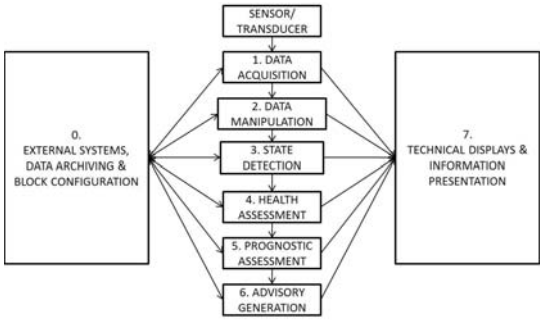


Fig. 2. The ISO 13374-1 processing model.

From the standard architecture one can identify an online implementation phase (see Fig. 3) and a preliminary offline phase for the CBM/PHM system. The offline phase consists of the required background studies that must be performed offline prior to implementation of the online CBM phase and includes determination of which features are the most important for machine condition assessment, the Failure Mode and Effect Critically Analysis (FMECA), the collection of machine legacy fault data to allow useful prediction of fault evolution and, finally, the specification of available resources to perform maintenance tasks. The online phase includes data processing functions: signal processing, extracting the features that are the most useful for determining the current status or fault condition of the

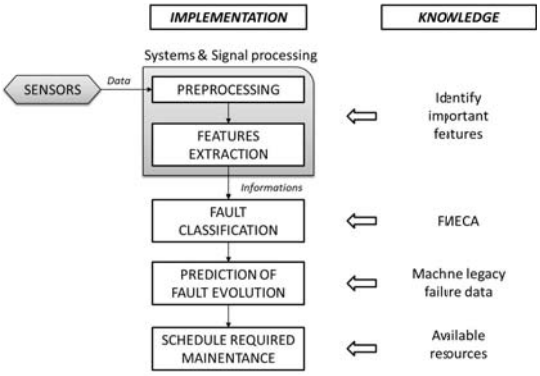


Fig. 3. CBM/PHM system.

machinery, fault detection and classification, prediction of fault evolution and scheduling of required mainentance.

The System and signal processing module can be seen as a virtual sensor (Marko et al. [1996]) that from simple and reliable measures, produces a value that quantifies the current magnitude of fault, generally not directly measurable. The Systems and signal processing functions implement all the calculations made on raw data to extract useful information. These functions can be divided in two subfunctions:

Pre-processing Set of techniques to improve signal quality in terms of signal to noise ratio
Features extraction Set of techniques to extract fault condition indicators.

The Fault Classification module performs fault diagnosis. Its output is the current faults, pending and incipient, or the failure conditions. It is necessary that the diagnosis is completed when the fault is in the incipient stages, otherwise prognosis, and all CBM/PHM system, becomes useless. Terminology for fault diagnosis refers to Isermann and Ball [1997].

Prediction of fault evolution is performed by the prognostic module. The following characteristics are required:

- direct or indirect measurements of the size of the fault at current time and for a number of past samples of times;
- algorithms for predicting the remaining time to failure to provide an estimate close to the actual value (in terms of probabilities);
- estimation of the uncertainty bounds or confidence limits associated with the prediction;
- learning strategies that enhance the richness of reference sample on the basis of new knowledge gained in operation, improving performances such as accuracy, precision and confidence.

3. FAILURE MODE ANALYSIS ON CHP PLANT

The most important features for a CHP plant can be derived analyzing the states of the methane valve, the mechanical transmission system and the heat exchanger.

3.1 Fault of methane valve

The methane valve is the actuator that provides the power for the combustion. The valve is modelled as a first order

system with delay and time costant $\tau_{valve} = 0.5\,\mathrm{s}$

$$\dot{m}_{\mathrm{CH_4}}(s) = \frac{a}{s+a}\dot{m}_{\mathrm{CH_4},des}(s) \tag{1}$$

where $a = \tau_{valve}^{-1}$ $[\mathrm{s}^{-1}] = 2\,\mathrm{s}^{-1}$, $\dot{m}_{\mathrm{CH_4}}(s)$ is the flow provided by the valve and $\dot{m}_{\mathrm{CH_4},des}(s)$ is the flow decided by regulator, both in Laplace domain. Fault mode is modeled as a loss of reactivity due to wear of the valve, i.e. a decrease of value a or an increase of value τ_{valve}. The value, where it is believed that the actuator performances have become unacceptable, is set to

$$\tilde{a} = 1\,\mathrm{s}^{-1}$$
$$\tilde{\tau}_{valve} = 1\,\mathrm{s}$$

The fault mode consists in the slow degradation of the time constant from nominal value τ_{valve} to critical value $\tilde{\tau}_{valve}$.

3.2 Fault of mechanical transmission system

The behavior of the motor shaft is described by the following differential equation

$$\dot{\omega}(t) = \frac{1}{J\omega(t)}\Big(P_{Turb}(t) - P_{Comp}(t) - \frac{1}{\eta_r}P_{el}(t)\Big) - \frac{\mu_{fr}}{J}\omega(t) - \frac{\mu_{lu}}{J}\omega(t)^2 \ [\mathrm{s}^{-2}] \tag{2}$$

where

- $\omega(t)$ $[\mathrm{s}^{-1}]$ is the shaft rotational speed;
- $P_{Turb}(t)$ [W] is the power output from turbine to shaft;
- $P_{Comp}(t)$ [W] is the compressor power;
- $P_{el}(t)$ [W] is the power absorbed by the alternator;
- $J = 0.03\,\mathrm{kg\,m}^2$ is the shaft inertia;
- $\eta_r = 98.5$ % is the alternator efficiency;
- $\mu_{fr} = 1.497 \times 10^{-4}\,\mathrm{kg\,m^2\,s^{-1}}$ is the coefficient of dry friction;
- $\mu_{lu} = 6.990 \times 10^{-9}\,\mathrm{kg\,m}^2$ is the coefficient of lubrificated friction.

The fault mode consists in a slow increase in the values of friction coefficients. Phisically, the increase of μ_{lu} is a sign of inadequate lubrification of the bearings supporting the shaft. Otherwise, an increase of μ_{fr} is the manifestation of phenomena of rubbing or play between mechanical parts related (e.g.: a deformation of turbine or compressor blades, a shaft misalignment due to bad bearings, deflection of the shaft, etc.). Critical values are fixed for

$$\tilde{\mu}_{fr} = 2 \times 10^{-4}\ \mathrm{kg\,m^2\,s^{-1}}$$
$$\tilde{\mu}_{lu} = 10 \times 10^{-9}\ \mathrm{kg\,m}^2.$$

3.3 Fault of heat exchanger

The critical parameter is the exchanger heat transfer coefficient that affects the amount of heat transfered between two fluids and hence the thermal efficiency of plant. In nominal condition the value of the coefficient is

$$K' = 40.07\ \frac{1}{\mathrm{m\,kg\,K}}$$

and the thermal efficiency does not exceed 48.6 %. The wear of the material that separates the fluids involves the gradual loss of its properties and thermal conductor heat transfer coefficient tends to zero. This entails the reduction of heat transfer and therefore the thermal efficiency of plant. Performance is unaccettable when

$$\tilde{K}' = 8.01\ \frac{1}{\mathrm{m\,kg\,K}}$$

and thermal efficiency of plant drops to 20 %.

3.4 The Fault Pattern Library

The Fault Pattern Library (FPL) (Skormin et al. [1994]) groups in a systematic way the condition of disturbance model parameters that can be interpreted as particular fault conditions. These conditions are arbitrary and can be defined by desingers, maintainers, or any process expert. FPL contains, therefore, the diagnostic information in terms of known patterns of failure. In this case, Table 1 may represent the FPL for failures that affect the life of the CHP plant, in terms of variations from nominal values deemed unacceptable. Note that the failures of the valve and the exchanger only affect system performance, therefore, disturbance conditions are less stringent. The shaft failures, instead, pose a greater criticality.

Table 1. Fault Pattern Library for CHP plant.

Failure modes	Parameter perturbations [%]			
	a	K'	μ_{fr}	μ_{lu}
Methane valve	−50			
Heat exchanger		−80		
Shaft (friction)			+30	
Shaft (lubrification)				+40

4. SYSTEM & SIGNAL PROCESSING MODULE

4.1 Pre-processing module

In the pre-processing module data from sensors are digitalized by an acquisition system and filtered with a moving average filter or Finite Impulsive Response (FIR) filter of order n:

$$y_k = \frac{1}{n}\sum_{i=1}^{n}\alpha_i u_{k-i} \tag{3}$$

The sampling time must be properly chosen in order to capture the entire slow dynamics of the process and filter out high frequency noise.

4.2 Features extraction

Following the approach of the virtual sensor, three independent algorithms of Digital Signal Processing (DSP) have been developed. These algorithms produces a signal that expresses the instantaneous magnitude of fault using the digitalized measurements of valve, shaft and heat exchanger variables. The output of the feature extraction module is a vector of characteristics according to the health state of plant, i.e., a set of four variables connected independently to a specific mechanism of fault

$$\boldsymbol{\varphi}_k = \begin{bmatrix} a_k \\ \varphi_{KS} \\ \mu_{fr,k} \\ \mu_{lu,k} \end{bmatrix} \in \mathbb{R}^4 \tag{4}$$

Methane valve • The amount of gas valve fault is given directly by the value of time constant τ_{valve} or its inverse a. It can not be measured directly, but is calculated by finite memory Recursive Least Square (RLS) algorithm. The discrete model of the valve can be written as an AutoRegressive Moving Average (ARMA) from eq. (1)

$$\dot{m}_{\mathrm{CH}_4,k} = \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{\vartheta} + \nu_k \tag{5}$$

where

$$\boldsymbol{h}_k = \begin{bmatrix} -\dot{m}_{\mathrm{CH}_4,k-1} \\ \dot{m}_{\mathrm{CH}_4,des,k} \end{bmatrix} \in \mathbb{R}^2, \quad \boldsymbol{\vartheta} = \frac{1}{1+aT_C}\begin{bmatrix} -1 \\ aT_C \end{bmatrix} \in \mathbb{R}^2$$

ν_k is the white noise representing measurement and modeling uncertainties and T_C [s] is the sampling interval. The algorithm is given by

$$\begin{cases} P_k = \frac{1}{\lambda}\left[P_{k-1} - P_{k-1}\boldsymbol{h}_k\boldsymbol{h}_k^{\mathrm{T}}P_{k-1}(\lambda + \boldsymbol{h}_k^{\mathrm{T}}P_{k-1}\boldsymbol{h}_k)^{-1}\right] \\ \hat{\boldsymbol{\vartheta}}_k = \hat{\boldsymbol{\vartheta}}_{k-1} + P_k\boldsymbol{h}_k(\dot{m}_{\mathrm{CH}_4,k} - \boldsymbol{h}_k^{\mathrm{T}}\hat{\boldsymbol{\vartheta}}_{k-1}) \end{cases} \tag{6}$$

where $P_k \in \mathbb{R}^{2\times 2}$ is the estimation error covariance matrix and $0 \leq \lambda < 1$ is the oblivion coefficient. In this way the estimated parameter of the model is influenced by the last $M = \frac{1}{1-\lambda}$ samples. The algorithm is initialized by choosing the initial estimate as the value of ARMA coefficients with nominal model parameters and initial error covariance matrix as $P_0 = \gamma I$ with γ chosen to adjust convergence speed and I indicates the identity matrix. Fig. 4(a) shows the evolution of the characteristic in a simulation of the failure mode.

Shaft • Estimates of friction coefficients μ_{fr} and μ_{lu} are used as fault indicators for the shaft. Expressing the nonlinear model of the shaft (eq. (2)) in the ARMA form in which the coefficients are linearly related to the measures results in:

$$\omega_k = \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{\vartheta} + \nu_k \tag{7}$$

where

$$\boldsymbol{h}_k = \begin{bmatrix} -\omega_{k-1} \\ a_k \\ b_k \end{bmatrix} \in \mathbb{R}^3, \quad \boldsymbol{\vartheta} = -\frac{1}{J+\mu_{fr}T_C}\begin{bmatrix} J \\ -T_C \\ \mu_{lu}T_C \end{bmatrix} \in \mathbb{R}^3$$

and ν_k is the white noise representing measurement and modeling uncertainties. a_k and b_k signals are obtained by combining system inputs and output as follows:

$$a_k = \frac{1}{\omega_{k-1}}\left(P_{Turb,k} - P_{Comp,k} - \frac{1}{\eta_r}P_{el,k}\right)$$
$$b_k = \omega_{k-1}^2$$

Applying to eq. (6) the finite memory RLS algorithm results in an equation similar to (7), in which the initial estimate is given by the nominal values of model parameters. A simulation of both failure modes is depicted in Fig. 4(c) and 4(d).

Heat exchanger • Let us consider a linearized discrete model of the exchanger in state space form which describes the variations of process variables from the system equilibrium point:

$$\begin{cases} \boldsymbol{\delta\dot{x}}_{k+1} = A_{ex}\boldsymbol{\delta x}_k + B_{ex}\boldsymbol{\delta u}_k \\ \boldsymbol{\delta y}_k = C_{ex}\boldsymbol{\delta x}_k \end{cases} \tag{8}$$

where $\boldsymbol{\delta x}_k \in \mathbb{R}^8$ is the state vector, $\boldsymbol{\delta u}_k \in \mathbb{R}^8$ is the input vector, $\boldsymbol{\delta y}_k \in \mathbb{R}^8$ is the output vector and $A_{ex}, B_{ex}, C_{ex} \in \mathbb{R}^{8x8}$ are the system matrices.

The degradation of exchange material is represented by a decrease in the heat transfer coefficient K' present in the first and fifth row elements of C_{ex} matrix, therefore, it affects only the temperature of output fluids (smoke and hot water). With an Unknown Input Observer (UIO by Chen and Patton [1999]) the system state is estimated and, using the C_{ex} matrix with the nominal value of the coefficients, the theoretical correct temperature outputs of the process are calculated. The differences with the measured outputs produce, therefore, a residual growing with the degradation of the material. The use of UIOs grants better decoupling properties with respect to classical Luenberger Observers. In particular, it avoids the measurement of specific heat at constant pressure (c_p) of inlet and outlet fluids and also avoids the measurements of flow and pressure of incoming fluids. To apply the UIO the linear model described by eq. (8) must be written as

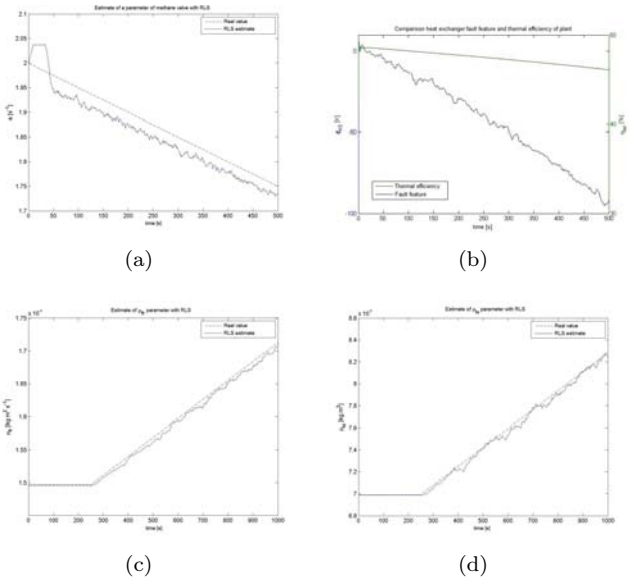


Fig. 4. Module output of feature extraction. In (a) feature a for methane valve fault. In (b) feature φ_{KS} for heat exchanger fault compared with efficiency η_{th} degradation. In (c) and (d) feature for mechanical transmission system faults, respectively μ_{fr} and μ_{lu}.

$$\begin{cases} \boldsymbol{\delta\dot{x}}_{k+1} = A_{ex}\boldsymbol{\delta x}_k + \tilde{B}_{ex}\boldsymbol{\delta\tilde{u}}_k + E_{ex}\boldsymbol{d}_k \\ \boldsymbol{\delta\tilde{y}}_k = \tilde{C}_{ex}\boldsymbol{\delta x}_k \end{cases} \tag{9}$$

where $\boldsymbol{\delta\tilde{u}}_k \in \mathbb{R}^2$ is the known input vector containing the fluid temperatures entering the heat exchanger, $\boldsymbol{d}_k \in \mathbb{R}^6$ is the unknown input vector including pressure, flow and specific heat at constant pressure of incoming fluids, $\boldsymbol{\delta\tilde{y}}_k \in \mathbb{R}^6$ is the measured output vector containing temperature, pressure and flow of exhaust fumes and hot water produced. $\tilde{B}_{ex} \in \mathbb{R}^{8\times 2}$ is the matrix composed of the first and the fifth column of the matrix B_{ex} and $E_{ex} \in \mathbb{R}^{8\times 6}$ is composed of the remaining columns. $\tilde{C}_{ex} \in \mathbb{R}^{6\times 8}$ is obtained removing from C_{ex} matrix the fourth and eight row. Since

$$rank(\tilde{C}_{ex}E_{ex}) = rank(E_{ex}) = 2$$

then the UIO exists and its structure is given by

$$\begin{cases} \boldsymbol{z}_{k+1} = F\boldsymbol{z}_k + T\tilde{B}_{ex}\boldsymbol{\delta\tilde{u}}_k + K\boldsymbol{\delta\tilde{y}}_k \\ \boldsymbol{\delta\hat{x}}_k = \boldsymbol{z}_k + H\delta\tilde{y}_k \end{cases} \tag{10}$$

where

$$H = E_{ex}\left[(\tilde{C}_{ex}E_{ex})^{\mathrm{T}}\tilde{C}_{ex}E_{ex}\right]^{-1}(\tilde{C}_{ex}E_{ex})^{\mathrm{T}}$$
$$T = I - H\tilde{C}_{ex}.$$

Since

$$A_1 = TA_{ex}$$

and since the couple $(\tilde{C}_{ex}, A_1)$ is detectable, matrix K_1 is derived using a pole-placement technique such that matrix

$$F = A_1 - K_1\tilde{C}_{ex}$$

has all eigenvalues on the segment $]-1, 0[$ on the real axis of the complex discrete plane. Matrix K results in

$$K = K_1 + FH.$$

The output observation error is given by

$$\boldsymbol{\varepsilon}_{y,k} = \boldsymbol{\delta\tilde{y}}_k - \tilde{C}_{ex}\boldsymbol{\delta\hat{x}}_k \in \mathbb{R}^6 \tag{11}$$

and only errors on temperature are affected by failure mode (the first and fourth element). Finally, the fault feature for the heat exchanger is given by the weighted sum of the residuals obtained from smoke and hot water temperatures. The merger combines the information contained in more features into a single feature, improving the robustness to disturbance and measurements errors. The fused feature is given by

$$\varphi_{KS} = \frac{r^2_{smoke} + r^2_{\mathrm{H_2O}}}{r_{smoke} + r_{\mathrm{H_2O}}} \tag{12}$$

Fig. 4(b) shows the evolution of the characteristic in a simulation of the failure mode, compared with the degradation of thermal efficiency of plant.

5. FAULT CLASSIFICATION MODULE

The four failure modes can define sixteen categories ranging from no damage to the situation where all three subsystems are affected by pending faults. The classifier must map the space formed by the characteristics vector $\boldsymbol{\varphi}_k$ in the space of sixteen classes of fault. To do this we use a neural network as data classifier. We chose a two-layer neural network with four neurons for each layer. The four inputs correspond to the four elements of characteristics vector, normalized in the interval $[-2, 2] \subset \mathbb{R}$. The four outputs form a binary number based on that, decoded into decimal base, returns the key to assign the fault class. The network is trained with a set of points that describe the FPL. For example, we want the network to assign the fault class 0, i.e. no fault, at the point where all the features assume nominal values. If, however, the data point is characterized by significant degradation of all fault indicators then we want the network to check the class 16 (all failures). The network is implemented using the Matlab® Neural Network Tool. The training ends in five epochs. The classifier is tested on a simulation of an hour and forty minutes during which the plant continuously provides 80 kW of electrical power and any performance degrades simultaneously up to the subsystems failure. The simulation represents also weather conditions (temperature and pressure), using data from various meteorological observation sites.

6. PREDICTION OF FAULT EVOLUTION MODULE

In this paper we use data-driven prognosis techniques, for which there is no need to have a model of the fault mechanism nor historical data. They are basically statistics techniques that project the trend of a variable in the future based on the current and past recorded values. These techniques hardly prescribe significant confidence limits, but their validity can be verified at the design stage. It is clear that a model or historical data describing the behavior of the process in fault situations are very useful. Data-driven techniques are therefore a compromise between the techniques based on model (powerful, but expensive and specific) and probabilistic techniques (less accurate but simpler and wide applicability). An intresting data-driven statistical technique is the autoregressive moving average, commonly used in economics to the Trend Analysis. The principle of the technique is depicted in Fig. 5.

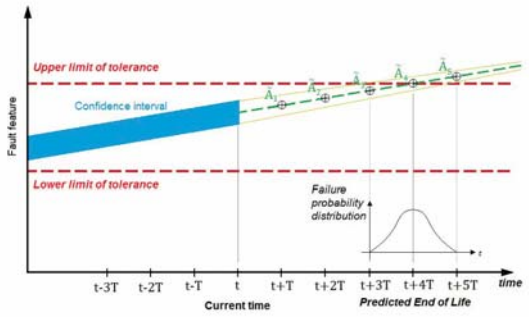


Fig. 5. Technique of forecasting with autoregressive.

A digital processing algorithm produces, every T_C seconds, the time sequence $\{A_{t-kT_C}\}_{k=1,2,3,\ldots}$ of parameter estimates. It contains valuable information that describes how the fault is deteriorating system performance. Let us introduce the autoregression equation

$$A_k = E\{A|A_{k-T_C}, A_{k-2T_C}, \ldots, A_{k-mT_C}\} \tag{13}$$

that expresses the fact that the estimate of the characteristics at current time step k is conditioned by the estimated value already registered in the last m time samples. A finite memory RLS procedure captures the failure modes and then the equation is simulated to generate a series of predictions of the estimate for the future time instants:

$$A_i = E\{A|A_{i-T_C}, A_{i-2T_C}, \ldots, A_k, A_{k-T_C}, \ldots, A_{k-mT_C}\} \tag{14}$$

with $i \geq k$. If the estimated prediction reachs the tolerance level in a reasonable time the system will declare that there is an End of Life (EOL). Otherwise, the system response is that there is no failure of the component in the horizon considered. The tolerance is established in FPL. It is important to clarify that the estimate of the characterisic is not known with certainty, but in a certain confidence interval. The confidence limits of prediction may be prescribed by assessing what is the precision that gives the estimate: looking to the future also mobile variance associated with it. In this way, an estimate which tends to improve its quality will produce a better prediction in the sense that the response of prognosis will have a distribution closer around the mean. The forecast error is evalutated in relative terms compared with actual EOL. A negative relative error indicates a postponed forecast. It is always preferable to have early predictions (positive error):

$$\varepsilon_r = \frac{\mathrm{EOL} - \hat{\mathrm{EOL}}}{\mathrm{EOL}} \tag{15}$$

Fig. 6 and Fig. 7 show the result of applying the algorithm for the coefficient μ_{fr} of the motor shaft (for the other fault modes the result is similar). The test is identical to that made in assessing the response of the classifier. Note in Fig. 7 as 50 min before component failure, EOL is provided 6 min in advance (relative error of 8 %). After 35 min of simulation the error lies constantly below 5 %; 40 min before failure error is predicted within 3 min.

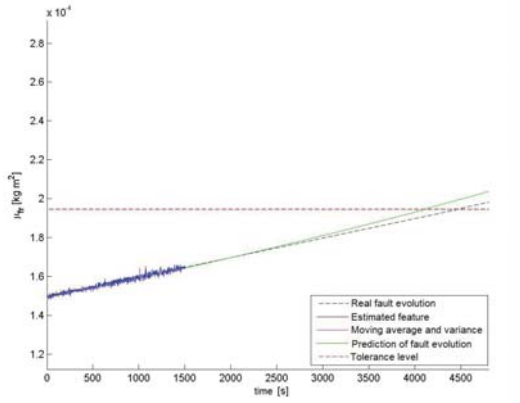


Fig. 6. Result of prognosis for μ_{fr} parameter: the forecast 25 min after the start of simulation.

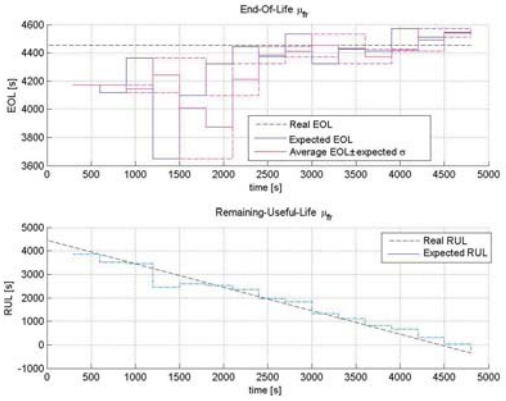


Fig. 7. Result of prognosis for μ_{fr} parameter: the EOL and RUL provided during the entire simulation and compared with real values.

7. RESULTS

To integrate the results reported for the various specific forms, an overall result of a simulation of CBM/PHM system operating on the plant affected by all the considered fault conditions is proposed. The CBM/PHM system must periodically send the following basic information:

- list of incipient fault conditions (diagnosed with EOL not available in finite time): class of fault and detection time t_r;
- list of pending fault conditions: fault class, detection time t_r, prognostic information (EOL e RUL);
- list of failure conditions: failed component, failure time t_f.

Table 2 shows an example of the type of information the system produces, reported to the case of CHP plant. Data are produced every 20 min to simulate the usual one hour and forty minutes. As can be inferred from Table 2, about 15 min from system failure, set at 75 min, the fault is correctly diagnosed and predicted with a relative error less then 5 %.

Table 2. Summary simulation results collected every 20 min.

t_{sim} [min]	Fault(s)	t_r or t_f [min]	EÔL [min]	ε_r [%]	RÛL [min]
20	Incipient for μ_{lu}	16	-	-	-
40	Pending for a	28	81	*-8*	41
	Pending for φ_{KS}	35	48	*36*	28
	Pending for μ_{lu}	25	66	*12*	26
	Pending for μ_{fr}	23	73	*3*	33
60	Pending for a	28	75	*0*	15
	Pending for φ_{KS}	35	73	*3*	13
	Pending for μ_{lu}	25	76	*-1*	16
	Pending for μ_{fr}	23	74	*1*	14
80	Failure for a	75	75	-	< 0
	Failure for φ_{KS}	75	75	-	< 0
	Failure for μ_{lu}	75	75	-	< 0
	Failure for μ_{fr}	75	75	-	< 0

8. CONCLUSION

The results show how the knowledge produced by the CBM/PHM system is designed as complete and comprehensive of the health status of process and allows the maintainer to plan intelligently activities to maintain the plant running continuously, eliminating emergency intervention. To do this, the maintainer must receive at its headquarters the health status of all the machines within its jurisdiction. Scheduling maintenance can take place automatically using a planning software or manually directly decided by maintainer. Furthermore, testing the proposed CBM/PHM system on a real cogeneration plant with gas turbine is also in this research future plans. In the future, different intelligent techniques will be embedded into this system, in order to suggest the level of maintenance action needed before the occurrence of failures.

REFERENCES

J. Chen and R. J. Patton. *Robust model-based fault diagnosis for dynamic systems.* Kluwer Academic Publisher, Norwell (Massachusetts), 1999.

P. Feliciotti. *Modellazione e controllo di sistemi di micro cogenerazione a microturbina ed a motore a combustione interna.* PhD thesis, Universit Politecnica delle Marche, Ancona (Italia), 2009.

R. Isermann and P. Ball. *Trends in the applications of model-based fault detection and diagnosis of techincal processes.* I.F.A.C. Safeprocess Techincal Comittee's, 1997.

K. A. Marko, J. V. James, T. M. Feldkamp, C. V. Puskorius, J. A. Feldkamp, and D. Roller. Applications of neural networks to the construction of virtual sensors and model-based diagnostics. In *Proceedings of ISATA 29th International Symposium on Automotive Technology and Automation*, Firenze (Italia), 1996.

V. A. Skormin, J. Apone, and J. J. Dunphy. *On-line diagnostics of a self-contained flight actuator.* IEEE Transactions on aerospace and electronic systems, 1994.

G. Vachtsevanos, F. Lewis, M. Roemer, A. Hess, and B. Wu. *Intelligent fault diagnosis and prognosis for engineering ststems.* John Wiley & sons, inc., Hoboken (New Jersey), 2006.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Periodic Linear Time-Varying System Norm Estimation Using Running Finite Time Horizon Transfer Operators

Przemyslaw Orlowski*

** West Pomeranian University of Technology, Szczecin, Department of Control and Measurements, Sikorskiego 37, 70-313 Szczecin, Poland, (Tel. +48 91 4495409, e-mail: orzel@zut.edu.pl).*

Abstract: A novel method for norm estimation for dynamical linear time-varying systems is developed. The method involves operators description of the system model i.e. transfer operator. The transfer operator defined for finite time horizon can be described by finite dimensional matrix whereas for infinite time horizon the operator is infinite dimensional. The norm estimate for infinite time horizon is based on analysis of a running series of the finite time horizon norm properties.

Keywords: norm estimation, discrete-time systems, time-varying systems, non-stationary systems.

1. INTRODUCTION

In order to describe the dynamics of time-varying discrete-time systems, one can employ state space equations with time-dependent matrices given by eq. (1)-(2):

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{v}(k)\,, \tag{1}$$

$$\mathbf{y}(k) = \mathbf{C}(k)\mathbf{x}(k) + \mathbf{D}(k)\mathbf{v}(k)\,,\ \mathbf{x}(k_0) = \mathbf{x}_0 \tag{2}$$

where $\mathbf{x}(k) \in \mathbb{R}^n$ is nominal state, $\mathbf{v}(k) \in \mathbb{R}^m$ is the nominal control, $\mathbf{y}(k) \in \mathbb{R}^p$ is the nominal output and $\mathbf{A}(k) \in \mathbb{R}^{n\times n}$, $\mathbf{B}(k) \in \mathbb{R}^{n\times m}$, $\mathbf{C}(k) \in \mathbb{R}^{p\times n}$, $\mathbf{D}(k) \in \mathbb{R}^{p\times m}$ are system matrices, $k = k_0, k_0+1, ..., k_0+N$ and N is length of the time horizon. For infinite time horizon $N = \infty$.

An LTV system can be equivalently described in terms of the matrix operators. There are two different approaches: one based on block diagonal operators Khalil (1996) and the other based on a lower triangular system matrix Orlowski (2004). Both approaches lead to an operator-based description of the system and a function which takes the role of a transfer function for time-varying systems. This function has many properties analogous to those of transfer functions of linear time-invariant (LTI) systems. In some cases, this allows one to apply to linear time-varying (LTV) systems techniques which have formerly been restricted to LTI systems.

Alternatively, the model may be described by means of operators. Equations (1)-(2) can be converted into following operators form:

$$\hat{\mathbf{y}} = \hat{\mathbf{C}}\hat{\mathbf{N}}\mathbf{x}_0 + \left(\hat{\mathbf{C}}\hat{\mathbf{L}}\hat{\mathbf{B}} + \hat{\mathbf{D}}\right)\hat{\mathbf{v}} = \hat{\mathbf{C}}\hat{\mathbf{N}}\mathbf{x}_0 + \hat{\mathbf{T}}\hat{\mathbf{v}} \tag{3}$$

In order that the system (3) be equivalent to the system (1)-(2), operators $\hat{\mathbf{T}} = \hat{\mathbf{C}}\hat{\mathbf{L}}\hat{\mathbf{B}} + \hat{\mathbf{D}}$ and $\hat{\mathbf{C}}\hat{\mathbf{N}}$ must be defined in one of the two equivalent notations: either an evolutionary one, where operators are written by means of sums and products Orlowski (2001):

$$\begin{aligned}\mathbf{y}(k) &= \left(\hat{\mathbf{C}}\hat{\mathbf{N}}\mathbf{x}_0\right)(k) + \left(\hat{\mathbf{C}}\hat{\mathbf{L}}\hat{\mathbf{B}}\hat{\mathbf{v}}\right)(k) + \mathbf{D}(k)\mathbf{v}(k) = \mathbf{C}(k)\phi_{k_0}^{k-1}\mathbf{x}_0 \\ &+\mathbf{C}(k)\left(\sum_{i=k_0}^{k-2}\phi_{i+1}^{k-1}\mathbf{B}(i)\mathbf{v}(i) + \mathbf{B}(k-1)\mathbf{v}(k-1)\right) + \mathbf{D}(k)\mathbf{v}(k)\end{aligned} \tag{4}$$

where $\phi_i^k = \mathbf{A}(k)\mathbf{A}(k-1)\ldots\mathbf{A}(i)$, or a matrix-based one, where each of the operators can be presented in terms of matrices. In order to analyze the stability of the system, one has to know operators $\hat{\mathbf{T}}$ and $\hat{\mathbf{N}}$ which can be expressed with the help of the following matrix operators:

$$\hat{\mathbf{L}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \phi_{k_0+1}^{k_0+1} & \mathbf{I} & \mathbf{0} & \vdots \\ \vdots & \ddots & \mathbf{I} & \mathbf{0} \\ \phi_{k_0+1}^{k_0+N-1} & \cdots & \phi_{k_0+N-1}^{k_0+N-1} & \mathbf{I} \end{bmatrix} \tag{5}$$

$$\hat{\mathbf{N}} = \begin{bmatrix} \phi_{k_0}^{k_0} \\ \vdots \\ \phi_{k_0}^{k_0+N-1} \end{bmatrix} \tag{6}$$

$$\hat{\mathbf{C}} = \begin{bmatrix} \mathbf{C}(k_0) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}(k_0+N-1) \end{bmatrix} \tag{7}$$

Operator $\hat{\mathbf{N}}$ can be neglected when initial conditions are zero. Following sequences: state $\hat{\mathbf{x}}$, output $\hat{\mathbf{y}}$ and input $\hat{\mathbf{v}}$ are constructed from state $\mathbf{x}(k)$, output $\mathbf{y}(k)$ and input $\mathbf{v}(k)$ signals rewritten in following block column vector form:

$$\hat{\mathbf{x}} = \left[\mathbf{x}^T(k_0+1)\cdots\mathbf{x}^T(k_0+N)\right]^T \tag{8}$$

$$\hat{\mathbf{y}} = \left[\mathbf{y}^T(k_0+1)\cdots\mathbf{y}^T(k_0+N)\right]^T \tag{9}$$

$$\hat{\mathbf{v}} = \left[\mathbf{v}^T \left(k_0 + 1\right) \cdots \mathbf{v}^T \left(k_0 + N\right) \right]^T \tag{10}$$

The input/output operator $\hat{\mathbf{T}}$ can be alternatively defined also using a set of impulse responses of a time-varying system taken at different times, e.g. for SISO system it may be written:

$$\hat{\mathbf{T}} = \begin{bmatrix} h_{k_0}^{k_0} & 0 & \cdots & 0 & 0 \\ h_{k_0}^{k_0+1} & h_{k_0+1}^{k_0+1} & \cdots & \vdots & \vdots \\ h_{k_0}^{k_0+2} & h_{k_0+1}^{k_0+2} & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & h_{k_0+N-2}^{k_0+N-2} & 0 \\ h_{k_0}^{k_0+N-1} & \cdots & \cdots & h_{k_0+N-2}^{k_0+N-1} & h_{k_0+N-1}^{k_0+N-1} \end{bmatrix} \tag{11}$$

where $h_{k_0}^{k_1}$ is the response of the system to the Kronecker delta $\delta\left(k - k_0\right)$ at time k_1 (after k_1-k_0 samples). In the case of a nonzero input-output delay operator, $\hat{\mathbf{D}} = \mathbf{0}$ and all diagonal entries of $\hat{\mathbf{T}}$ are equal to zero.

For further considerations in the paper following definitions of norms for sequences and operators are used. The norm of a sequence in the Hilbert-space is understood as Euclidean norm:

$$\left\|\hat{\mathbf{v}}\right\| = \left\|\hat{\mathbf{v}}\right\|_2 = \sqrt{\left\langle \hat{\mathbf{v}}, \hat{\mathbf{v}} \right\rangle} = \sum_k \mathbf{v}^{\mathrm{T}}\left(k\right)\mathbf{v}\left(k\right) = \hat{\mathbf{v}}^{\mathrm{T}}\hat{\mathbf{v}} \tag{12}$$

The ∞-norm of a sequence in the bounded sequences space is understood as:

$$\left\|\hat{\mathbf{v}}\right\|_\infty = \max_k \left(\left|v\left(k\right)\right|\right) \tag{13}$$

Norms of operators are defined in following way:

$$\left\|\hat{\mathbf{T}}\right\| = \left\|\hat{\mathbf{T}}\right\|_2 = \sup_{\hat{\mathbf{v}} \neq 0} \frac{\left\|\hat{\mathbf{T}}\hat{\mathbf{v}}\right\|_2}{\left\|\hat{\mathbf{v}}\right\|_2} \tag{14}$$

For systems defined on finite time horizon all operators are represented by finite dimensional matrices and signals by finite dimensional vectors. Moreover the input-output operator is a compact, Hilbert-Schmidt operator from l_2 into l_2 and actually maps bounded signals $v \in \mathcal{M} = l_2\left[k_0, k_0 + N\right]$ into the signals $y \in \mathcal{P}$.

2. COMPUTATION THE NORM OF THE TIME-VARYING SYSTEM

Stability and performance criteria for analysis and robust control design of linear systems, are often expressed by norms of appropriately defined transfer functions or transfer operators, especially for time varying systems. Norms of the linear time-invariant systems defined on infinite time horizon can be easily computed using algorithms described in Bruisma et al. (1990), Bryson et al. (1975). The algorithms are also implemented in Matlab Control Toolbox Trefethen (2000). They needs only conversion of the system operator into state-space description. Although many methods for computing norms for linear time-invariant systems Boyd et al. (1990), Bruisma et al. (1990), Genin et al. (2002) which are essential in a computer aided control system design Zhou et al. (1995) there are very difficult to find methods applicable for linear time-varying systems.

Norm of transfer operator defined on infinite time horizon can be computed for periodic linear time-varying systems employing lifting technique. The paper (Bittanti et. al. 2000) is an overview and comparison of techniques which allows to rewrite time-varying systems using time-invariant representation with increased but finite dimensions. Norm of the transfer operator for such system can be computed in similar way as for linear time-invariant systems. More description for the lifting technique for periodic time-varying systems can be found in Bamieh et. al. (1991), Flamm (1991), Laub (1981), Meyer et al. (1975), Varga (1989).

Nevertheless norm of systems non periodic time varying systems cannot be easily computed. In such case the norm of transfer operator can be estimated using general operator theory Baladi et al. (1995), Descombes et al. (1999), Dewilde et al. (1993), Gohberg et al. (1984), Leblond et al. (1998) or the technique based on parameterised functional minimization. The main idea is based on the following general result given inOrlowski et al. (1999).

2.1. Parameterised functional based norm estimation

Theorem 1. *Let $\mathcal{M}, \mathcal{P}$ be real Hilbert spaces, $\hat{\mathbf{T}} \in \mathcal{L}\left(\mathcal{M}, \mathcal{P}\right)$, $\hat{\mathbf{C}}\hat{\mathbf{N}} \in \mathcal{P}$, $\gamma \in \left(0, \infty\right)$ and $J\left(\hat{\mathbf{v}}\right)$ be a functional defined on $\mathcal{M}$ and given by*

$$J(\hat{\mathbf{v}}) = \left\|\hat{\mathbf{T}}\hat{\mathbf{v}} + \hat{\mathbf{C}}\hat{\mathbf{N}}\right\|_{\mathcal{P}}^2 - \gamma^2 \left\|\hat{\mathbf{v}}\right\|_{\mathcal{M}}^2 \tag{15}$$

(a) $\left\|\hat{\mathbf{T}}\right\| < \gamma$ if and only if there exists $\beta > 0$, such that

$$\left\|\hat{\mathbf{T}}\hat{\mathbf{v}}\right\|_{\mathcal{P}}^2 - \gamma^2 \left\|\hat{\mathbf{v}}\right\|_{\mathcal{M}}^2 \leq -\beta \left\|\hat{\mathbf{v}}\right\|_{\mathcal{M}}^2 \qquad \forall \hat{\mathbf{v}} \in \mathcal{M} \tag{16}$$

Consequently, if $\left\|\hat{\mathbf{T}}\right\| < \gamma$, then (15) *always achieves a unique finite maximum over $\mathcal{M}$.*

(b) If $\left\|\hat{\mathbf{T}}\right\| > \gamma$ then (15) *does not achieve a finite maximum over $\mathcal{M}$, i.e. $\sup_{\hat{\mathbf{v}} \in \mathcal{M}} J\left(\hat{\mathbf{v}}\right) = +\infty$.*

It mean that $\left\|\hat{\mathbf{T}}\right\| = \inf\gamma$ over all γ such that the maximization of (15) has a finite solution. The required value of γ can be found with arbitrary accuracy, e.g. by means of the bisection method. Equivalence between the maximization of the functional (15) and the existence of a solution to the corresponding Riccati difference equations can be exploited.

Estimation of the operator norm using the method of parameterised functional minimization in general can takes large computational power.

2.2. *Running finite time horizon based norm estimation*

In order to make computationally efficient norm estimation, following approach based of finite-time horizon norm is proposed.

Definition 1. Amplification energy factor k_e for system with zero initial condition $\mathbf{x}_0=\mathbf{0}$ is given in following way

$$k_e = \frac{\|\hat{\mathbf{y}}\|}{\|\hat{\mathbf{v}}\|} = \sqrt{\frac{\hat{\mathbf{y}}^{\mathrm{T}}\hat{\mathbf{y}}}{\hat{\mathbf{v}}^{\mathrm{T}}\hat{\mathbf{v}}}} = \sqrt{\frac{\sum_{i=1}^{N} y^2(i)}{\sum_{i=1}^{N} v^2(i)}} \tag{17}$$

For systems unstable in the input-output sense output energy grows unboundedly for bounded input signals, i.e. $\sup_{\hat{\mathbf{v}}\neq 0}(k_e) = \infty$. It implies infinite value of the norm of transfer operator, i.e.

$$\left\|\hat{\mathbf{T}}\right\| = \sup_{\hat{\mathbf{v}}\neq 0}(k_e) \tag{18}$$

where the norm $\left\|\hat{\mathbf{T}}\right\| \to \infty$.

For systems stable in the input-output sense output energy is bonded for bounded input signals, i.e. $0 \le k_e < \infty$. It implies finite value of the norm of transfer operator $\left\|\hat{\mathbf{T}}\right\|$.

Let us assume that a system defined on infinite time horizon will be considered as a system defined on finite time horizon with length N. The norm of transfer operator of the system defined on finite time horizon N be denoted in following way:

$$\left\|\hat{\mathbf{T}}_{[N]}\right\| \tag{19}$$

where

$$\underset{N\in\mathbb{Z}}{\forall} \quad \left\|\hat{\mathbf{T}}_{[N-1]}\right\| \le \left\|\hat{\mathbf{T}}_{[N]}\right\| \tag{20}$$

If the norm of transfer operator defined on infinite time horizon is finite $\left\|\hat{\mathbf{T}}\right\| = c$ then there exist a limit c such that:

$$\lim_{N\to\infty}\left\|\hat{\mathbf{T}}_{[N]}\right\| = c \tag{21}$$

Thus for large enough lengths of the time horizon it may be concluded that finite time horizon norm is an approximation of the infinite time horizon norm, i.e.:

$$\underset{N\ge N_0}{\forall} \left\|\hat{\mathbf{T}}_{[N]}\right\| \cong \left\|\hat{\mathbf{T}}\right\| \tag{22}$$

Relative approximation error can be expressed by following equation:

$$\delta\left(\hat{\mathbf{T}}, N\right) = \left| \frac{\left\|\hat{\mathbf{T}}_{[N]}\right\|}{\left\|\hat{\mathbf{T}}\right\|} - 1 \right| \tag{23}$$

Although it is impossible to find simple relation between the relative error δ and the length of the time horizon N for general time-varying system $\hat{\mathbf{T}}$, we show that the method is relatively simple and efficient for discrete-time, time-varying systems norm estimation.

3. NUMERICAL ANALYSIS FOR PERIODIC TIME-VARYING SYSTEM

The system under consideration is special case of the linear time-varying system whereas $\mathbf{A}(k)$ is the time-varying system matrix with invariant eigenvalues. The system is characterized by constant (time-invariant) eigenvalues of the system matrix despite changes in its entries. This idea is borrowed from De La Sen (2002), Khalil (1996). The additional parameter ε allows changes of the system with a degree of non-stationarity as well as the pole location. Eigenvalues of matrix $\mathbf{A}(k)$ are inside the unitary circle, but can be either stable or unstable with respect to switching in the structure of the system. The deciding factor is the switching interval defined by the parameter ε. System matrices (1)-(2) are the following:

$$\mathbf{A}(k) = \mathbf{A}_{\kappa}, \quad \mathbf{B}(k) = \begin{bmatrix} 1 & 0 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{C}(k) = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \mathbf{D}(k) = 0 \tag{24}$$

where

$$\begin{aligned} &\mathbf{A}_0 = \begin{bmatrix} 2 & 1.2 \\ -2 & -1 \end{bmatrix}, \ \mathbf{A}_1 = \begin{bmatrix} -1 & -2 \\ 1.2 & 2 \end{bmatrix}, \ \mathbf{A}_2 = \begin{bmatrix} -1 & 1.2 \\ -2 & 2 \end{bmatrix}, \\ &\mathbf{A}_3 = \begin{bmatrix} 2 & -2 \\ 1.2 & -1 \end{bmatrix}, \ \kappa = \mathrm{floor}\left(\mathrm{rem}\left(\frac{k}{\varepsilon}, 4 \right) \right) \end{aligned} \tag{25}$$

Variable κ denotes rounding towards negative infinity (floor) of the remanent (signed remainder of k/ε after division by 4). Eigenvalues of the matrix $\mathbf{A}(k)$ are independent of the parameter ε and equal to $0.5 \pm 0.3873\mathrm{i}$ for all k.

In fact value of the parameter ε significantly changes properties of the system. Small values $\varepsilon<2.8$ implies unstable character of the system whereas large values results in stable, switching system. Figure 1 shows values of the transfer operator norm $\left\|\hat{\mathbf{T}}_{[N]}\right\|$ vs. length of the time horizon N for $\varepsilon = 5$. Value estimated using lifting techniques is equal to $\left\|\hat{\mathbf{T}}\right\| = 12.9849$ and depicted by dotted line. As can be seen

from fig. 1 estimated norm fast reach neighbourhood of the real value. It takes only about 27 time steps.

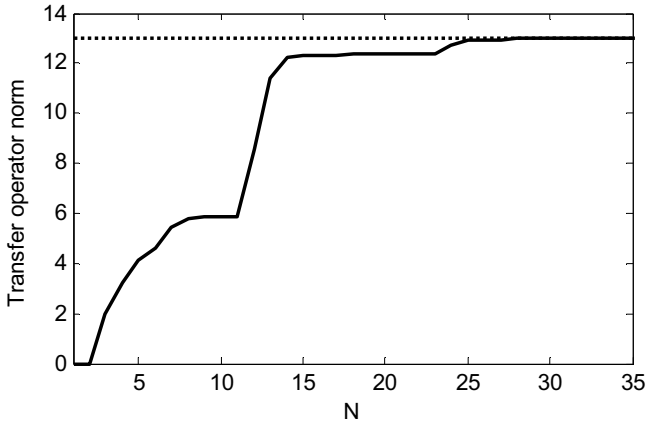


Figure 1. Norm of transfer operator for finite time horizon discrete switching system (24)-(25) with $\varepsilon = 5$ vs. the length of the time horizon *N*.

Relative error for the same system computed for the length of the time horizons up to 500 is depicted in fig. 2. From practical point of view relative error for norm estimation below 10^{-2} is in most cases sufficient, in this case it takes only 27 time steps what is relatively fast, even for second order system but with variability period of $4\varepsilon = 20$ time steps.

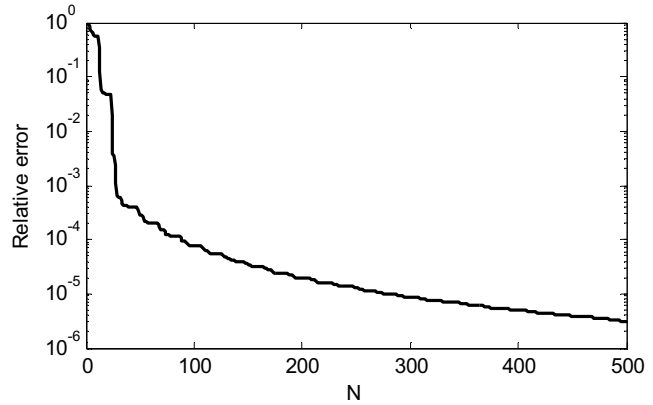


Figure 2. Relative error of the transfer operator norm computed on finite time horizon for discrete switching system (24)-(25) with $\varepsilon = 5$ vs. the length of the time horizon *N*.

4. CONCLUSION

In the paper a novel approach for the estimation of the operator norm is proposed. Particularly infinite dimensional transfer operator norm of dynamical discrete-time, periodical time-varying stable systems can be estimated using block matrix operator notation for transfer operator defined on finite time horizon. The minimal length of the time horizon required for computations is dependent both on the dominant time constant of the system and the variability period of the system matrices.

Open problems are connected mostly with estimating the minimal length of the time horizon required for computations. Further investigations should concern extending the method for wider class of the time-varying systems e.g. for other common classes, i.e. almost periodic systems etc.

REFERENCES

[1] Baladi V., Isola S., Schmitt B. (1995). Transfer operator for piecewise affine approximations of interval maps. *Annales de l'IHP Physique théorique*.

[2] Bamieh B., Pearson J.B., Francis B.A., Tannenbaum A. (1991). A lifting technique for linear periodic systems with applications to sampled-data control, *Systems & Control Letters*, Vol. 17(2), pp. 79-88.

[3] Bittanti S, Colaneri P (2000). Invariant representations of discrete-time periodic systems. *Automatica* vol. 36, 1777-1793.

[4] Boyd S., Balakrishnan V. (1990). A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L1-norm, *Systems & Control Letters*, vol.15, pp.1–7.

[5] Bruisma, N.A., Steinbuch M. (1990). A Fast Algorithm to Compute the H∞-Norm of a Transfer Function Matrix, *System Control Letters*, vol. 14, pp. 287-293.

[6] Bryson, A.E., Ho Y.C. (1975). *Applied Optimal Control*, Hemisphere Publishing, pp. 458-459.

[7] De La Sen M. (2002). Robust stability of a class of linear time-varying systems, *IMA Journal of Mathematical Control and Information* 19, 399-418.

[8] Descombes S., Dia B.O. (1999). An Operator-Theoretic Proof of an Estimate on the Transfer Operator. *Journal of Functional Analysis*, Vol. 165(2), pp. 240-257.

[9] Dewilde P., Van Der Veen A.J. (1993). On the Hankel-norm approximation of upper-triangular operators and matrices. *Integral Equations and Operator Theory*.

[10] Flamm D.S. (1991). A new shift-invariant representation of periodic linear systems. *Systems & Control Lett.* 17, 9–14.

[11] Genin Y., Stefan R., Van Dooren P. (2002). Real and complex stability radii of polynomial matrices, *Linear Algebra and its Applications*, vol.351-352, pp.381–410.

[12] Gohberg I., Kaashoek M.A. (1984). Time varying linear systems with boundary conditions and integral operators. I. The transfer operator and its properties. *Integral Equations and Operator Theory*, Vol. 7(3), pp. 325-391.

[13] Khalil H.K. (1996). *Nonlinear Systems (Second Edition)*. Englewood Cliffs, NJ: Prentice-Hall.

[14] Laub A. J. (1981). Efficient multivariable frequency response computations. *IEEE Trans. Autom. Control* 26, 407–408.

[15] Leblond J., Olivi M. (1998). Weighted H 2 approximation of transfer functions. *Mathematics of Control, Signals, and Systems*

[16] Meyer R.A., Burrus C.S. (1975). A unified analysis of multirate and periodically timevarying digital filters. *IEEE Trans. Circuits and Systems* 22, 162–168.

[17] Orlowski P. (2001). Applications of Discrete Evolution Operators in Time-Varying Systems. *European Control Conference*, Porto, pp. 3259-3264.

[18] Orlowski P. (2004). Selected problems of frequency analysis for time-varying discrete-time systems using singular value decomposition and discrete Fourier transform. *Journal of Sound and Vibration*, Vol. 278, pp. 903-921.

[19] Orlowski P., Emirsajlow Z. (1999). Analysis of Finite Horizon Control Problems for Uncertain Discrete-Time Systems, *IX KKA (National Automation Conference),* Opole, vol. I, pp. 107-112.

[20] Trefethen L.N. (2000). Spectral methods in MATLAB, volume 10 of Software, Environments, and Tools, *SIAM.*

[21] Varga A. (1989). Computation of transfer function matrices of generalized state-space models. *Int. J. Control* 50, 2543–2561.

[22] Zhou, K., Doyle J.C., Glover K. (1995). *Robust and optimal control*. Prentice Hall.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

An Application of Model Based Fault Detection in Power Plants

Goran Kvaščev*, Predrag Tadić*, Željko Đurović*

**Signals and Systems Department, School of Electrical Engineering, University of Belgrade Belgrade, Serbia (e-mail:kvascev@etf.bg.ac.rs)*

Abstract: Early detection of faults within subsystems of complex plants enables a timely reaction of the plant staff, thus alleviating the ensuing problems with the plant operation and possibly avoiding the need for a complete shut-down. An application of a model-based fault detection and isolation (FDI) algorithm to the superheating system of a coal-fired thermal power plant is presented in this paper. Data from a real process were used to obtain a piecewise affine model of the process, which was then used to generate a residual carrying the information of a possible fault in the system. In order to eliminate the effects of plant parameter fluctuations due to operating point changes, a residual post-filter was implemented. The results of experimental verification are presented and briefly discussed.

Keywords: System identification, Fault detection, Thermal power plants

1. INTRODUCTION

Fossil-fuel fired thermal power plants remain one of the main sources of electrical energy in many countries around the world. In Serbia, for example, more than 60% of production comes from coal-fired thermal power plants. Since most of these plants are more than 20 years old, equipment reliability is a major issue. Sensor and actuator failure lead to prolonged shut-downs, which have severe financial consequences. Having this in mind, investment in fault detection and isolation (FDI) systems becomes more attractive. If the staff is able to receive early warnings about problems that are just beginning to develop, they might have enough time to react in such a way as to prevent the need for a complete shut-down of the plant or its subsystems. The additional system start-up costs and losses due to off-line time are thus greatly reduced, increasing the overall plant efficiency.

The system under consideration here is the steam temperature control mechanism. The temperature needs to be kept around a prescribed value in order to optimize the turbine working conditions, and in order to fulfil obvious safety conditions. Control is accomplished through a process known as attemperation, which involves injecting cold water into the steam flow, thus lowering its temperature appropriately. Valves used to regulate the amount of injected water often get stuck, and this is the fault we aim to identify and isolate by using the approach described in this paper.

The variety of different types of FDI algorithms found in the literature (see, for example, Gertler(1998), Ding(2008)), can roughly be divided into two groups: model-based and data-driven. The former approach is used here: a model of the process is identified, and then run in parallel with the plant. The residuals, in the form of differences between the model and plant outputs, are used to conclude if the type of fault under consideration has actually occurred. Typically, if the residual value significantly differs from zero, a fault is supposed to have occurred. However, operating-point changes lead to plant parameter fluctuations, which in turn affect the residual, leading to false alarms. Such effects are suppressed by post-filtering the residuals, but in such a way as to preserve the desired behaviour in case of actual faults.

The rest of the paper is organized as follows. Chapter 2 describes in detail the attemperation system and its components. Identification and modelling of the plant are presented in chapter 3, while chapter 4 explains the applied FDI mechanism and presents the experimental results. Chapter 5 concludes the paper.

2. ATTEMPERATION SYSTEM

A characteristic feature of the once-through boiler is that the pumps force the feedwater/steam through the boiler tubing, which in principle is arranged a continuous pipe. The boiler process includes several steam superheating processes. Each of these processes serves as an energy transferring system-energy being transferred from the flue gas to the steam. Each superheater is equipped with an attemperator device (water injection at the inlet) for control of the steam outlet temperature, Flynn(2003), Brkic(2005), Jovanovic(1982). A superheater process is shown in Fig. 1 and Fig. 2.

The control of steam temperatures in power plants is one of the most widely discussed control problems in power plants. The reasons for the extensive attention to this problem are mainly found in issues such as:

- Lifetime of plant: The steam temperature control has a significant influence on the variation of the steam temperatures and accordingly on the thermal stress of the plant.
- Efficiency: If the steady-state variations can be reduced significantly, the outlet set-point can be

increased and the turbine efficiency will increase accordingly.

- Load-following capability
- Availability: The improved overall stability and the resulting reduced probability of forced plant outage is an indirect advantage of improving the steam temperature control. Nevertheless, it is an important advantage - e.g. a forced outage of a coalfired base-load unit will imply additional fuel costs for restart, lack of power sales, increased wear of the plant and reduced availability.

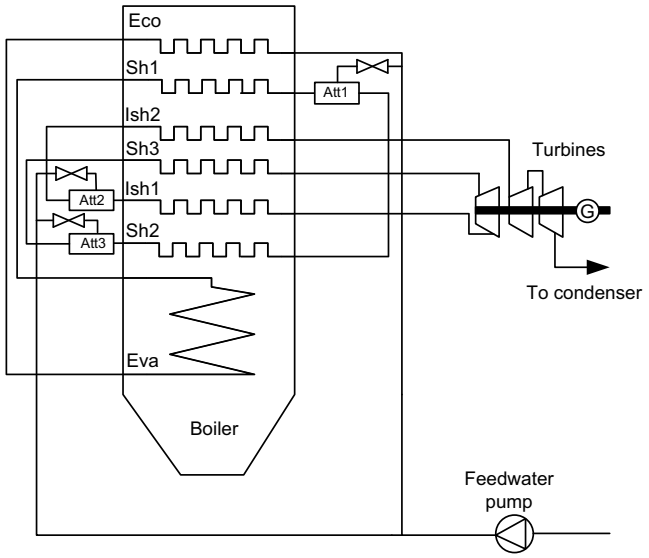


Fig. 1. Outline of steam power plant. Eco: economiser; Eva: evaporator; Sh: high-pressure superheater; Ish: intermediate pressure superheater; Att: attemperator.

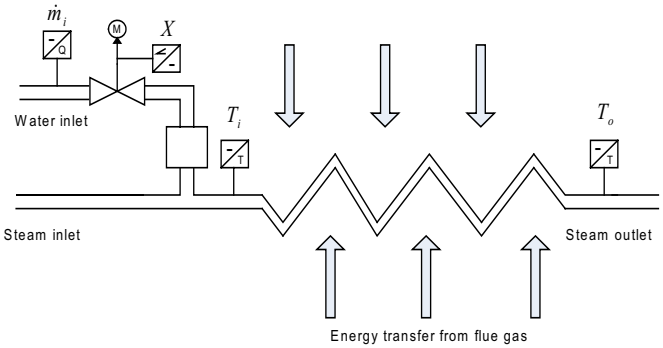


Fig. 2. Superheater process with typical instrumentation: T_o - outlet steam temperature (°C); T_i -inlet steam temperature (°C); X-valve position (%); $\dot{m}_i$ - water injection flow

First step in FDI model-based approach consists of providing a mathematical description of the system under investigation which shows all the possible faulty conditions.

3. MODELING AND IDENTIFICATION

Basic idea is how to find suitable nonlinear model for real processes, with the aim to provide a general treatment of the problem. One solution is to focus attention on the nonlinear black-box parametric models.

The main idea underlying the mathematical description of nonlinear dynamic systems is based on the interpretation of single input single output, nonlinear, time invariant regression models such as

$$y(t+n) = F(y(t+n-1),\ldots,y(t),u(t+n-1),\ldots,u(t)) \quad (1)$$

where u(t) and y(t) are input and output of process, respectively, and F(·) is continuous nonlinear function that represents functional relation from inputs to outputs.

The identification of the nonlinear system can be translated to the approximation of the mathematical model given previous equation using a parametric structure that exhibits arbitrary accuracy interpolation properties.

A piecewise model defined through the composition of simple models having local validity is the natural candidate to perform this task, as it combines function interpolation properties with mathematical tractability. The piecewise SISO model is formed by a collection of parametric submodels of the type:

$$y(t+n) = \sum_{j=0}^{n-1} -a_j^{(i)} y(t+j) + \sum_{j=0}^{n-1} b_j^{(i)} u(t+j) + y_0^{(i)} \quad (2)$$

in which the system operating point is described by the input and output samples $y(t+n-1),\ldots,y(t)$ and $u(t+n-1)$, that can be collected with vector $x_n(t)=[y(t),\ldots,y(t+n-1),u(t),\ldots,u(t+n-1)]^T$. The switching function is

$$\chi(x_n(t)) = \begin{cases} \chi(x_n(t)) = 1 \quad if \quad x_n(t)\ in\ Range_n^{(i)} \\ \chi(x_n(t)) = 0 \quad otherwise \end{cases} \quad (3)$$

First order model for proper describing of valves in attemperation subsystem is given with

$$y(t+1) = \sum_{i=1}^{7} \chi_i(x(t))[-a_1^{(i)} y(t) + b_1^{(i)} u(t) + y_0^{(i)}] \quad (4)$$

In attemperation subsystem, water injects through controlled servo valve and nozzle in live (fresh) steam. Measured values, also inputs in fault detection system are:

- Valve position (u(t)=X) in %
- Water injection flow (y(t)=m;˙i) in tons/hour

For experimental verification of FDI techniques two attemperation subsystems are chosen:

- Servo valve NB12C101 in main pipeline (last water injection in right side of armature)
- Servo valve NB10C101 in backup pipeline (last water injection in left side of armature)

Different valves with different constructions and servo controllers were used to demonstrate system modelling, identification and FDI principles.
For parameters identification of piecewise affine model following procedure was carried out:

- System is switched in manual regime
- Control signal was sequence of step signals in range of 0-100%
- Parameters for seven different affine models were obtained using LS method,Ljung(1987), Bosch(1994), and shown in Table 1., for valve NB12C101

Table 1. Estimated parameters of piecewise affine model

	Model no.	1		2		3		4		5		6		7	
		Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
u – control	Position Range	4.86	12.26	36.24	24.86	48.47	36.24	59.98	48.47	71.84	60.15	83.35	71.86	96.00	83.35
y - measurement	Flow range	3.51	2.54	5.83	3.51	7.53	5.84	9.16	7.53	11.49	9.17	15.47	11.48	21.01	15.50
Model parameters	a1	-0.694		-0.699		-0.769		-0.81		-0.796		-0.787		-0.805	
	b1	0.0241		0.0558		0.0316		0.0265		0.0395		0.071		0.083	
	y0	0.474		-0.287		0.204		0.152		-0.506		-2.627		-3.923	
Td (estimated)		2.74		2.798		4.6521		4.7466		4.4056		4.1866		4.632	
K (estimated)		0.0788		0.1858		0.47		0.1397		0.1945		0.3341		0.430	

Results of system identification are shown in Fig. 3 and 4. After identification, output error between process and model was formed in manner to check goodness of identification procedure. Autocorrelation function of this residual was presented in Fig. 5, and it was shown that identification error is almost white noise. Because of that, we can conclude that proposed model can properly describe process.

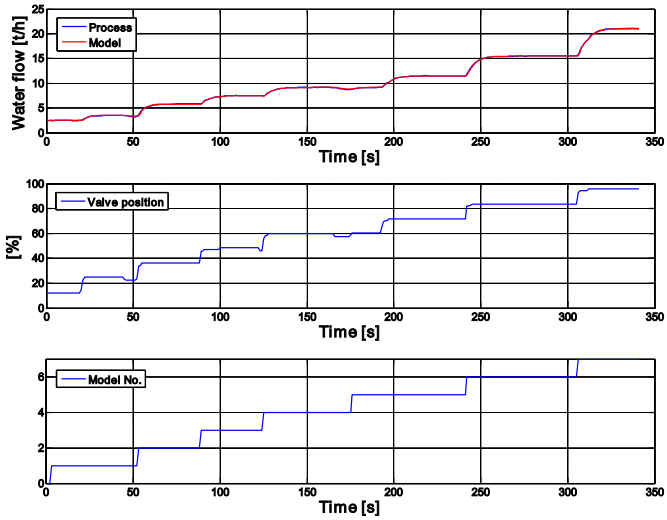


Fig. 3. System identification for input sequence in openloop experiment (Plant and model – first subfigure, position of valve – second subfigure, Model number used in piecewise affine structure – last subfigure)

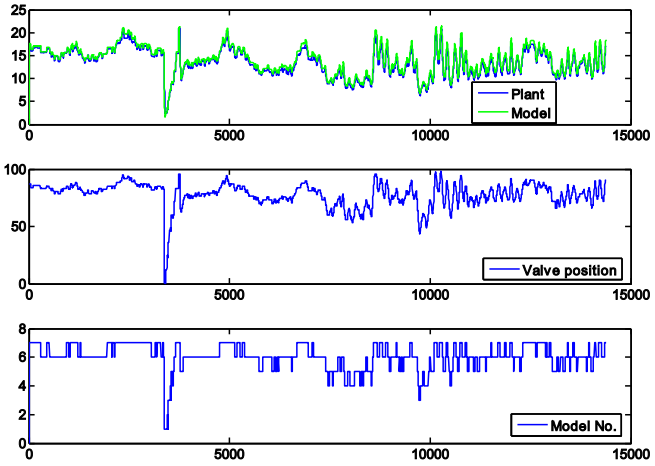


Fig. 4. Plant and model, valve position, model number used in piecewise affine structure

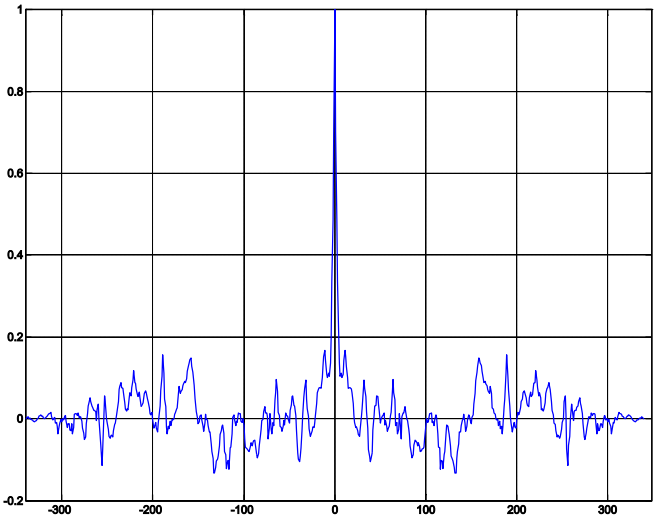


Fig. 5. Autocorrelation function of identification residual signal

4. MODEL BASED FAULT DETECTION APPLIED TO THE ATTEMPERATION SYSTEM

Most of the residual generation techniques are based on discrete system models Isermann(1997), Paton(2000). The basic idea of the parity relations approach is to provide a proper check of the parity of the measurements acquired from the monitored system.

4.1 Output error

A straightforward model-based method of fault detection is to take a model and run it in parallel to the process, thereby an output error vector is obtained

$$r_{oe}(k) = e(k) = y(k) - \frac{\hat{B}(z)}{\hat{A}(z)} \tag{5}$$

The methodology is depicted in following Fig. 6.

When a fault occurs in the plant, the residual r(t) will be different from zero.

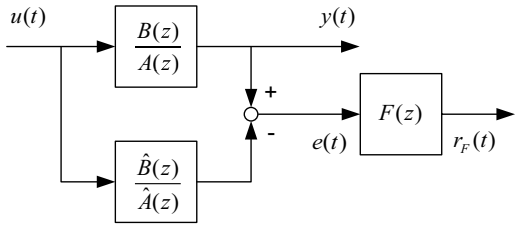


Fig. 6. Scheme for output error via parity equation method

4.2 Experimental results

After modelling and identification, residual generator was applied for two attemperation systems (NB10C101, NB12C101) (Fig. 7 and 8, respectively). It is remarkable that there is a time-varying mean-value of the obtained residual signal, even during time intervals without any faults. This phenomenon may be explained by the fact that the parameter

identification has been done on nominal power setpoint. After power setpoint change the nature of residual signal is changed as well being nonzero without any fault presence. Offset in residual is caused by changing pressure in water injection nozzle, due to feedwater setpoint changing in function of desired electric power of plant. Consequently some corrections in the residual generator structure have to be done.

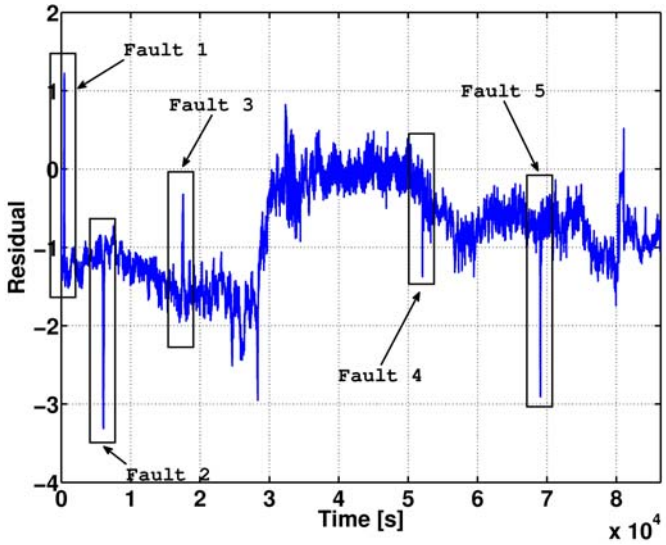


Fig. 7. Output error residual for valve NB10C101, left armature, Fault1-Actuator is blocked (time~400s), Fault2, Fault3, Fault4, Fault5 - Data acquisition system fault (time~6000s,17000s, 52000s and 69000).

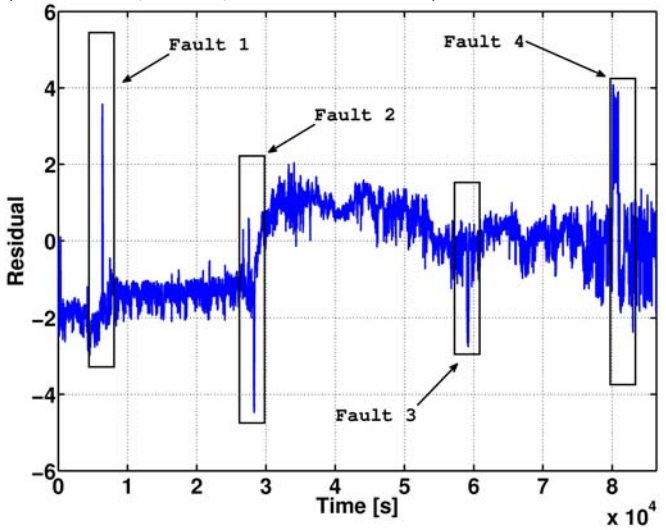


Fig. 8. Output error residual for valve NB12C101, right armature, Fault1, Fault3 - Data acquisition system fault (time~6000s, 59000s), Fault2, Fault4 - Actuator is blocked (time~28000s, 83000s).

In attemperation subsystem there are several major faults:

1. Sensors faults
   - Sensor failure – sensor output signal is not in range 4-20mA. This fault is not considered because it is detected with comparing signal with valid signal range
   - Sensor offset or saturation is failure due to wrong selection of type of measurements or due to changes in measurement environment, set-point or conditions.
2. Actuator faults
   - A typical valve malfunction is a position block or jamming, which occurs when the actuator no longer responds to the control signals sent by the regulator. The position of the actuator may be locked in three different ways: in the same position as when the fault happens, completely open, and completely closed.
3. Data acquisition system fault
   - Input data are frozen and control system has information from last valid measuring without notice of bad measuring.

In Fig. 7 and 8, there are output errors presented, on date February, 16th 2009, for control valves NB10C101 and NB12C101, left and right armature respectively.
Faults are clearly visible, but we cannot set the threshold that would be able to process failure detection, because the residual has a mean value due to electric power setpoint changes from 205MW to 320MW. To eliminate this deficiency residual post filtering was carried out.

4.3 Residual post- filtering

Residual itself has two components of the signal of interest. The first comes from failure that occur in the system, and the other is due to changes in the operating regime, which depends on the given power setpoint of the plant. As the nature of these two signals is different in frequency domain, it is possible to carry out a post filter design that will make the removal of the signal originating from changes in setpoint, and thus eliminate the variables mean signal, which is located at low frequencies.

Post-filter design is done in order to maximize the speed of elimination low frequency component signal and pass-through hi frequency signal that comes from failure. The results are shown in Fig. 9 and 10, which are filtered residual, for a one day measurement.

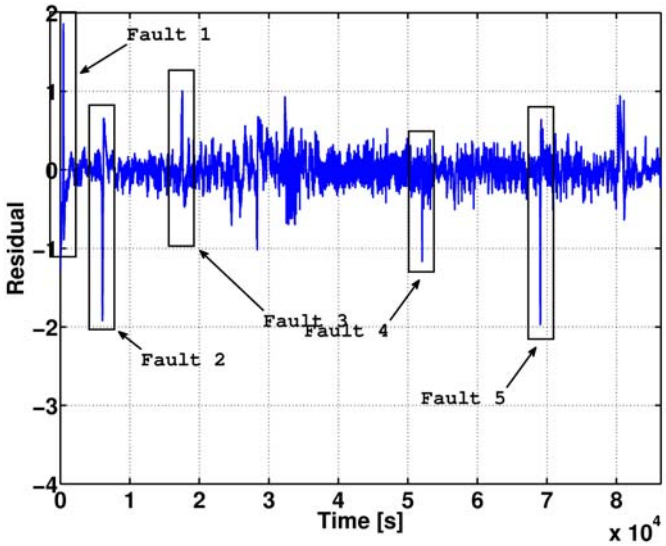


Fig. 9. Residual post filtering for valve NB10C101, left armature, Fault1-Actuator is blocked (time~400s), Fault2, Fault3, Fault4, Fault5 - Data acquisition system fault (time~6000s,17000s, 52000s and 69000).

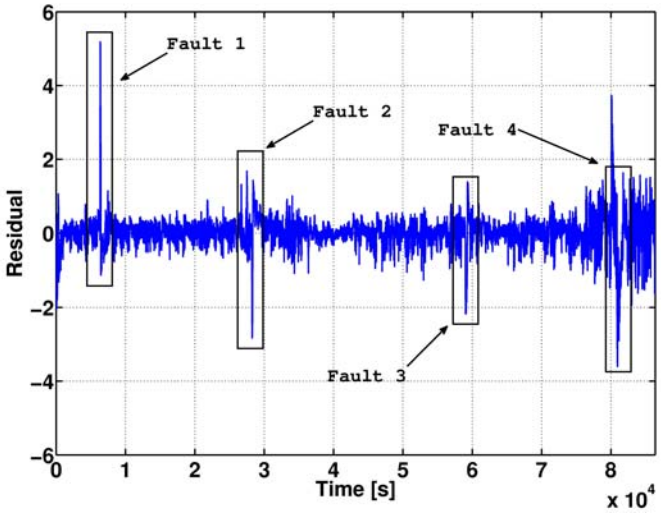


Fig. 10. Residual post filtering for valve NB12C101, right armature. Fault1, Fault3 - Data acquisition system fault (time~6000s, 59000s), Fault2, Fault4 - Actuator is blocked (time~28000s, 83000s).

It is now possible to set a constant threshold that will perform failure detection with great reliability. Because modelling of the attemperator system is not ideal and a simple method of detection failure is used, there is possibility of false failure detection.

5. CONCLUSION

The paper shows that a relatively straightforward FDI procedure can be used to effectively identify problems with the attemperation valve operation. A simple step-response based LS identification method was used to produce a piecewise-affine first order model of the plant, which successfully captures the characteristics of plant behaviour relevant to FDI purposes. The effects of operating-point changes were suppressed by using a high-pass residual post-filter, while preserving the desired performance with respect to the actual fault. The results obtained with data taken from a real-world process were satisfactory, considering the relative simplicity of the algorithm.

6. ACKNOWLEDGMENTS

This work is supported by the European Commission's Seventh Framework Programme, as part of the PRODI project (INFSO-ICT-224233).

REFERENCES

Ljung, L.(1987). *System Identification, Theory for the user*, New Jersey: Englewood Cliffs, Prentice Hall

Bosch, P.P.J. van den, Klauw, A.C. van der (1994). *Modeling, Identification and Simulation of Dynamical Systems*, Boca Raton: CRC Press.

Gertler, J. (1998), *Fault Detection and Diagnosis in Engineering Systems*, Marcel Dekker.

Ding, S.X. (2008). *Model-based Fault Diagnosis Techniques*, Springer-Verlag.

Flynn, D. (2003). *Thermal Power Plant Simulation and Control*, The Institution of Electrilal Engeneer, London.

Brkic, Lj., Zivanovic T. (2005). *Parni Kotlovi*, Mašinski fakultet Univerziteta u Beogradu, Beograd.

Jovanovic, M. (1982). *Pogonski Propisi za Kotao 920t/h, 186bar, 543ºC – Obrenovac 3*, Termoelektrane "Nikola Tesla" Obrenovac.

Isermann, R., Ballé P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes, *Control Engineering Practice*, 5(5):709-719.

Patton, R.J., Clark, R., Clark, R.N. (2000) *Issues of fault diagnosis for dynamic systems*. Springer-Verlag, Berlin Heidelberg, New York

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Validation of a New Time Delay Estimation Method for Control Performance Monitoring

M. Stockmann*, R. Haber*, U. Schmitz**

**Department of Process Engineering and Plant Design, Laboratory of Process Control, Cologne University of Applied Science, D-50679 Köln, Betzdorfer Str. 2, Germany. fax: +49-221-8275-2836 and e-mail: {markus.stockmann; robert.haber}@ fh-koeln.de*
*** Shell Deutschland Oil GmbH, Rheinland Raffinerie, D-50389 Wesseling, Ludwigshafener Strasse 1, e-mail: ulrich.schmitz@shell.com.*

Abstract: Time delay estimation is a very important topic for control performance monitoring, as it is required for several performance indices (e.g. Harris index) and for fault propagation analysis in the presence of plantwide faults. Many classical methods for time delay estimation are only valid for linear Single Input Single Output (SISO) processes or require advanced model assumptions. In the present paper, a new and simple method based on k nearest neighbour imputation is validated and demonstrated on an industrial application.

Keywords: performance monitoring, time delay estimation, significance test

1. INTRODUCTION

Control performance monitoring and assessment is a very important topic as it enables plant operators to increase plant productivity and to decrease the production of rejections. The procedure of control performance monitoring and assessment has been summarised in many reputable overviews (e.g. Jelali, 2006) and had been started with the introduction of the Harris index (Harris, 1989) which objectively measures the performance of a single control loop related to the best achievable performance of a comparable minimum variance controller with the same time delay. In complex plants, single control loop faults are often propagated by production stream or heat coupling to many other control loops and cause plantwide faults. Hence, fault propagation analysis is also a very important topic of control performance monitoring and its aim is to determine the source of plantwide faults under the use of propagation path models (Bauer and Thornhill, 2008). Fault propagation analysis is often realized by time delay estimation (Bauer et al., 2004). It can be seen that time delay estimation is a basic step of control performance monitoring and requires high accuracy. A very reliable procedure was introduced by Stockmann and Haber (2010) and allows dead time estimation for both time invariant linear and nonlinear Multiple Inputs Single Output (MISO) systems. Stockmann and Haber (2010) do not give any rules about the validation of this new method. This validation is shown in the present paper and aim of it is to determine strict thresholds in order to give evidence about the quality and repeatability of time delay estimation. The paper is structured as follows: In section 2 the procedure for time delay estimation based on k nearest neighbour (*knn*) imputation is shortly presented. In section 3 the validation is shown based on significance testing and in section 4 the method is demonstrated on an industrial example.

2. TIME DELAY ESTIMATION USING K NEAREST NEIGHBOUR IMPUTATION

The idea for time delay estimation based on k nearest neighbour imputation has been introduced by Stockmann and Haber (2010). Its mode of operation is briefly explained for a static nonlinear SISO system without loss of generality (see reference for more details). Assume two signals $u(k)$ and $y(k)$. Where y is a function depending on u and a discrete time delay d, e.g. (1)

$$y(k) = 5 + u(k-d) + 0.25u^2(k-d) + 0.01u^3(k-d) \qquad (1)$$

For example, let be $d = 3$ and let u be a standard normal distributed random number. The courses of u and y are given in Fig. 1.

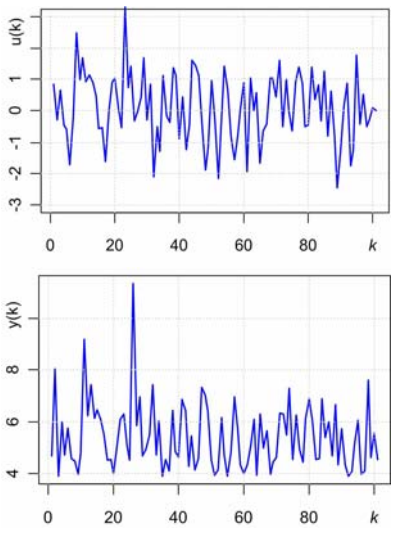


Fig. 1. Simulated course of two time series u and y fulfilling (1) with discrete time delay $d = 3$

Furthermore let $u(k = i)$ be a certain value (e.g. 1) and let n_d be an estimated time delay then $y(i+n_d)$ will have a defined value by (1) ($y(i+3) \approx 6.26$) if the estimated n_d value is chosen correctly. If the value of $u(i) = 1$ occurs again at a later sample time j, then $y(j+n_d)$ will have the same value as $y(i+n_d)$ in the absence of noise. That means for imputation theory that $y(i+n_d)$ can be expressed by the nearest neighbour of $u(i)$ only which is $u(j)$ and its corresponding output $y(j+n_d)$. This can also be visualised in the u, y scatter plot for different assumed time delays n_d shown in Fig. 2.

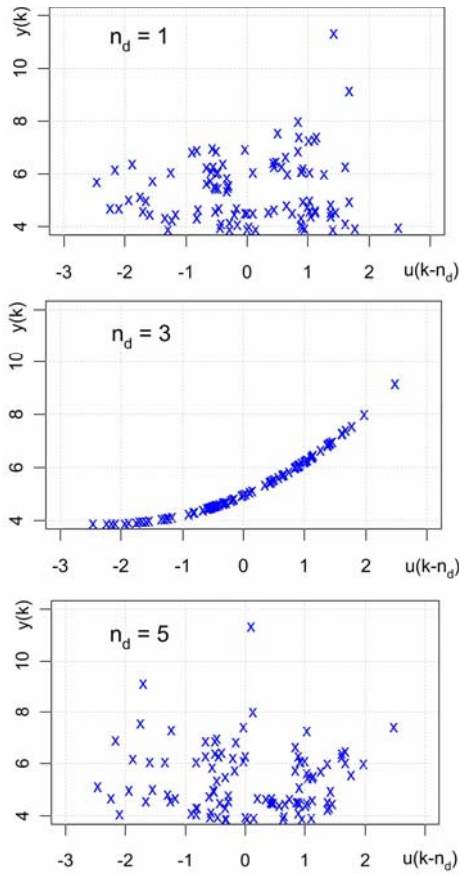


Fig. 2. Scatter plot $u(k\text{-}n_d)$ vs. $y(k)$ for different assumed dead times n_d

Theory of the automated time delay estimation based on the presented idea is to omit the output signals $y(k=1,...,N)$ for every sample time after another and to impute it ($\hat{y}(k)$) by the k_{imp} nearest input neighbours. Looking at Fig. 2, one can see clearly that if the estimated time delay n_d is chosen correctly ($n_d = d = 3$) then the imputed value will be nearly the same as the omitted value. Stockmann and Haber (2010) proposed the following ten steps to automate this procedure:

1. Let the start value of n_d be the minimal assumed time delay d_{min} minus one, then:
2. Increment n_d by one.
3. Shift the input vector u by n_d.
4. Let the counting index k and the summing variable s be equal to zero.
5. Increment k by one.
6. Delete y(k).
7. Replace $y(k)$ by a *knn* imputation $\hat{y}(k)$.
8. Calculate the deviation $s = s + \left(y(k) - \hat{y}(k)\right)^2$.
9. Replace $\hat{y}(k)$ by $y(k)$ and go to step 4 until k has reached the length of the vectors N.
10. Calculate $R_{uy}^{knn}(n_d) = \left(1 + \frac{\sqrt{s}}{N}\right)^{-1}$. Go to step 1 and repeat it until n_d has reached the maximal assumed time delay.

$R_{uy}^{knn}(n_d)$ is a normalized function which lies between 0 ($n_d \neq d$) and 1 ($n_d = d$). In Fig. 3 the course of $R_{uy}^{knn}(n_d)$ is illustrated for the simulated example shown in Fig. 1. The maximal value of $R_{uy}^{knn}(n_d)$ can be found for the assumed time delay 3, which is equal to the simulated time delay. Even in the absence of noise the maximal value of $R_{uy}^{knn}(n_d)$ is always smaller than 1. The reason for this is that the number of $\{u(k\text{-}n_d), y(k)\}$ pairs is finite and hence the imputation is always defective ($\Rightarrow s > 0$).

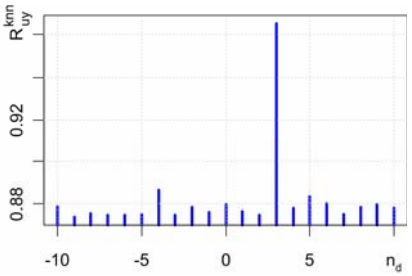


Fig. 3. $R_{uy}^{knn}(n_d)$ course for the two time series u and y shown in Fig. 1

Stockmann and Haber (2010) showed that it is also possible to determine time delays for MISO systems with different time delays for the inputs.

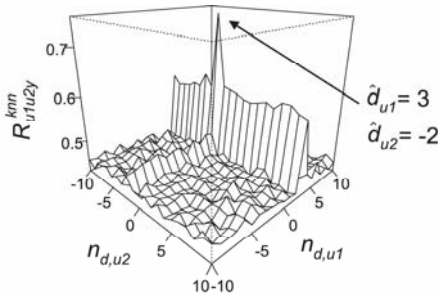


Fig. 4. Time delay estimation based on knn imputation for two inputs with $d_{u1} = 3$ and $d_{u2} = -2$

Explanation: Assume a Two Inputs Single Output (TISO) system. The ten steps of the *knn* based time delay estimation need one modification only: instead of shifting one input (u_1), two inputs have to be shifted separately. Assume first in the SISO case that $n_d = (-10, -9, ..., 9, 10)$ then $R_{uy}^{knn}(n_d)$ has the length of 21. By shifting two inputs R_{uy}^{knn} becomes a function of $n_{d,u1}$ and $n_{d,u2}$. A plot of $R_{u1u2y}^{knn}(n_{d,u1}, n_{d,u2})$ with two inputs hence becomes a three-dimensional plot, see Fig. 4. For MISO systems with *m* inputs the function $R_{uy}^{knn}(n_{d,u1}, n_{d,u2}, ..., n_{d,um})$ must be interpreted as a function of a *m*-dimensional array. The calculation time increases drastically.

3. VALIDATION VIA SIGNIFICANCE TESTING

Aim of this section is to determine strict thresholds for the quality of time delay estimation. Therefore two parameters are introduced for the mean centred series $\tilde{R}_{uy}^{knn}(n_d)$ by (2) and (3) calculated based on two scaled time series *u* and *y* :

$$\rho^{knn} = \max\left(\tilde{R}_{uy}^{knn}(n_d)\right) \tag{2}$$

$$\sigma^{knn} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M}\tilde{R}_{uy}^{knn}(n_{d,i})^2} \tag{3}$$

In the present paper, k_{imp} always equals 4 (see Stockmann and Haber (2010) for more details). In Fig. 5 the results of two simulations are shown with their corresponding $R_{uy}^{knn}(n_d)$ and $\tilde{R}_{uy}^{knn}(n_d)$courses. The question is which of them indicates a better reliable time delay estimation.

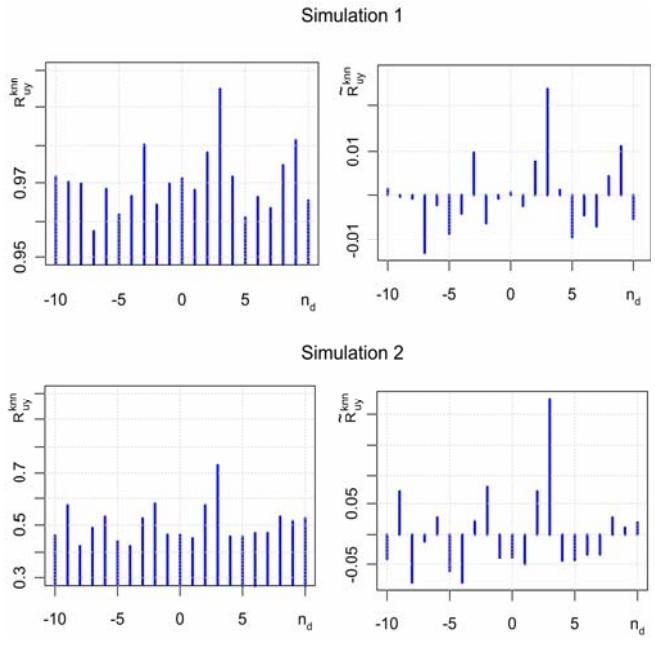


Fig. 5. $R_{uy}^{knn}(n_d)$ (left) and $\tilde{R}_{uy}^{knn}(n_d)$ course (right) for two simulation examples

At first sight, it may seem that the time delay estimation for simulation 1 is better as the $R_{uy}^{knn}(n_d = 3)$ value is closer to 1 than in the second simulation. However, this is not the case because the mean value of $R_{uy}^{knn}(n_d)$ is relatively high (0.97). The reason for this is always just a good imputation which is not necessarily connected to good time delay estimation. A good imputation may be achieved for example in the presence of a nearly constant input and a constant output signal. In this case, the imputation will be rather accurate. A good time delay estimation will always be found if the imputation for one assumed dead time is much better than the imputations for all other dead times. Hence, the quotient ε given in (4) can be used as a measure of a good quality for time delay estimation based on *knn* imputation of two scaled time series *u* and *y*:

$$\varepsilon = \rho^{knn} / \sigma^{knn} \tag{4}$$

The higher ε is, the more outstanding the time delay estimation is for one certain assumed dead time. For the two simulations shown in Fig. 5 the following parameters can be calculated:

Simulation 1:

$$\rho^{knn} = 0.0247,\ \sigma^{knn} = 0.0083 \Rightarrow \varepsilon = 2.9630$$

Simulation 2:

$$\rho^{knn} = 0.2269,\ \sigma^{knn} = 0.0667 \Rightarrow \varepsilon = 3.4020$$

Looking at these values it can be seen that the time delay estimation of the second simulation is a bit better than for the first simulation.

It seems that ε also depends on the number of measurements which were included for time delay estimation (*N*) and on the number of assumed time delays for which the imputation is performed (*M*). In the two examples shown in Fig. 5 *N* is equal to 100 and *M* equals to 21.

The conclusion about quality of time delay estimation is reduced to one sided significance testing as ε is a positive value. A comparable method for the interpretation of the cross-correlation function was proposed by Bauer and Thornhill (2008).

For the analysis of $\tilde{R}_{uy}^{knn}(n_d)$ a null hypothesis of no causal dependency based on ε is made. The null hypothesis will be kept or replaced by the alternative hypothesis (causal dependency) under the use of a certain level of significance α. For this significance testing the distribution of ε has to be known. Without deriving it theoretically, a Monte Carlo simulation showed that ε is log-normal distributed in case of no dependency between two scaled Gaussian white noise time series with mean value μ^{LN} and standard deviation σ^{LN}, see (5)

$$\varepsilon \sim LN(\mu^{LN}, \sigma^{LN}). \tag{5}$$

Explanation of the used Monte Carlo simulation

For the Monte Carlo simulation the time delay estimation of two causally independent and scaled time series, created by standard normal distributed random numbers with N = 100 and M = 21, was performed 10.000 times. ε was calculated for every single simulation. The distribution of ε is shown in Fig. 6. The parameters $\mu^{LN} = 0.7158$ and $\sigma^{LN} = 0.2404$ were empirically estimated based on the results of the Monte Carlo simulation.

With M = 21 and N = 100 the null hypothesis can be formulated as follows: *Under the use of a significance level α = 0.025 two scaled time series will be independent if the calculated ε value is smaller than 3.2772. In this case the time delay estimation based on k nearest neighbour imputation delivers no retraceable result and can not be used for fault propagation analysis.*

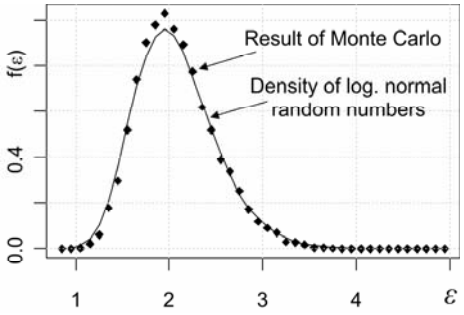


Fig. 6. Distribution of ε based on the result of the Monte Carlo simulation (•) and distribution of log-normal random numbers with same μ^{LN} and σ^{LN} value (-)

3.2772 is the 0.975 quantil of the log-normal distribution with $\mu^{LN} = 0.7158$ and $\sigma^{LN} = 0.2404$. This quantil $q_{0.975}^{LN}$ can be calculated under the knowledge of the location parameters μ^{LN} and σ^{LN} by (6).

$$q_{0.975}^{LN} = e^{\mu^{LN} + q_{0.975} \cdot \sigma^{LN}} \quad (6)$$

where $q_{0.975}$ is the corresponding quantil of the standard normal distribution (1.9600).

This calculated quantil $q_{0.975}^{LN}$ will be called ε_{krit}. For the two simulation examples shown in Fig. 5 one can see that the time delay estimation for simulation 1 ($\varepsilon = 2.9630$) can not be considered as a reliable result because $\varepsilon = 2.9630 < \varepsilon_{krit} = 3.2772$. This result might have been also achieved for two independent time series. The time delay estimation results for simulation 2 can be used for fault propagation analysis as $\varepsilon = 3.4020 > \varepsilon_{krit} = 3.2772$.

In the following section the influence of N and M on ε_{krit} will be determined with α = 0.025.

3.1 Changing the number of measurements

For determining the influence of the number of measurements N on ε_{krit} the same Monte Carlo simulation was used as in section 3 but with varying N, $N \in \{50,100,...,1000\}$ and constant M = 21. For every number of measurement, ε_{krit} was calculated based on (6). The results were shown in Fig. 7.

Based on a nonlinear regression the formula $\varepsilon_{krit} = f(N)$ given in (7) can be derived.

$$\varepsilon_{krit} = 5.4673 \cdot N^{-0.1121} \quad (7)$$

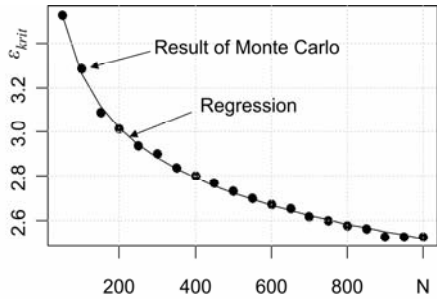


Fig. 7. ε_{krit} as a function of the number of measurements (N) based on the result of the Monte Carlo simulation (•) and nonlinear regression (-)

3.2 Changing the number of assumed time delays

In this section the influence of the number of assumed time delays M on ε_{krit} will be determined with N = 100. The results of the Monte Carlo simulation are given in Fig. 8.

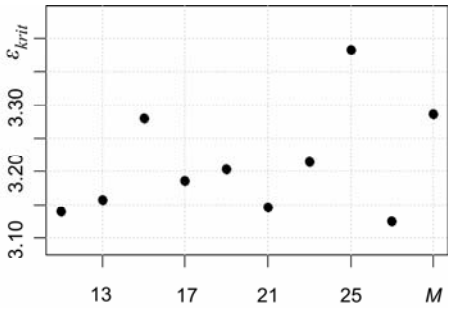


Fig. 8. ε_{krit} as a function of the number of assumed dead times (M) based on the result of the Monte Carlo simulation (•)

There seems to be no coherence between ε_{krit} and M. Therefore, ε_{krit} is only considered as a function depending on α and N.

For a better understanding, the validation process of time delay estimation is demonstrated on the first simulated example shown in Fig. 1 and its corresponding $R_{uy}^{knn}(n_d)$ course shown in Fig. 3.

First, the time series have to be scaled and $\widetilde{R}_{uy}^{knn}(n_d)$ has to be calculated by $\widetilde{R}_{uy}^{knn}(n_d) = R_{uy}^{knn}(n_d) - \overline{R_{uy}^{knn}}(n_d)$, where $\overline{R_{uy}^{knn}}(n_d)$ is the mean value of $R_{uy}^{knn}(n_d)$.

Afterwards, the variables shown in (2) and (3) have to be calculated. For this simulation $\rho^{knn} = 0.0625$ and $\sigma^{knn} = 0.0149$. Based on the calculated variables the quotient ε given in (4) has to be determined. In this case $\varepsilon = 4.1968$.

In the last step, the calculated quotient has to be compared to a critical threshold $\varepsilon_{\text{krit}}$ which might have been also achieved if the two time series were independent. $\varepsilon_{\text{krit}}$ is a function depending on the number of measurements and in this simulation $\varepsilon_{\text{krit}} = 5.4673 \cdot 100^{-0.1121} = 3.2627$. Due to the fact that ε is greater than $\varepsilon_{\text{krit}}$ this time delay estimation can be considered as a reliable and repeatable result.

4. INDUSTRIAL APPLICATION

The presented method for time delay estimation was developed for fault propagation analysis of a hydrocracker. Because several processes were nonlinear or consist of multiple inputs the classical methods could not be used. This was also shown by Stockmann and Haber (2010). In this section, the results of two time delay estimations are shown.

Hydrocrackers are used to crack high-boiling hydrocarbons to more valuable lower-boiling hydrocarbons like naphtha, kerosene and diesel under the presence of high-pressure hydrogen and fixed-bed catalysts. The analysed hydrocracker consists of four catalytic beds and five quenches in which cooling hydrogen is injected to work against the highly exothermic reaction of hydrogenation and to assure an optimal temperature, which lies between 400 and 450°C.

A flow chart of the hydrocracker plant is given in Fig. 9 and a scheme of the second quench is shown in Fig. 10.

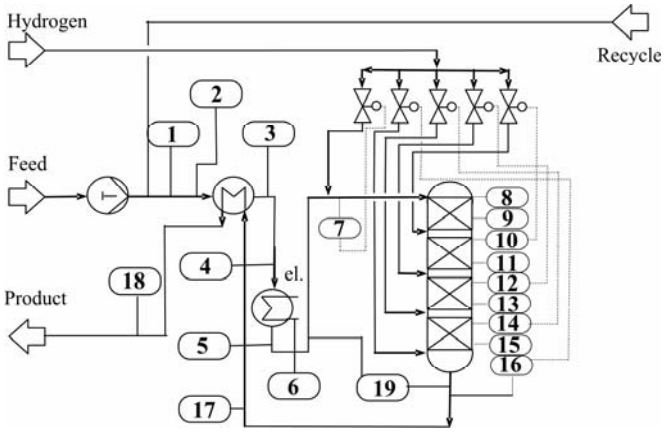


Fig. 9. Flow chart of the hydrocracker plant

4.1 Time delay estimation in case of causal dependency

In this section, the method is tested for the time delay estimation from measurement 9 to control loop 10. Measurement 9 detects the outlet temperature of the first catalytic bed and control loop 10 controls the entry temperature of the subsequent bed by injection of hydrogen.

The time delay between measurement 9 and control loop 10 represents the transport time of the product through the quench. Due to the fact that the manipulated variable in this quench was kept constant, the simple SISO method can be applied.

The courses of the scaled measured and controlled variables are shown in Fig. 11. One can see that there is a causal dependency between these two variables.

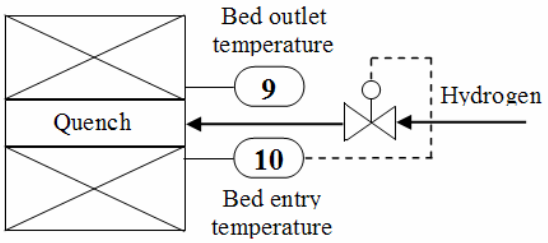


Fig. 10. Scheme of one quench (static mixer) from the hydrocracker shown in Fig. 9

The time series consist of 500 measurements. Hence, ε_{krit} can be calculated based on (7) as 2.7241. The result of the time delay estimation is shown in Fig. 12.

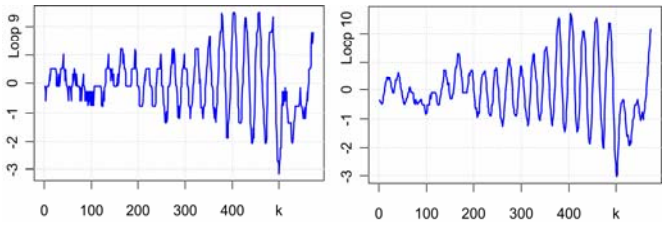


Fig. 11. Course of scaled controlled variables from loop 9 (left) and loop 10 (right)

The course $R_{uy}^{knn}(n_d)$ shows its maximal value at $n_d = 1$. Based on the results shown in Fig. 12 the following parameters can be estimated empirically:

$$\rho^{knn} = 0.02485,\ \sigma^{knn} = 0.00686 \Rightarrow \varepsilon = 3.62395$$

Due to the fact that $\varepsilon = 3.62395 > \varepsilon_{krit} = 2.7241$ with $N = 500$ this result can be considered as a reliable estimation of time delay. Hence, the transport time from measurement 9 to control loop 10 equals one sampling time.

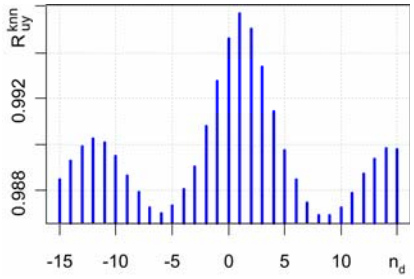


Fig. 12. Time delay estimation between loop 9 and 10

The time delay from control loop 10 to measurement 11 was also estimated as a part of the industrial application. This time delay represents the transport time through the second catalytic bed and is an important indicator of the chemical reaction progress. The maximal value of $R_{uy}^{knn}(n_d)$ can be found at n_d = 3. With $\varepsilon = 2.7358$ and $\varepsilon_{krit} = 2.7241$ this time delay estimation is also validated. The $R_{uy}^{knn}(n_d)$ course of this time delay estimation is shown in Fig. 13.

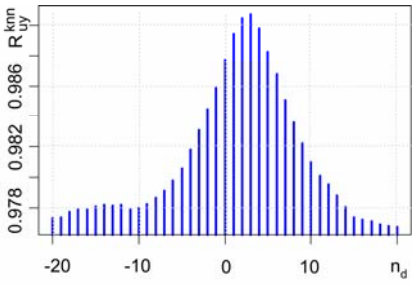


Fig. 13. Time delay estimation between loop 10 and 11

4.2 Time delay estimation in case of no causal dependency

In section 4.1 the time delay estimation delivers a reliable and retraceable result. In this section, the method and its threshold value is tested in case of no causal dependency. Therefore the time delay is estimated between the measured power of the electronic preheater (control loop 6) and the difference pressure measurement (control loop 19). The courses of the scaled time series are shown in Fig. 14.

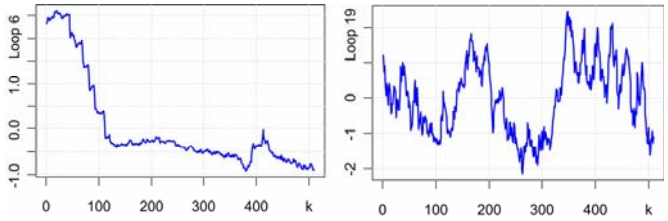


Fig. 14. Course of scaled controlled variables from loop 6 (left) and loop 19 (right)

The time series shown in Fig. 14 consist of 500 measurements as well. The result of the time delay estimation is shown in Fig. 15. The course $R_{uy}^{knn}(n_d)$ in this case shows his maximal value at n_d = -11. Based on the results shown in Fig. 15 the following parameters can be empirically estimated:

$$\rho^{knn} = 0.00336,\ \sigma^{knn} = 0.00138 \Rightarrow \varepsilon = 2.43928$$

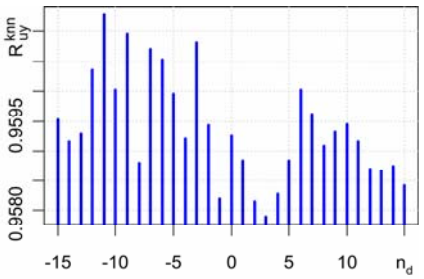


Fig. 15. Time delay estimation between loop 6 and 19

Because $\varepsilon = 2.43928 < \varepsilon_{krit}$ this result can not be considered as a reliable estimation of time delay. Regarding the process engineering this conclusion is evident.

5. CONCLUSIONS

In the present paper, a new method for time delay estimation based on k nearest neighbour imputation is validated and demonstrated on a simulation and on an industrial example. State of the art methods for time delay estimation would have failed generally or higher model assumption would have been needed. A simple index based on the results of the SISO time delay estimation has to be calculated in order to give evidence about its quality and repeatability. Aim of future research will be to test this threshold even for the MISO time delay estimation.

6. ACKNOWLEDMENTS

The authors gratefully acknowledge the support by the Ministry for Innovation, Science, Research and Technology of North Rhine-Westphalia (Germany) and the Cologne University of Applied Science in the framework of the competence platform "Sustainable Technologies and Computational Services for Environmental and Production Processes" (STEPS).

REFERENCES

Bauer, M., Thornhill, N. F., Meaburn, A. (2004). Specifying the directionality of fault propagation paths using transfer entropy. Presented at *DYCOPS 7*, Cambridge, Massachusetts, USA.

Bauer, M., Thornhill, N.F. (2008). A practical method for identifying the propagation path plant-wide disturbances. *Journal of Process Control*, vol. 18, pp. 707-719.

Jelali, M. (2006). An overview of control performance assessment technology and industrial applications. *Control Engineering Practice*, vol. 14, pp. 441-466.

Harris, T. (1989). Assessment of control loop performance. *Canadian Journal of Chem. Eng.*, vol. 67, 1989, pp. 856-861.

Stockmann, M., Haber, R. (2010). Determination of fault propagation by time delay estimation using k nearest neighbour imputation. *Proceedings of SYSTOL'10*, Nice, France.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Estimation and prediction of solar radiation by Meteosat image processing

Zaher A.*, Thiéry F.*, N'Goran Y., Traore A.***

**Laboratoire ELIAUS, Université de Perpignan Via Domitia.
Perpignan, France, (tel : +33 0468662240 ; email : thiery@univ-perp.fr)
** Laboratoire de Physique de la Matière Condensée et Technologie (LPMCT),
UFR SSMT, Université de Cocody, 22 BP 358 Abidjan, Côte d'Ivoire*

Abstract: This paper contains two sections: the first one presents a procedure based on the fuzzy logic inference systems, to optimize the GISTEL method in order to estimate the hourly global solar radiation with better accuracy. In the second section, a new algorithm is proposed to predict this same parameter depending on clouds tracking and the optimized GISTEL method.
To demonstrate the efficiency of the optimizing process, a statistical study is done to compare the results obtained using GISTEL method to those obtained using the optimized one. In order to validate our prediction algorithm, another statistical study is done to show its performance over two forecasting horizons.

Keywords: solar radiation prediction, image processing, optimization, fuzzy inference systems

1. INTRODUCTION

Today, the development of solar energy technology becomes the interest of the most researchers for responding to the growing energy needs in the world. Many applications of solar energy systems depend on the measurement of solar radiation, such as, the control of CSP (concentrating solar power) plants and PV (photovoltaic power) plants. However, the limited numbers of measuring stations for solar radiation are insufficient for use to overcome this problem. Thus, during the last two decades, different methods are proposed for estimating global solar radiation using Meteosat satellite images. One of these methods is GISTEL, which is applied in France and in different countries in Africa using B2, HR and Meteosat Wefax images (Chaâbane et al.,1996; Ben Djemaa et al., 1992).
The last method exploits the visible images (monospectral analysis), but Bachari et al. (2001), presented the importance of using bispectral analysis in the evaluation of solar radiation.
Metfi et al. (2008), compared the method GISTEL to SICIC (solar irradiation from cloud image classification) and they found that the error for the first approach is greater than the error for the second one.

Therefore, the first objective of this work is to optimize the method GISTEL using fuzzy inference system for combining the data extracting from visible images and infrared images for the estimation of hourly global solar radiation.
The second aim is to develop a short-term prediction algorithm for solar irradiance data basing on the detection of the clouds, the measurement of motion in image sequences and the GISTEL method.
The last goal needs a robust algorithm for the estimation of cloud motion field taking into account the semi-fluid behavior of clouds. Many approaches are used to compute the motion vectors in image sequences, such as Block-matching algorithm, Pixel recursive algorithms, optical flow (Ezhilarasan et al., 2008; Horn et al., 1981). The method used in this article based on Block Matching Algorithm combined with a best candidate block search, which is proposed by BRAD et al. (2002) for Extracting Cloud Motion from Satellite Image Sequences.
The estimation and the prediction results are presented and discussed in this work, a comparison between the optimized GISTEL and the GISTEL method is done regarding the accuracy to demonstrate the efficiency of the optimizing process. Another comparison between the results obtained of two forecasting horizons (one hour and two hours) is also done to show the validity of our algorithm.

2. ESTIMATION OF GLOBAL SOLAR RADIATION

2.1 GISTEL Method

GISTEL is a methodology used to evaluate global solar radiation using Meteosat visible channel images. In these kinds of images, the brightness varies from 0 to 255. The minimal value of brightness refers either to the case of clear sky or to the clouds shadow and the maximal value of the brightness refers to overcast sky.
However, GISTEL Model is based on comparing the brightness of a given pixel with reference value in case of clear sky at the same time in order to estimate the irradiation fraction (k). So, hourly and monthly reference images ($Image_{clear\text{-}sky}(hour, month)$) are constructed to represent the case of clear sky for every month of the year using an archive of images.

To compute the irradiation fraction, we use the following rules that classify the weather into three categories of sky: clear, partly covered and overcast sky, using the reflection coefficient (a), as proposed by Chaâbane et al. (1996), Ben Djemaa et al. (1992), Muselli et al. (1998) and Chaâbane et al. (2002):

$$\begin{aligned} &\text{If } (a \le a_{min}) \text{ then } k = 1 \\ &\text{If } (a_{min} < a < a_{max}) \text{ then } k = \\ &\qquad 1 - (1-k_0)\,[(a-a_{min}) / (a_{max}-a_{min})] \\ &\text{If } (a \ge a_{max}) \text{ then } k = k_0 \end{aligned} \quad (1)$$

The cloud cover index (n) is given by:

$$n(x, y, h) = (a-a_{min}) / (a_{max}-a_{min}) \quad (2)$$

a_{min} and a_{max} are respectively the minimum and the maximum reflection coefficient, $k_0 = 0.2$, These rules represent respectively the clear, the partly covered and the overcast sky.
Then the global solar irradiance (G) will be obtained using:

$$G(x, y, d, h) = k(x, y, d, h).G_c(x, y, d, h) \quad (3)$$

Where G_c is the global solar irradiance in the case of clear sky, this value is calculated using the model recognised by the WMO (World Meteorological Organization). (1981). x and y are pixel coordinates, d is the day and h is the hour. Let's note that in the expression of G_c, are taken into account many parameters such as the Linke turbidity factor, the angular elevations of the sun and the satellite.

2.3 Gistel method optimization

This work focuses on the use of the cloud top temperature extracted from infra-red images provided by EUMETSAT to optimize the GISTEL method**.** On these images, the gray levels ranging from 0 to 255 and refer to the temperatures from -75 to 45 °C.
To improve the precision of this method, the cloud cover index calculated using (2) from visible images and the temperature attributed to pixels (*T)* will be used in a fuzzy inference system (FIS) in order to estimate the irradiation fraction as shown in Figure 1.

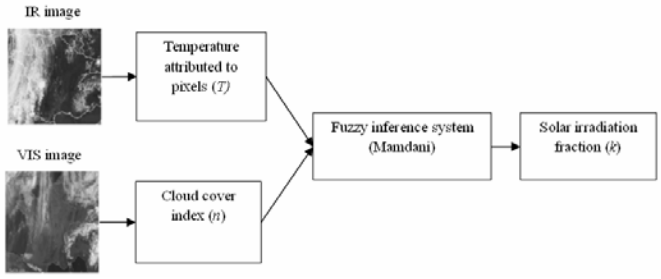


Fig. 1. Inference system block diagram.

Fuzzy inference systems **(***FIS***)** have a multidisciplinary nature and they have been successfully applied in a wide variety of fields, such as automatic control, data classification, decision analysis, expert systems, and computer vision. There are two main structures of fuzzy inference systems: Mamdani-type and Sugeno-type. Here, Mamdani's fuzzy inference system (Mamdani et al., 1975) is used in order to compute the irradiation fraction from the two inputs: the cloud top temperature and the cloud cover index. The first step when making up a FIS is the fuzzification.
The cloud cover index varies from 0 to 1, the cloud top temperature varies from -75 to 45 °C and the irradiation fraction varies from 0 to 1. These three variables were fuzzified into Gaussian membership functions as presented in figure 2(a, b and c). The inferences are realized by the following rule base:

If (T is small) and (n is very small) then (k is very big)

If (T is small) and (n is small) then (k is big)

If (T is small) and (n is big) then (k is small)

If (T is small) and (n is very big) then (k is very small)

If (T is big) and (n is very small) then (k is very big)

If (T is big) and (n is small) then (k is big)

If (T is big) and (n is big) then (k is big)

If (T is big) and (n is very big) then (k is small) (4)

As presented in figure 2(d), a non-linear relationship can be noticed between the input and output variables. In the described fuzzy inference system, the fuzzy rule structure is predetermined by the interpretation of the characteristics of the variables and the membership functions are empirically tuned depending on the data.

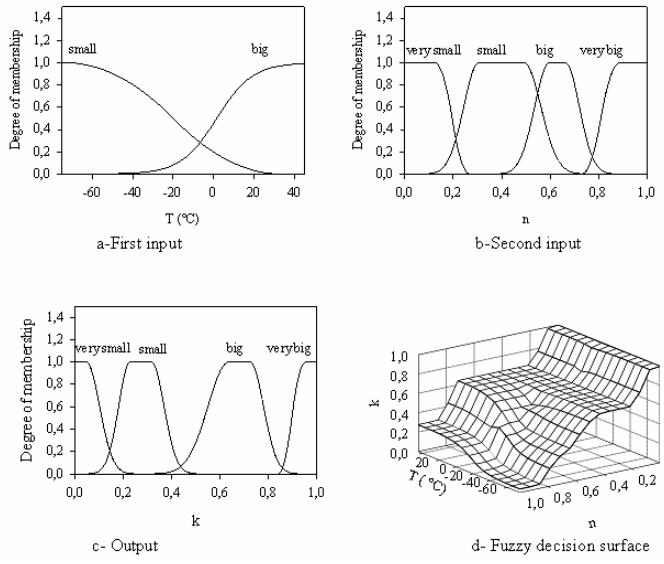


Fig. 2. Input and output membership functions and the decision surface of the FIS.

3. PREDICTION OF THE GLOBAL SOLAR RADIATION

In this section, the interest is a short-term global solar radiation prediction depending on GISTEL method and on clouds speed estimation. The prediction algorithm is presented in Figure 3**.** In this algorithm, we have to predict the global solar radiation at pixel $p_0\ (x_0, y_0)$ at the time $(t + \Delta t)$, thus, the first step is to estimate the case of sky in the studied zone in the future by segmentation of the visible *image (t)* using the cloud cover index. The result is an image of cloudy pixels and clear pixels, if the number of cloudy pixels is smaller than threshold that means the sky at the pixel $p_0(x_0, y_0)$ will be clear at the time $(t + \Delta t)$, else the sky will be cloudy and then we must determine the cloud cover index and the temperature which will be attributed to the pixel p_0 at the time $(t + \Delta t)$. For this reason, the next step is to compute the motion vectors u and v in order to estimate the position of the pixel $p_1(x_1, y_1)$. The clouds at p_1 will move towards p_0 during Δt**.** The coordinates for the pixel p_1 are given by:

$$x_1=x_0-u.\Delta t$$
$$y_1=y_0-v.\Delta t \quad (5)$$

The last step is to evaluate the global solar radiation $G\ (x_0, y_0, t+ \Delta t)$ using the following equation:

$$G(x_0, y_0, t + \Delta t) = k_0\ (x_0, y_0, t+\Delta t).\ G_c(x_0, y_0, t + \Delta t) \quad (6)$$

Where k_0 is given by:

$$k_0\ (x_0, y_0, t+\Delta t) = \alpha.\ k_1\ (x_1, y_1, t+\Delta t) \quad (7)$$

Where α is a coefficient taking into account the attenuation of k_1 and the indices 1 and 0 in this section refer respectively to pixels p_1 and p_0.The irradiation fraction k_1 can be calculated by the optimized GISTEL method or by GISTEL method as mentioned in the previous section.

Now, the motion vectors must be estimated correctly to obtain a good result of the prediction algorithm. Therefore, the block matching algorithm **(***BMA***)** combined with a best candidate block search is used.

The basic concept of *BMA* is to divide the frame into small blocks then, the *BMA* finds the optimal motion vectors (*MVs*) that minimize the difference between reference block of the current frame and candidate block from the search area of previous frame.

The *SAD* is used to match criteria, and is defined as (Liaw et al., 2009; Song et al., 2000):

$$SAD(x_0, y_0) = \sum_{i=0}^{L-1}\sum_{j=0}^{M-1}\left[\begin{matrix} I_n(x_0 + i, y_0 + j) - \\ I_{n-1}(x_0 + x + i, y_0 + y + j)\end{matrix}\right] \quad (8)$$

$I_n(x_0, y_0)$ is the block of the current frame (n) and (x_0, y_0) is the coordinate of its upper left corner. $I_{n-1}(x, y)$ is the block of the previous frame (n-1) and (x, y) is the coordinate of its upper left corner. Each of these blocks is of $L \times M$ pixels. But in this method, the future evolution of the block is not taken into account, thus a big error occurs in the estimation. For this reason, (*BMA*) is combined with a best candidate block search (BRAD et al., 2002).

In the last method, the *BMA* is applied and a list of the best candidates is stored with their scores *Sc* computed using (8) and the Euclidean distance between the reference block and the candidate (*Cc*), for each search region, for each *nth* image pair of the *N-1* image pairs in the sequence.

The Euclidean distance between the reference block and the block candidate is obtained by:

$$C_c(x_c, y_c) = \sqrt{(x_c - x_0)^2 + (y_c - y_0)^2} \quad (9)$$

The criterion used to match the candidates to the reference block in the search zone is a cost function of the distance and the scores. This function is given by:

$$Cost = p.S_c + (1 - p)C_c \quad (10)$$

With p is a weigth parameter equal to 0.8.

The best block match is obtained by minimizing the cost function along the *N-1* image pairs using:

$$Match = \min \sum_{n=1}^{N-1} Cost_{n,c} \quad (11)$$

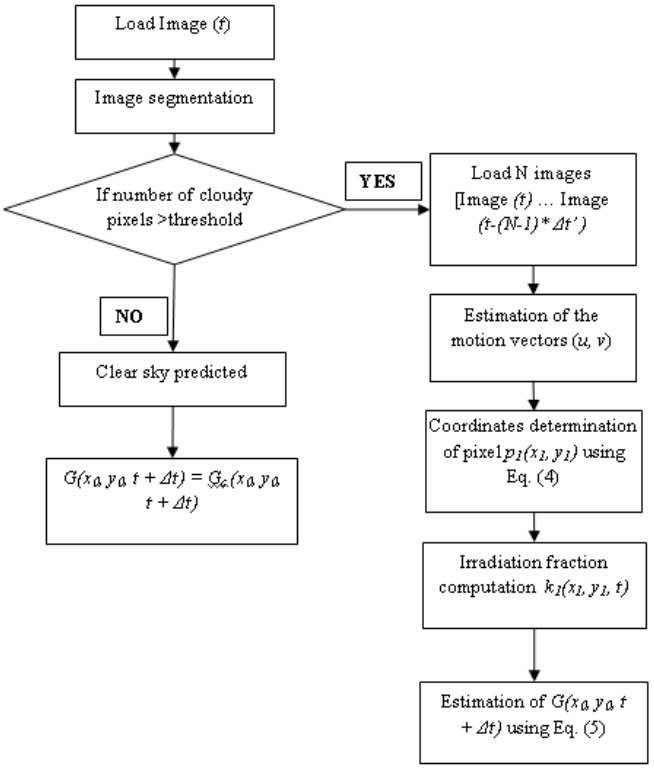


Fig. 3. Flow diagram of prediction algorithm. Δt' being the time interval between two successive images.

4. RESULTS AND DISCUSSION

4.1 Estimation

The images used in this study are H.R. (high resolution) images provided by EUMESAT with precision of 2.5×2.5 km^2 for visible images and 5×5km^2 for IR images and the time interval between two successive images is one hour.
The global solar radiation is measured in the University of Perpignan (42.700 N, 2.900 E) using CS300 pyranometer with precision of 0.01 %, in addition to different meteorological variables, such as wind speed, the temperature and the humidity. These data can be found online at **http://eliaus.univ-perp.fr/solaire/mesures/**. The data in this work covered the period from 1 April to 31 October 2009.
In order to evaluate the accuracy of the optimization method, described previously, we use the statistical criteria: the Relative Mean Bias Error (*RMBE*)**,** the Relative Root Mean Square Error (*RRMSE*) and the correlation coefficient**(***R***)**, these indicators are used to evaluate GISTEL method (Chaâbane et al., 1996; Ben Djemaa et al., 1992) Table 1**.**
The smaller values of *RMBE* and *RRMSE* and the bigger value of *R* then the better the performance.
Depending on this rule, the results in Table 1 show a significant improvement by using the optimized method.
The scatter diagrams plotted in Figure 4, present another means to validate the optimizing process. In this figure, the results of the optimized method indicate a strong correlation between the estimated and the measured data (the closer the points to the straight line then the stronger the correlation).

Table 1. Statistical results of comparison between measured and estimated hourly solar radiation using GISTEL method and the optimized approach.

Method	GISTEL	Optimized GISTEL method
R (%)	94	97
RRMSE (%)	16	10
RMBE (%)	0.9	0.5

4.2 Prediction

To estimate the clouds velocity the following parameters are used in every step of estimation:

- Four visible images, at 15 minutes time interval.
- Block size equals to 4×4 pixels
- Research window equals to 8×8 Blocks

In order to evaluate the performance of the prediction algorithm, the data were predicted using two forecast horizons and we use the statistical indicators presented above, *RRMSE*, *RMBE* and *R***,** to compare the results.
Results presented in the table 2, show a good efficiency of the prediction algorithm for the two forecast horizons. But in the same time, these results are better using a prediction horizon of one hour.
The predicted data compared to the measured ones are presented in the figure 5. This figure shows a better correlation for a forecast horizon of one hour.

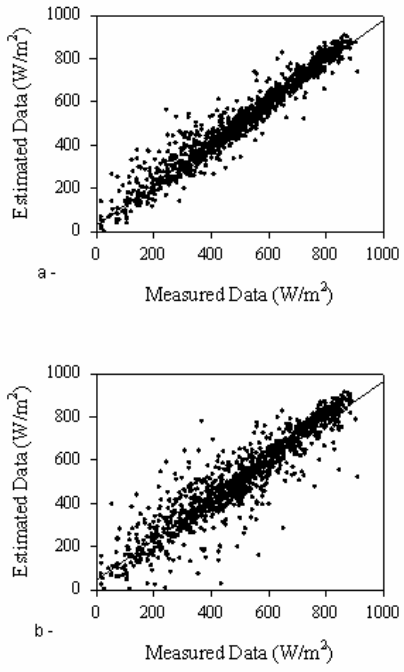


Fig. 4. Scattering diagram of measured and estimated hourly solar radiation: a- using optimized GISTEL method, b- using GISTEL method.

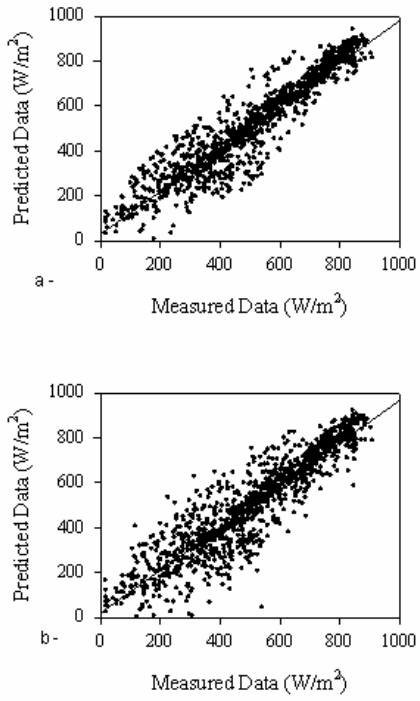


Fig. 5. Scattering diagram of measured and predicted hourly solar radiation: a- using one hour forecast horizon, b- using two hours forecast horizon.

Table 2. Statistical results of comparison between measured and predicted hourly solar radiation using one hour and two hours forecast horizon

Forecast horizon	One hour	Two hours
R (%)	93	90
RRMSE (%)	14.63	18
RMBE (%)	0.8	-1.4

5. CONCLUSION

In this paper, a procedure was used to optimize the GISTEL method basing on the fuzzy inference system. The statistical study demonstrated the smaller values of the indicators (RRMSE, RMBE) and the bigger value of correlation coefficient comparing to the other indicator resulted by using GISTEL approach.

In the second section of this paper, was proposed a new algorithm to predict the hourly global solar radiation depending on the estimation of the clouds velocity and the optimized GISTEL method, and the statistical study which is done for two forecast horizons showed the good performance of this algorithm.

REFERENCES

Bachari N., Benabadji N., Razagui . A. and Belbachir A.H., (2001), Estimation et Cartographie des Différentes Composantes du Rayonnement Solaire au Sol à Partir des Images Météosat, *Rev. Energ. Ren*. Vol. (4), 35-47.

Ben Djemaa A. and Delorme C., (1992), Comparison between one year of daily global irradiation from ground based measurements vs Meteosat images from seven locations in Tunisia, *Solar Energy*; 48(5), 325-33.

BRAD R. and LETIA I. A., (2002), Cloud Motion Detection from Infrared Satellite Images, *The Second International Conference on Image and Graphics (SPIE)*.

Chaâbane M., Ben Djemaa A. and Kossentini A., (1996), Daily and hourly global irradiations in Tunisia extracted from Meteosat Wefax images, *Solar Energy*; 57(6), 449-57.

Chaâbane M. and Ben Djemaa A., (2002), Use of HR Meteosat images for the mapping of global solar irradiation in Tunisia: preliminary results and comparison with Wefax images, *Renewable Energy*, (25), 139-151.

Ezhilarasan M. and Thambidurai P., (2008), Simplified Block Matching Algorithm for Fast Motion Estimation in Video Compression, *J. Computer Sci.*, 4 (4), 282-289.

Horn B. K. P. and Schunck B. G., (1981). Determining optical flow, *Artificial Intelligence*, 17(1-3):185–203.

Liaw Y.C., Lai J. Z.C. and Hong Z .C., (2009), Fast block matching using prediction and rejection criteria, *Signal Processing*, (89), 1115-1120.

Mamdani E.H. and Assilian S., (1975), An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies*, 7(1), 1-13.

Mefti A., Adan A. and Bouroubi M.Y., (2008), Satellite approach based on cloud cover classification: Estimation of hourly global solar radiation from meteosat images, *Energy Conversion and Management*, (49), 652-659.

Muselli M., Poggi P., Notton G. and Louche A.,(1998), Improved procedure for stand-alone photovoltaic systems sizing using Meteosat satellite images, *Solar Energy,* 62(6), 429-444.

Song B. C. and Ra J.B., (2000), A fast multi-resolution block matching algorithm for motion estimation, Signal Processing: *Image Communication,* (15), 799-810.

WMO, (1981), Meteorological aspects of the utilization of solar radiation as an energy source, *Technical note WMO*. 172, 57-85.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Advanced and Predictive Diagnosis on the Example of Pump Systems

S. Kleinmann*, A. Dabrowska **, D. Leonardo*,**
R. Stetter, A. Koller-Hodac*****

* Allweiler AG, 78315 Radolfzell, Germany (e-mail: S.Kleinmann@allweiler.de)
** Hochschule Ravensburg-Weingarten, 88250 Weingarten, Germany
(e-mail: dabrowska(stetter)@hs-weingarten.de)
*** Hochschule Rapperswil, CH-8640 Rapperswil, Switzerland (email: dleonardo@hsr.ch)

Abstract: In many industrial facilities pump systems play a crucial role. The serve for lubrication purposes which are needed in a large share of applications requiring motion, for cooling and for the supply of all kinds of fuels for combustion e. g. in power plants. In contrast to this importance, today only rather basic diagnosis systems are used in industrial applications. This paper presents some of the results of the project which goal is increasing energy efficiency as well as reliability and effectiveness and efficiency of service. Earlier work has shown that an integrated approach considering control and diagnosis simultaneously is most likely to be adopted by industry due to several reasons. This paper is focused on the viewpoint of diagnosis reporting about the developed concept which is relying on a hierarchical concept ant the systematic development of diagnostic functions. Here the concept of predictive diagnosis is elucidated. A core element of the methods of advanced and predictive diagnosis is a mathematical model of the pump system. This model is therefore described in detail.

Keywords: risk assessment, diagnosis, control, monitoring, pump systems, burner industry

1. INTRODUCTION

Pump systems play a significant role in whole industry environment. They are used in power plants for instance for fuel injection system, lubrication or cooling purposes. In technological processes they are the main driving source of all processes. Unfortunately until now the majority of pump systems in industry are not equipped with diagnosis systems. This is even more astonishing if one is aware that the failure of pump systems will usually lead to a shutdown of the respective industrial facility. This problem is currently tackled by three time and cost efficient approaches:

- regularly servicing the pump systems (even if they do not yet need it, installing redundant pumps) leading to high service costs and shut-down times of the industrial facilities,
- installing redundant pump systems leading to increased installation cost and space demand and, in the case of hot redundancy, lower efficiency of the system,
- and regularly replacing the pumps leading to costs and unnecessary waste.

Systems relying on advanced and predictive diagnosis dispose of the potential to dramatically increase the reliability and serviceability of technical systems. In order to offer the possibility to be implemented in pump systems the strategies, methods and tools of advanced and predictive diagnosis need to be adapted and refined. For this purpose two Universities (Hochschule Ravensburg-Weingarten, Germany and Hochschule Rapperswil, Switzerland) together with a leading German pump manufacturing company (Allweiler AG) engaged their employees to start research project called “Smart Pumps”. The main goal of the project is development of a well founded prognosis of future control and diagnosis technologies for intelligent pump systems.

An in-depth investigation (Kleinmann et al. 2009) showed that a large increase of the application rate of control and diagnosis systems in industry can only be realized if a modular system can be created which will connect three fundamental functions of the optimization in order to increase reliability and serviceability: control, monitoring and diagnosis. Especially the last function shows wide field of research, when it is based on self-learning the typical behaviours in the pump system. The successful application of advanced and predictive diagnosis in whole system could lead to early fault detection and identification like e. g. cavitation which occurrence causes loss of pump efficiency and damage of mechanical parts and integrity of the pump.

In the next section the notions diagnosis, advanced diagnosis, predictive diagnosis, control and monitoring are briefly explained. The third section presents the hierarchical concept for monitoring, control and diagnosis. The system model is discussed in the fourth section. Section five describes the systematic development of diagnostic functions.

2. CLARIFICATION OF NOTIONS

2.1 Diagnosis, Advanced Diagnosis

Diagnosis is usually understood as the process of estimating the object condition. Diagnosis is carried out by the

estimation of important parameters and the determination what should be done in case when faults occur. Over the last three decades, the growing demand for safety, reliability, and maintainability in technical systems has drawn significant research in the field of diagnosis. Since more than two decades methods of **advanced diagnosis** were developed which are usually characterized by the application of some kind of model of the system which is to be diagnosed. Such efforts have led to the development of many techniques; see for example the most recent survey works (Blanke et al. 2006, Isermann 2005, Witczak 2007, Zhang and Jiang 2008, Korbicz et al. 2004). The application of a collection of these techniques gathered in one system (AmandD - Koscielny et al. 2006) was analyzed by Dabrowska und Kleinmann (2009). For fault compensation in general fault tolerant control methods are proposed which can be classified into two types, i.e. Passive Fault Tolerant Control Scheme (PFTCS) and Active Fault Tolerant Control Scheme (AFTCS) (Blanke et al. 2006, Zhang and Jiang 2008).

2.2 Predictive Diagnosis

The term "predictive diagnosis" is in contrast to "predictive control" rarely used in the technical domain (it is widely used e. g. in medicine). One example for the usage of this term is the presentation of an automated system for fault diagnosis based on vibration data recorded from an main power transmission (Diwakar et al. 1998). For pump systems predictive diagnosis (essentially in the meaning of failure detection and identification before these failures even occur) presents a promising field of research. Five main problems in the operation of pumps bear the possibility to be identified early:

- wear of the seals leading firstly to increased leak-age and finally to system failure;
- wear of the surfaces for flow or pressure genera-tion (e. g. the spindles of screw pumps) leading to back leakage, reduced efficiency and finally to system failure;
- changes of the viscosity of the hydraulic medium, e. g. as a consequence of degradation leading to changed operation conditions, reduced efficiency and power (if a pump system was optimized with regard to a certain viscosity) and in extreme cases to destruction of the pump;
- agglomeration of particles in the hydraulic medium usually as a consequence of wear of elements in the hydraulic circuit leading to increased wear of the seals and the surfaces for flow or pressure generation;
- vibration which could be caused from high air content for lube oil applications or system suction conditions which does not allow operating the pump according to NPSH required.

Pumps are usually part of larger systems. A failure of the larger system which is caused by a failure of a pump usually leads to enormous consequences in terms of cost of idleness (e. g. of a whole power plant). Therefore preventive maintenance is desirable for industrial pumps; however such preventive maintenance today is aggravated by the fact the up-coming failures can usually not be detected. The only preventive maintenance systems possible are time based but not state based. A predictive diagnosis system would allow scheduling maintenance and service in dependence of the current state of a pump (wear of seals and surfaces) and the state of the medium (viscosity, agglomeration of particles).

2.3 Control

The term "control" names activities intended to manage, command, direct or regulate the behaviour of devices or systems and has been the core of extensive research for many decades. In recent years the techniques of predictive control have found rising attention (compare e.g. Camacho&Bordons 2008). Predictive control usually relies on dynamic models of the process, most often linear empirical models obtained by system identification. In the area of pump systems predictive control can pursue three different objectives:

- smoothing changes of system states,
- better coordination of multiple pumps and
- evaluating decision alternatives.

2.4 Monitoring of Operation Data

The notion monitoring summarizes all kinds of systematic observation, surveillance or recording of an activity or a process by any technical means. In the area of pump systems monitoring could be understood as a systematic collection of data concerning the state of certain physical (usually hydraulic or electrical) characteristics such as flow, pressure, temperature, viscosity, vibrations, torque, velocity, currents, voltage, current gradient, velocity gradient, etc. In leading industries such as computer chip production or car manufacturing today usually nearly all operation data of the productions systems are being monitored for the three main reasons safety, efficiency and planability:

- The safety of production systems can be enhanced because a reliable safety system with a fast reaction can be realized on the basis of a real-time monitoring system. The role of coincidence for detecting possibly dangerous faults is diminished if a continuous monitoring is in place.
- The efficiency of production systems can be enhanced because any kind of waste (of energy, time and production goods) will be detected and can subsequently be prevented or reduced.
- The planning possibilities and planning quality can be enhanced if accurate data from a real-time continuous monitoring system are available as realistic prognosis is enabled by such data.

Such monitoring of pumps is currently not realized but could be an additional function of a control and diagnosis system.

3. HIERACHICAL CONCEPT

As reported before that a large increase of the application rate of control and diagnosis systems in industry can only be realized if a modular system can be created which will

connect three fundamental functions of the optimization in order to increase reliability and serviceability: control, monitoring and diagnosis. This modular system needs to be integrated in the IT-infrastructure of the industrial facility. In connected work a hierarchical concept including an advanced monitoring, control and diagnosis system (MCDS) was proposed (Kleinmann et al. 2000). Figure 1 shows a proposal of a sensible hierarchy of the levels of these information systems in form of a pyramid.

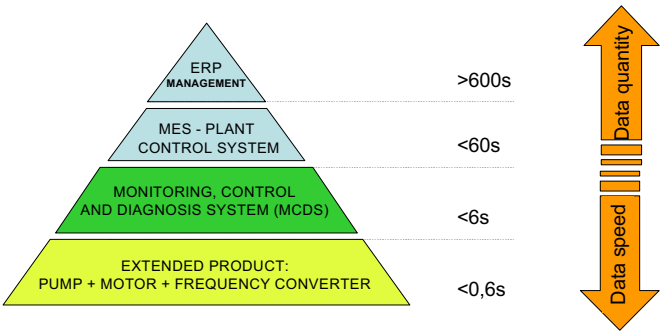


Fig. 1. Hierarchical information system structure for pump systems.

On the highest level the Enterprise resource planning (ERP) system can be found. It is not present in all kind of enterprises and is sometimes referred to with different names in the area of pumps. The main function is always the same: this level concerns the planning of the entities to be produced (e. g. amount of heat in burner applications, amount of energy of power plants) on a not time-critical level. On the next lower level the production of these entities is executed by a plant control system which fulfils similar tasks than a Manufacturing Execution System (MES); in current pump applications a number of names is given to these systems. Some-times the two highest levels are realized in only one system.

On the next level is the first core of the proposed concept – the monitoring, control and diagnosis system (MCDS) for a section of a hydraulic system usually including a number of pumps. Hydraulic systems usually contain more than one pumps e. g. for reasons of redundancy or high pressure which cannot be easily realized with only one pump. The lowest level is called extended product. In industrial applications frequently pumps are used in combination with an electrical motor. Such components are the basis for this level named ”extended product”. This notion means the pump in combination with a appropriate electrical motor, the necessary electronic equipment to use this motor (frequency converter, …) and the control systems which may be delivered with this package. On this level the real-time control has to take place and the most important safety functions should be realized on this level for the sake of a quick reaction.

4. SYSTEM MODEL

Currently, there are many designs of pump systems, depending on type, capacity and especially on further application. A detailed analysis of many different factors, like market requirements and extending knowledge, moved project into direction of research focused on three screw spindle pumps SPF in burner industry application. As a technical challenge, this application creates also many interesting aspects for further applications. A sensible level of investigation is the so-called “extended product”, composed from pump, asynchronous motor with frequency converter and consumer part (process load). Each of these components must be taken into consideration for modelling this system. Each of the model blocks was modelled separately in Matlab/Simulink and later the blocks were connected into an overall system (Figure 2). The motor speed n_2 is the input to the pump system block and the torque load from the pump side M_{load} is the input to the motor converter system block. The theoretical flow rate Q_{act} from the pump is the input to the consumer system block and the required pressure p is the input to the pump system block. These three components are described in the following sections.

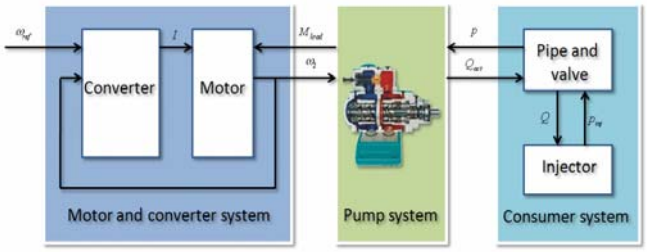


Fig. 2. Overall pump system.

4.1 Motor model

Because of good dynamic, the so-called vector control methods of asynchronous motors are widely used in industrial circuits where the control quality is the determining factor in the whole technological process. In this research project the vector method called DTC (Direct Torque Control) invented by Takahashi and Noguchi (1985) was used. Concerning the principle of control this method is based on the control of flux and torque. Set values of stator flux and electromagnetic torque (exit of speed regulator) compared with real, measured values of suitable signals, are transmitted to the nonlinear comparators. Optimal switches are tabulated in the memory addressed by the state of two-level comparator of main flux regulation and three-level torque comparator, depending on the sector $N(\pi/3)$, which in flux vector Ψ_M currently is. The use of switching table for voltage vector selection provides fast torque response, low inverter switching frequency and low harmonic losses without the complex field orientation by restricting the flux and torque errors within respective hysteresis band with the optimum selection being made.

The DTC scheme is much simpler than other vector control methods due to absence of coordinate transformation between stationary frame and synchronous frame and PI regulators. It also does need the PWM (Pulse Width Modulation).

The motor model is equal (the same parameters) to the real asynchronous motor with nominal power P_n= 2.2kW and nominal angular velocity n_n= 1415rpm. This motor control model has good results in static and also in dynamic states.

4.2 Pump model

Beside the main goal of this project, it is intended to achieve also additional scientific benefits. One of the most important is the creation of a dynamic model of the three screw spindle pump. Unfortunately, scientific literature and research not contain many descriptions concerning three screw spindle pumps (in contrary to other pump types). Because of the lack of a mathematical model of these pumps, the model which was realized also in Matlab/Simulink will produce the desired output from the given input merely based on rather basic mathematical equations prepared by employees of Allweiler AG. This means that the present model of the pump is not yet a real dynamic model. Currently, the project team is working on the validation of the existing model in the real test field and on the expansion this model to the dynamic considerations.

SPF pump system model contains four main important (calculation) blocks:

- the theoretical flow rate Q_{th},
- leakage loss Q_v,
- the theoretical power P_{th},
- the friction power P_r .

The actual flow rate Q_{act}, total power consumption P_{act}, the volumetric- η_{vol} , the mechanical- η_{mech} and the pump efficiency η_{pump} are merely calculated from the model blocks mentioned above.

4.3 Process load

Screw pumps have been used in many different application domains what result in variety of construction, from a simple load to an extensive system.

The model is built with using such components by means of which it can be easily combined to build almost every model of a process load. The main library in the model consists of five standard hydraulic parts of process load like a pipe, a pipe bend, a regulating valve, a y-pipe and a nozzle and creates possibility to model each part individually and also allow observing conditions between components. Using a simple user interface, each component model can be parameterized and adjusted according to the real component properties. The y-pipe is the only component with one inlet point and two outlet points. Depending on the hydraulic resistance on the outlet points the y-pipe divides the incoming medium to the outlets in the correct way. For example if the regulating valve produces more resistance to the return thread in the tank, so the largest amount of fluid will pass through the nozzle.

Based on the actual flow rate Q of the pump, the simulation program calculates the whole back pressure of the system Δp_{tot} that will affect the pump in different ways. Each load n calculates its pressure loss Δp_n by the actual flow rate Q and passes the actual incoming flow rate Q to the next component in the chain. Furthermore each load element adds to the own pressure loss Δp_n the pressure loss of the successive component Δp_{n+1}. Consequently the first component in the chain directly after the pump contains the total back pressure $\Delta p_{tot}=\Delta p_n+\Delta p_{n+1}+\Delta p_{n+2}$ on the pump.

The flow velocity w of the fluid is a linear function of the flow rate Q as described in equation (1). This also depends on physical properties such as the diameter of a pipe.

$$w = Q \cdot b \quad (1)$$

The pressure loss Δp depends on the fluid flow velocity w, the friction factor ζ and the density ϱ of the fluid passing through a component. The pressure loss can be calculated using equation (2).

$$\Delta p = \zeta \cdot \frac{e}{2} \cdot \overline{w}^2 \qquad (2)$$

The friction factor ζ varies in function to the component geometry as well as the viscosity of the medium and accordingly its temperature. This factor can be derived by a calculation, in the simple case of a straight pipe, or can be obtained in an empiric way for components with a complicated inner construction. This model was developed on basis of the equations and empiric deviations described in literature (Bohl et al. 2008). The roughness of the interior wall surface as well as hydrostatic effects has also been taken in to account in the calculation.

The mathematical model of the process loads previously described has been implemented in Matlab/Simulink. Each component model has been designed in a similar way. It consists of an input Q for the incoming flow rate, an output Q for the successive component, an input p for the pressure loss of the successive component and an output p to give back the pressure loss to the previous component. The equations have been implemented graphically and structured using sub-functions. An overview of the system is for the sake of readability given in Appendix A.

5. SYSTEMATIC DEVELOPMENT OF DIAGNOSTIC FUNCTIONS

The scientists involved in the project, together with expert from Allweiler AG and lead user experts from another company developed a table of possible diagnostic functions in the pump system called “Risk-table” in order to be able to systematically explore the diagnosis possibilities. A short fragment of this large table is for the sake of readability included in this publication as Appendix B (without columns). Currently, the table discerns eleven different possible fault groups which can occur in the process load part of the system, e. g. pump break, oil leak, empty oil tank or in-coming air in to system.

The table presents the results of a systematic analysis of the possible occurrence of faults in the process load (burner application), their detection and their localization possibilities. The goal of this analysis is to research the best possible sensor combinations and locations which guarantee the best diagnosis of system (Figure 3). However, as the fundamental principle of this analysis the pump itself is taken as a main sensor in the system (actuator as sensor – this is possible using the methods of advanced diagnosis). To get better results also virtual sensors were applied and real sensors localized in different, strategic places in the system. The table shows exactly how many faults can be distinguished during one sensor combination and precisely indicate how the fault is localized in the process load part.

5.1 Systematic Development of Diagnostic Functions

The project behind the presented research is aimed at developing a long time perspective for the monitoring, control and diagnosis of pump systems. It is therefore desirable to cover a large field of possible solutions for these three general application types and to be aware of "white fields", i. e. possible applications in this field which will no be covered. Therefore a two-sided approach was chosen. On the one hand it was systematically determined which strategies, methods and tools are available or currently researched. For this purpose concise literature surveys (e. g. Blanke et al. 2006, Isermann 2005, Witczak 2007, Zhang and Jiang 2008, Korbicz et al. 2004) and certain systems (e. g. AmandD - Koscielny et al. 2006) were analyzed. Some results are reported in Dabrowska&Kleinmann 2009. On the other hand possible functionalities were systematically collected by means of a so-called "risk-table". This table is described in the next section.

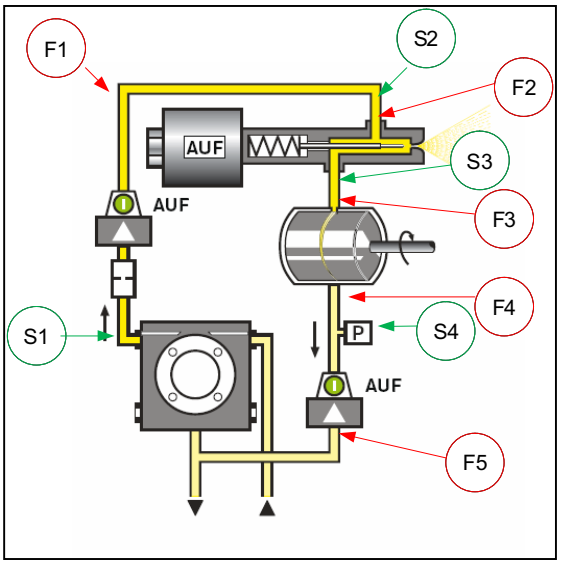


Fig. 3. Possible sensor *S* combinations and fault *F* occurrence in the process load part.

5.2 Explanation of the "Risk-table"

Technical faults in industrial processes create the danger of system break down and expose the industry for additional costs. According to this conditions, were created "Risk-table" which distinguish possible fault appear during functioning of pump system in the burner industry. Additionally, were analyzed, what kind of system behaviour appear during incorrect principle of operation in the system and what are the consequences of given faults occurrences.

By means of this table it is shown the best possible configuration of sensor replacement in the process load for typical burner application. Currently, it is taken into consideration leakage pressure and his changes. Table give specific information what kind of fault could be hidden behind the typical system behaviour. For example, through pressure decrease is possible to detect pipeline rupture or clogged pipeline when value of this signal is increasing. By expanding this table for different strategic signals in the pump system, like flow or speed, it is possible to increase the precision of applied diagnosis.

Up to this time, presented table is prepared on the base of complete burner simulation model in the Matlab / Simulink environment. System can be diagnose in regard to diagnosis, by using signal threshold examination or taking inlet real pump parameters which are put in the simulation and make comparison the several output in the model with the sensors in the real system.

Currently, project team is working on real stand-up to compare accuracy of the results given by simulation model with real technical units. The challenge is to adjust every part of the simulation model to get the same behaviours which appear in the real system. Additionally, necessity is to make a deep and exact research of signal changes. Therefore the systematic work with the risk table is first step to prepare a fault identification and localization in the system. In the next steps this system will be expanded using predictive diagnosis and focusing on the state of the pump as well as the driving electrical motor and its frequency converter.

ACKNOWLEDMENTS

The project is founded by the European Union in the scope of the European Fond for regional development in the Interreg IV program Alpenrhein-Bodensee-Hochrhein together with the Schweizer Bund and the Fürstentum Lichtenstein as well as the Internationale Bodensee-Hochschule.

REFERENCES

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M.: Diagnosis and Fault–Tolerant Control. Berlin: Springer, 2006.

Bohl W., Elmendorf W.: *Technische Strömungslehre*, Vogel Verlag und Druck, 2008 (in German).

Camacho, E., Bordons, C.: Model Predictive Control. Berlin: Springer, 2004.

Dabrowska, A., Kleinmann, S.: Analysis of Possible Applications of the Advanced Modelling and Diagnosis System AMandD in German Industry. In: Proceedings of ACD'2009.

Diwakar, S. Essawy, M.A. Sabatto, S.Z.: An intelligent system for integrated predictive diagnosis. In: Proceedings of the Thirtieth Southeastern Symposium on System Theory, 1998.

Isermann, R.: *Fault–Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Berlin: Springer 2005.

Kleinmann, S.; Dabrowska, A.; Agathe Koller-Hodac, A.; Leonardo, D.: Model of a combined pump and drive system for advanced control and diagnosis. Accepted for presentation during the Conference "Mechatronic Systems and Materials" (MSM 2010).

Kleinmann, S., Koscielny, J.M, Koller-Hodac, A., Paczynski, A., Stetter, R. (2010). Concept of an andvanced monitoring, control and diagnosis system for posistive displacement pumps., Accepted for SysTol 2010, Nice, 6th – 8th October 2010.

Kleinmann, S., Dabrowska, A., Hoffmann, M., Kühn, H., Koller-Hodac, A., Stetter, R.: Advanced Control of Industrial Pump Systems. In: Proceedings of the 7th Workshop on Advanced Control and Diagnosis, ACD'2009, Zielona Góra, Poland.

Korbicz, J., J.M. Kościelny, Z. Kowalczuk and W. Cholewa: *Fault Diagnosis: Models, artificial intelligence methods, applications*. Springer: Berlin, 2004.

Koscielny J..M., Syfert M., Wnuk P.: Advanced monitoring and diagnosis system "AmandD", In: Proceedings of SafeProcess, Beijing, 2006.

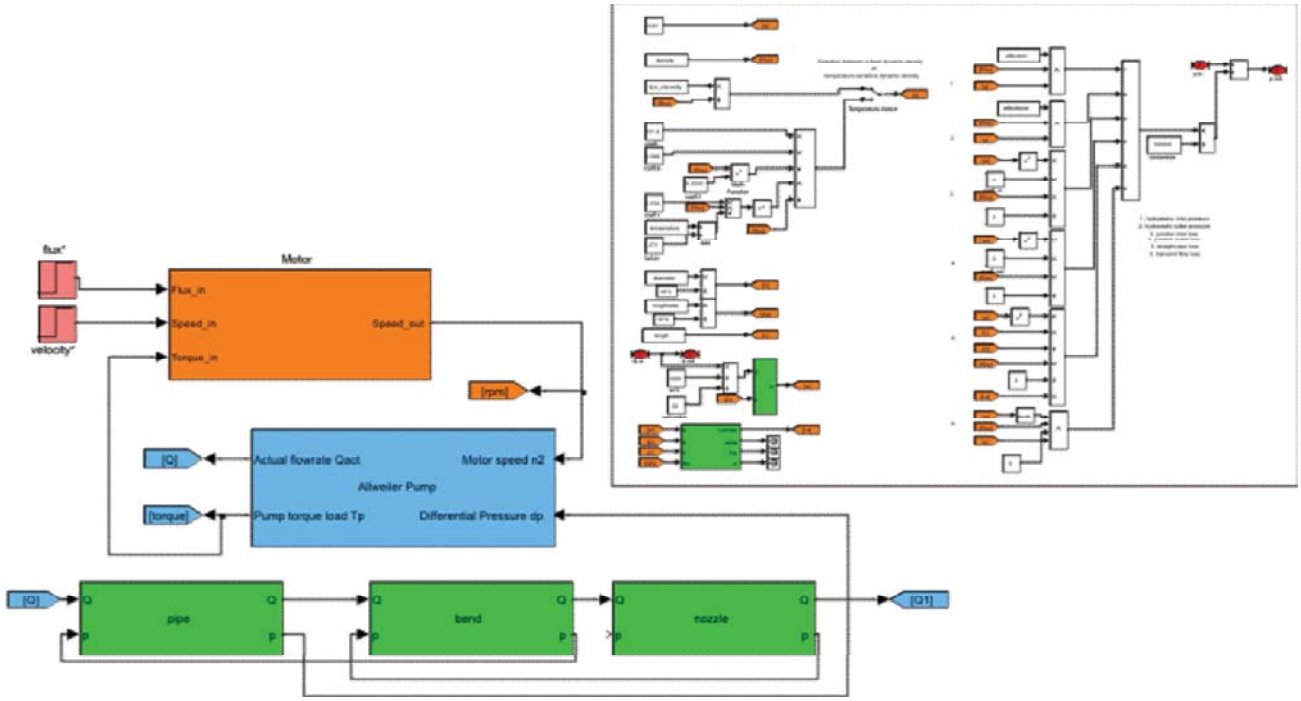
Takahashi I., Noguchi T.: A new quick response and high efficiency control strategy of an induction motor, IEEE Transactions on Industry Applications, vol. IA-22, no. 5, Sep/Oct 1986,s. 820-827

Wang, Y., Boyd, S.: Fast Model Predictive Control using Online Optimization. In: Proceedings of the 17th World Congress. The International Federation of Automatic Control. Seoul, Korea, July 6-11, 2008.

Witczak, M.: Modelling and Estimation Strategies for Fault Diagnosis of Non–Linear Systems: From Analytical to Soft Computing Approaches. Lecture Notes in Control & Information Sciences. Berlin: Springer 2007.

Zhang, Y., Jiang, J.: Bibliographical review on reconfigurable fault–tolerant control systems. Annual Reviews in Control, 32, 229–252, 2008.

Appendix A. Complete System in the simulation environment and example of implementation of pipes part in Matlab/Simulink



Appendix B. Risk Table

					Features in the pressure system				
No.	Possible fault	Possible reason of the fault occurrence	Danger / Consequence of the incident	Feature to identification	Feature and identification only with pump as a virtual sensor (without physical sensors)	Feature and identification with sensor configuration (S1)	Feature and identification with sensor configuration(S2)	Feature and identification with sensor configuration(S3)	Feature and identification with sensor configuration (S4)
1.1	Pressure fall in oil supply	Pipe break (F1)	Flow out of oil and under the circumstances and no burning, oil is sprayed to the air	Sudden loss of pressure. Consequently, sudden drop of torque on pump	[A] Pressure loss in the virtual sensor	[C] The same pressure loss in the physical and in the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor
1.2		Leak (F1)	Leak of oil		[A] Pressure loss in the virtual sensor	[C] The same pressure loss in the physical and in the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor
1.3		Empty oil tank	No flame in the burner and pump runs without lubrication	Sudden loss of pressure. Consequently, sudden drop of torque on pump	[A] Pressure loss in the virtual sensor	[C] The same pressure loss in the physical and in the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor
1.4		In-coming air	Temporarily, no flame in the burner	Sudden loss of pressure. Consequently, sudden drop of torque on pump	[A] Pressure loss in the virtual sensor	[C] The same pressure loss in the physical and in the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor	[E]Different pressure loss in the physical and the virtual sensor

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Evaluation Scheme of Task Allocation in Mesh Connected Processors with Metaheuristic Algorithms

Wojciech Kmiecik*. Leszek Koszalka*. Iwona Pozniak-Koszalka*. Andrzej Kasprzak*.

*Dept. of Systems and Computer Networks, Wroclaw University of Technology,
50-370 Wroclaw, Poland (e-mail: leszek.koszalka@ pwr.wroc.pl).

Abstract: This paper focuses on applying three metaheuristic local search algorithms to solve the problem of allocating two-dimensional tasks within a two-dimensional processor mesh in a period of time. The objective is to maximize the level of mesh utilization. To achieve this goal we adapt three algorithms: Tabu Search, Simulated Annealing and Random Search, as well as we design an auxiliary algorithm Dumb Fit and adapt another auxiliary algorithm named First Fit. To measure the efficiency of the algorithms we introduce our own evaluating function called Cumulative Effectiveness and a derivative Utilization Factor. Finally, we implement an experimentation system to test these algorithms on different sets of tasks to allocate. Moreover, a short analysis based on results of series of experiments conducted on three different categories of task sets (small tasks, mixed tasks and large tasks) is presented.

Keywords: Task allocation, algorithm, meta-heuristic, mesh structure, experimentation system.

1. INTRODUCTION

Recently, processing with many parallel units is gaining on popularity very rapidly. It is applied in various environments, ranging from multimedia home devices to very complex machine clusters used in research institutions. In all these cases, success depends on a wise task allocation e.g. Koszalka (2006), enabling the user to utilize the power of a highly parallel system. Research has shown, that in most cases, parallel processing units give only a fraction of their theoretical computing power e.g. Buzbee (1983) and Kasprzak (1999), which is a multiplication of the potential of a single unit used in the system. One of the reasons for this is high complexity of task allocation on parallel units.

Meta-heuristic algorithms have been invented to solve a subset of problems, for which finding an optimal solution is impossible or far too complex for contemporary computers. Algorithms like Tabu Search invented by Glover (1989) and Simulated Annealing proposed by Kirkpatrick (1983) and developed by e.g. Laarhoven (1987) and Granville et. all (1994) are among the most popular. They are capable of finding near-optimum solutions for a very wide range of problems in a time incomparably shorter than the time that it would take to find the best solution, e.g. Glover and Kochenberger (2002).

We decided to adapt three main algorithms for solving the allocation problem: the aforementioned – Tabu Search and Simulated Annealing as well as a simplified local search – Random Search used for comparison. In our approach, we decided that we would use also an existing solution for task allocation on processor meshes – the First Fit algorithm. The difference from typical approach is that it is not used as a solution by itself, but only as an algorithm to help evaluate the results of the three main algorithms (in each iteration). In our product we actually use two incarnations of the First Fit algorithm, see Goh and Veeravalli (2008). We consider one which is here named First Fit – it is the simplest form of First Fit and one which is here named Dumb Fit – it is actually a richer form of the classical First Fit, which enables reorganization of the task set. Our Dumb Fit is also used to generate results that we use as reference when examining the efficiency of the main algorithms. Finally, we had to invent an evaluating function for the main algorithms. We called it Cumulative Effectiveness. The function and its derivative Utilization Factor are further explained in following sections of the article. To examine our solutions' efficiency in various conditions (mesh sizes, task sizes, task processing times etc.) we implemented an experimentation system. It gives many possibilities to generate task lists, conduct series of tests and to save their results.

The rest of the paper is organized as follows. Section 2 exactly defines the problem to be solved, Section 3 describes the considered algorithms and pointed out their roles, Section 4 describes the experimentation system. In Section 5 the results of investigations are discussed. Finally, in Section 6 appear conclusions and final remarks.

This paper is a development of our previous research published in Kmiecik (2010). We have created a new version of algorithms, upgraded our experimentation system and conducted new, more detailed experiments.

2. PROBLEM STATEMENT

2.1 *Basic Terms*

A **node** is the most basic element which represents a processor in a processor mesh. It is a unit of the dimensions of a mesh, sub-mesh or task. Such node can be busy or free.

A **processor mesh**, which thereafter will be simply referred to as 'mesh', is a 2-D rectangular structure of nodes distributed regularly on a grid. It can be denoted as M (w, h, t)**,** where w and h are the width and height of the mesh and t is the mesh lifetime. The value of t may be zero or non-zero. A zero value means that the mesh will be active until the last task from the queue is processed. This value also determines the choice of evaluating function, which will be further explained later in this article.

A **position (x, y)** within a mesh M refers to the node positioned in column x and row y of the mesh, counting from left to right and top to bottom, starting with 1.

A **sub-mesh *S*** is a rectangular segment of a mesh M – a group of nodes, defined in a certain moment of time, denoted as $S(a, b, e, j)$ with its top left node in the position (a, b) in the mesh M, and of width e and height j. This entity, as a separate being, has only symbolic value. If a sub-mesh is occupied, it means that all its nodes are busy.

Tasks, denoted ***T*(*p*, *q*, *s*)**, are stored in a list. All the contents of the list are known before the allocation. Tasks are taken from the list and allocated on a mesh. There, they occupy a sub-mesh S of width p height q for s units of time (thus s is their processing time).

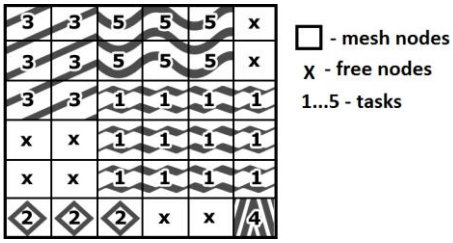


Fig. 1. A sample depiction of a mesh with 5 allocated tasks.

A mesh in a certain moment of time – $M(w, h, t_1)$, can be depicted as a matrix of integers, where each number corresponds to a node. Zero can be denoted as a X and it means a free node. Non-zero numbers (same for a sub-mesh processing one allocated task) indicate a busy node, their value is the time left to process the task. Such depiction is portrayed in Fig. 1. There, we can see five various tasks allocated on a small mesh.

2.2 *Evaluation Functions*

The main evaluating function introduced is the *Cumulative Effectiveness* (1). Knowing it and the parameters of used mesh we can count a more self-descriptive factor: the *Usage Factor* (2). In (1) p_i q_i and s_i are width, height and processing time of the i-th of n processed tasks. In (2) w, h, t are width, height and time of life of the used mesh.

$$CE = \sum_{i=1}^{n} \left(p_i \cdot q_i \cdot s_i \right) \tag{1}$$

$$U = \frac{CE}{w \cdot h \cdot t} \cdot 100\% \tag{2}$$

A task, as well as a mesh can be treated as 3D entities when we assume that time is the third dimension. Then CE function is the cumulative volume of all allocated tasks and U is the percentage of mesh's volume used by the processed tasks. It allows us to easily and objectively determine how much of the mesh's potential was "wasted" but how much utilized.

The creation of the CE function and derivative U factor is based on assumption that a company, using a processor mesh, has a set of tasks to process on their equipment, which exceeds the number of tasks possible to process in one atomic period of time (mesh's lifetime, e.g. a day), in the beginning of which a single allocation process is conducted. In such case it is essential to utilize as much of the mesh's power as possible – this would make its work most effective.

But there is also another approach, to testing allocation algorithms, in which we assume that lifetime of the mesh is unlimited, and it is desired to process all tasks in the list as soon as possible. Then, as the evaluating function, the *Time of Completion* (3) is introduced. In (3) t_{fin} is the moment of time, since the start of processing, when the last of all tasks has been executed.

$$T = t_{fin} \tag{3}$$

This factor can only be used for comparing algorithms, not for objectively evaluating their efficiency.

3. THE ALGORITHMS

3.1 *Basic Ideas*

In the paper, we implemented three meta-heuristic local search algorithms as the main algorithms: SA – *Simulated Annealing* explained e.g. by Laarhoven (1987), TS – *Tabu Search* explained e.g. by Glover (1989), RS – *Random Search (*not to be confused with simple evaluating of a random solution), explained in Kasprzak (1999). All of them work for a number of iterations. In each iteration, they operate on a single solution and its neighbourhood and evaluate the results. Fig. 5 shows the process of a single experiment with a main algorithm.

A **solution** is defined here as a permutation of tasks to be allocated, stored in a list. Such permutation is to be found using one of two atomic algorithms (*First Fit* and *Dumb Fit*) and calculating one of the evaluation functions (the index of performance) explained above.

There are also various kinds of neighbourhood to be explored by the main algorithms. We implemented two of them: *insert* and *swap*. In case of the first one, a neighbouring solution is found by taking one element of the permutation and putting it in some other position. In case of the second one, two elements are taken and their positions are swapped.

The success level for each of the 3 main algorithms highly depends on the instance of the problem (mesh's and task's dimensions and life/processing times) as well as on algorithms' specific parameters, the atomic algorithms used to initiate them as well as those used during their work.

(1) (1)

3.2 Random Search (RS)

RS is the simplest local search algorithm. In each iteration, it finds a new solution (calculating the chosen index of performance) from the neighbourhood of the current one.

In the next iteration, the new one becomes the current one and the process continues. In RS there are no additional parameters except for the number of iterations. This algorithm is highly resistant to local minima but it does not improve certain solutions too precisely.

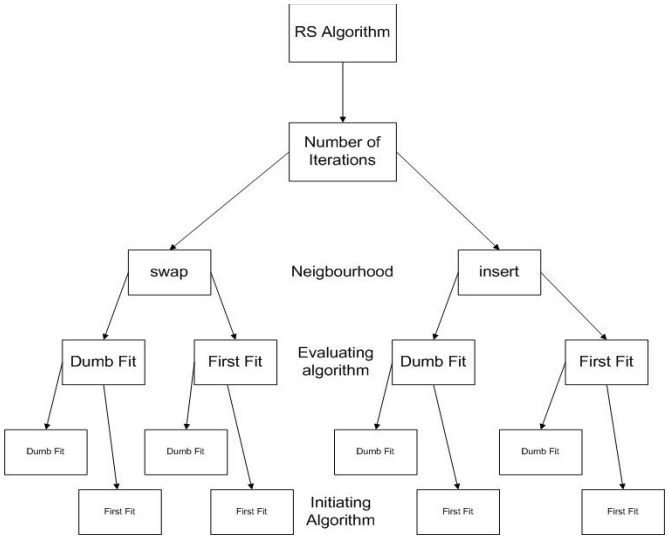


Fig. 2. Block-diagram of *RS chain algorithm.*

3.3 Simulated Annealing (SA)

SA works in a more complex way. Its main parameters are initial and final temperatures. During the course of its work the temperature drops (logarithmically or geometrically). In each iteration, a random solution from the neighbourhood of the current one is found and evaluated. When the temperature is high there is high probability to accept the new solution as the current one, even if it is worse. When the temperature is low only these solutions are accepted as new current ones which are better. Such approach makes this algorithm resistant to local minima in the beginning and precisely improving a current solution in the end, going down to the nearest local minimum.

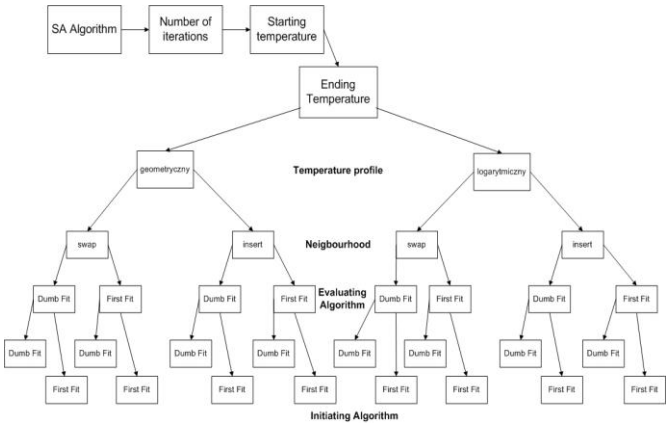


Fig. 3. Block-diagram of *SA chain algorithm.*

3.4 Tabu Search (TS)

Our implementation of the TS algorithm is similar to the SA algorithm with low temperatures, except for the fact that it does not accept a new solution as the current one, if the same solution is in the taboo list. Whenever a new current solution is set it is added to the taboo list. The taboo list has limited length which is the main parameter of the algorithm. This algorithm is forced to leave the vicinity of a local minimum. This vicinity is limited by the length of the taboo list. At the same time TS tries to precisely improve a current solution.

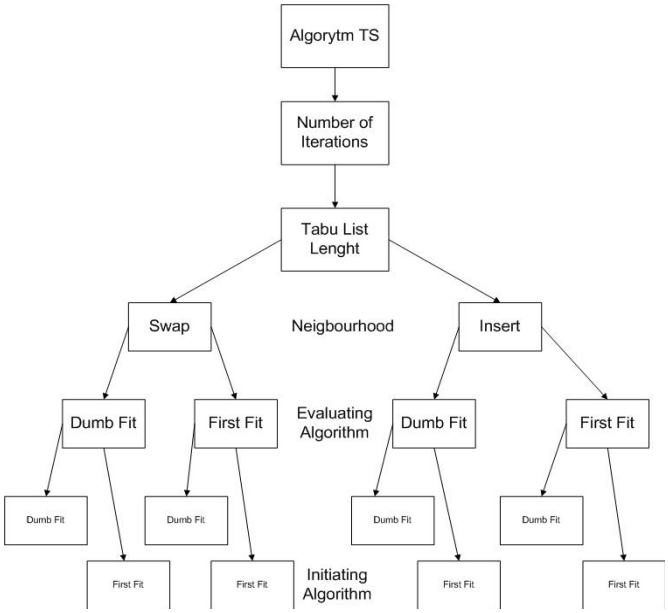


Fig. 4. Block-diagram of *TS chain algorithm.*

3.5 Atomic Functions

The atomic algorithms that we use are: FF – *First Fit* and DF – *Dumb Fit*. Their block-diagrams were presented in our previous paper – Kmiecik (2010).

The FF takes a solution and does not modify the order of task permutation. It scans through the mesh from top to bottom, left to right. If it encounters a free node, it checks whether there is enough free nodes right from it and below it, to allocate the first task from the list. If this try fails, it rotates the first task from the list by ninety degrees and tries to fit it again. If it succeeds, the task is taken from the list, a corresponding sub-mesh is allocated and the algorithm keeps scanning the mesh and tries to allocate next tasks until the mesh ends. This process is repeated in each moment of mesh's lifetime or until the last task is allocated (if mesh's lifetime is not limited).

The DF algorithm works very similarly to FF, except for the fact, that upon encountering a free sub-mesh, it does not limit itself to trying to allocate only the first task from the list, but tries each of them. Therefore, it can modify the permutation.

During DF or FF algorithm's work, the appropriate evaluating function value is calculated and then it is returned to the one of the three main algorithms that is currently performed.

3.6 Concept of chain algorithms

Large number of parameters of algorithms and problems with representing such structure inspired us to create chain form of metaheuristic algorithms. The idea of chain algorithms is showed on Fig. 2,3 and 4. Consistently we use them in our work, for example in batch files, where algorithms are defined by chains of parameters:

```
RS 30000 insert Dumb_Fit Dumb_Fit
TS 15000 100 swap First_Fit Dumb_Fit
```

4. EXPERIMENTATION SYSTEM

Our goal was to design such simulation environment that would be able to evaluate all combinations of parameters for various problem instances. As a result we have developed a GUI application, written in the C++ language using QT framework, with various abilities to read parameters for the experiments and to write their results. The created application named CATAM can be run under Microsoft Windows OS. It has two main working modes - Single Experiment and Command Line Mode.

Experiment design:

- Task parameters
- Mesh parameters
- Main allocation algorithm used
- Atomic algorithm for generation of a initial solution
- Atomic algorithm for evaluating a solution
- Inner parameters of allocation algorithms used
- The number of iterations

Initialization
Generating initial solution
Loop of the algorithm for a given number of iterations:

- Getting a new solution from the neighborhood
- Solution processing and evaluation
- Comparison of the result with the best solution

Returning the results

Fig. 5. A single experiment concept.

4.1 Input Variables

Generally, in all modes of operation, our software allows the user to set certain input parameters. First group of them defines the problem. It allows the user to choose ranges of dimensions (p, q) and processing times (s) for the tasks and the task-list length. The user can also define the size and lifetime of the mesh: w, h, t. All the parameters from the first group allow the program to randomly create a task-list and define a mesh, which, together, form a problem instance.

The other group consists of specific parameters of the chosen algorithm like: the number of iterations, initial and final temperatures for SA, temperature profile for SA, taboo list length for TS, etc. Specifying both groups of parameters makes it possible to solve a predefined problem with a chosen, configured algorithm.

4.2 Single Experiment Mode

It allows the user to specify all the parameters easily without using an input file and to watch the algorithm work (it's progress and Usage factor value is shown). This mode is created to give the user a possibility to check how metaheuristic algorithms respond to different parameter sets.

4.3 Batch File Mode

This mode is the preferred one for running a series of experiments for a certain research. It allows the user to specify path to a file with a pre-designed test series so-called multistage experiment design, see Pozniak-Koszalka (2006). Such file begins with a set of parameters defining the problem instance. Also the number of repetitions for each test is specified. When using the command line mode, the user can create a batch file (*.txt*) for a series of series of tests.

4.4 Outputs

For a single experiment (see the last module in Fig. 4), in either execution mode, an output file contains a line with used parameters for the experiment and lines showing Cumulative Effectiveness and Usage Factor values for each experiment.

5. RESEARCH

5.1 *Experiment Design*

Firstly, a preliminary experiment for finding the best parameters for each considered metaheuristic algorithm, was carried out. Then, three distinct cases were checked for different categories of tasks, including relatively small tasks (as compared with the mesh size), big tasks, and mixed tasks (a composition of small and big tasks). In each test, an obtained result with Dumb Fit was used as a level of reference..

Table 1. Experiment Design

Parameter	**Test**		
	Mixed tasks	***Small tasks***	***Large tasks***
p	2÷12	2÷6	6÷12
q	2÷12	2÷6	6÷12
s	2÷12	2÷6	6÷12
w	12	50	12
h	12	50	12
t (task-list length when t=0)	1000	0 (1000 tasks)	1000
tested algorithms	SA, TS, RS	SA, TS, RS	SA, TS, RS
evaluating algorithms	FF	FF	FF
initiating algorithm	DF, FF	DF, FF	DF, FF
evaluating function	*CE*	*T*	*CE*
neigbourhood	*swap*, *insert*	*swap*, *insert*	*swap*, *insert*

5.2 *Preliminary Experiments*

First of our preliminary experiments was to find more effective neighbourhood. Results are shown in Fig. 6 :

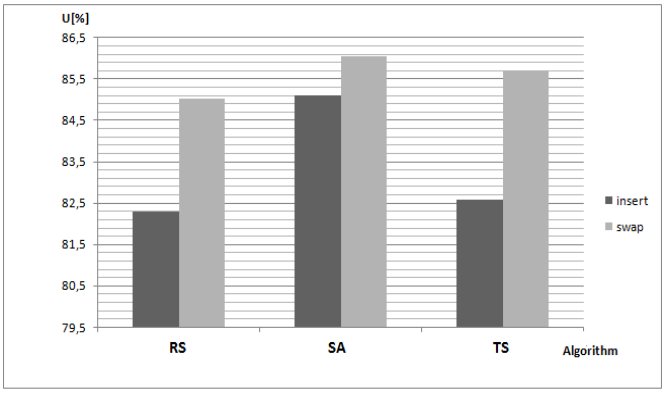


Fig. 6. Performance of the metaheuristic algorithms for swap and insert neighbourhoods.

Swap neighbourhood gives better results than insert. For both Random Search and Tabu Search difference was 3% and for Simulated Annealing it was 1% Usage factor.

After analysis of the results of other experiments we observed that:

- 30000 should be chosen as number of iterations for the algorithms. Above that number results are improving very slow, but time of processing increases linearly.
- The best atomic algorithm for setting a list of tasks for RS and TS algorithms is Dumb Fit (3% better results than First Fit). For SA algorithm First Fit gives better results (0,2%) than Dumb Fit.
- We chose First Fit as fitting parameter, because Dumb Fit gives better results (4%) but also increases the time of processing by over 1000%.
- For SA algorithm the best parameters were : geometrical temperature profile and initial temperature T=250.
- For Tabu Search algorithm, the best Tabu List length was 1500.

5.3 *Mixed Tasks*

In this case almost all tests were performed for 30000 iterations (except for a few with 5000 iterations) for all main algorithms for the same task set. For tests in which DF was the evaluating algorithm, which significantly increases evaluation time, 5000 iterations were tested. The aim was to keep all algorithms running for about 100 seconds. Each test was repeated 100 times and the mean values are used below unless it is specified otherwise. The evaluating function used was CE which allowed counting the Usage Factor.

Results obtained:
(i) the best result: SA, swap,
evaluation function FF,
initial temperature=250, geometrical profile
U=84.05%,
(ii) the difference between the best result and the result obtained by DF is 9,64%.

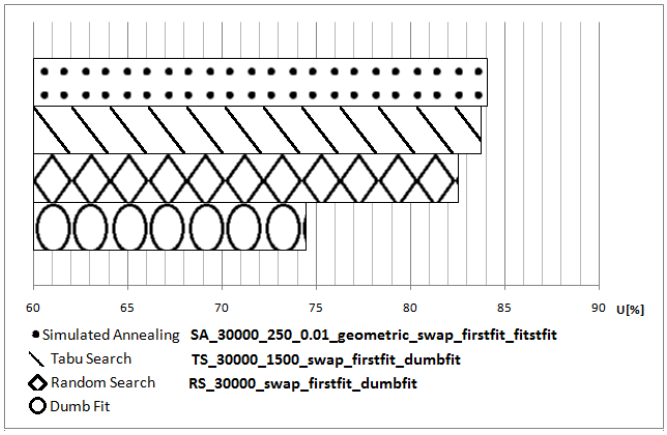


Fig. 7. Performance of the metaheuristic algorithms and Dumb Fit for mixed tasks.

Observations:
It may be observed, after analysis of the results of series of experiments, that:

- SA algorithm performed best outperforming TS, RS and single DF execution (see Fig. 7).
- The main factor affecting the effectiveness of SA was the initial temperature.
- It is a good idea to use DF as a generator for the initial permutation in the main algorithms. Nevertheless, this does not apply to SA: mean *U* value for the same settings, for SA (T0=250, *swap* neighbourhood) starting from a random permutation was 84.05%, but when starting from the permutation generated by DF it was 83.86%.

5.4 *Small Tasks*

In this case we decided to use the second evaluating factor *T* expressed by (3). It is less productive than (1) but still allows comparing the algorithms and gives much better ability to spare experimentation time. It is so, because for a large mesh and small tasks it would be needed to process a huge list of tasks, so as not to run out of them in e.g. 1000 units of mesh's lifetime. This would make a series of experiments very long to conduct in our conditions. Moreover, due to semi-random characteristics of the tested metaheuristic, ones that started from a random solution and did not use FF for evaluation, gave even worse results, e.g. SA, in such case, gave a result of *T*=124.

Results obtained:
(i) the best result: algorithms SA/TS/RS: *T*=98,
(ii) the difference between the best result and FF: 0.

These experiments also gave a conclusion, namely one, saying that using metaheuristic algorithms for allocating small tasks does not make sense. Tasks are here small enough, in comparison to the size of the mesh that the FF algorithm manages to fit a task from the list into almost every free sub-mesh. Therefore, even metaheuristics based on FF's cannot achieve any better result (example plot in Fig. 8).

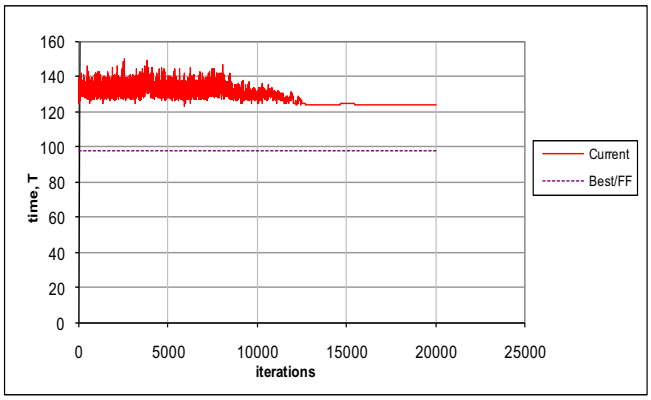


Fig. 8. The values of current and best results for small tasks.

5.5 Large Tasks

In this case, achieved results and behavior of algorithm were very similar to the general case of mixed tasks (we also used the same scheme of testing as then). The achieved result of the winning algorithm was marginally better (U=85.44). Also, as in the case of mixed tasks, SA algorithm was the best and the same parameters (as for mixed case) caused the best performance.

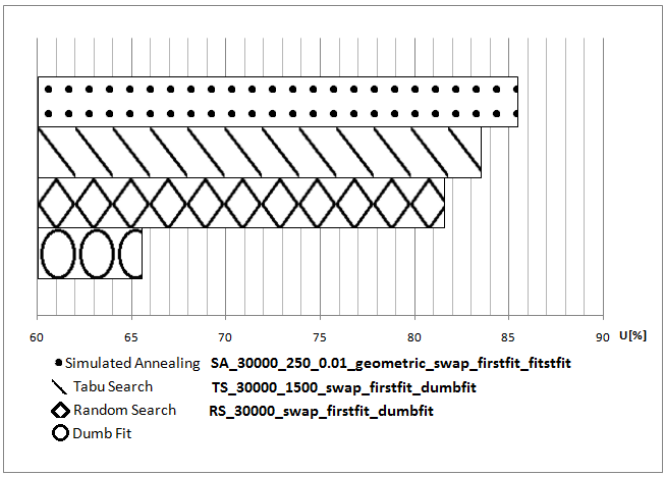


Fig. 9. Performance of the metaheuristic algorithms and Dumb Fit for big tasks.

Results obtained:
(i) the best result: SA, swap,
evaluation function: FF,
initial temperature= 250, geometrical profile
U=85.44%.
(ii) the difference between the best result and the result obtained by DF is 19.94%.

6. CONCLUSIONS

In the paper, there are described and discussed: (i) three implemented metaheuristic local search algorithms for task allocation on a processor mesh and two algorithms for generating their initial conditions, (ii) the created experimentation system for testing the algorithms, and (iii) results of the experiments.

The experiments showed that in general, local search metaheuristic algorithms are a good tool for solving the considered problem. Only for allocating small tasks on a large mesh, it is needless to use them as they do not achieve better results than the classic FF algorithm which itself performs well, due to the easiness of fitting small tasks into free sub-meshes. The leader of all our tests was the SA algorithm. It defeated all others for tasks of mixed and large sizes. It also achieved reasonable results of over 85% of mesh usage, which we find quite satisfactory, even though it was achieved in a relatively small number of iterations. The implemented experimentation environment allows the user to easily design whole series of experiments and to check many combinations of parameters.

In the further research in this area we are planning to construct far more thorough and versatile testing environment and to implement more algorithms, e.g. the Genetic Algorithm, Ant Algorithm, etc..

REFERENCES

Buzbee, B.L. (1983). The Efficiency of Parallel Processing. In *Frontiers of Supercomputing*, Los Alamos.

Glover, F. (1989). Tabu Search – part I. *ORSA Journal on Computing*, vol. 1, no. 3.

Glover, F. and Kochenberger, G.A. (2002). *Handbook of Metaheuristics*, Springer, Heidelberg, New York.

Goh Lee Kea and Veeravalli, B. (2008). Design and Performance Evaluation of Combined First-Fit Task Allocation. *Parallel Computing*, vol. 34, pp. 508-520.

Granville, V., Krivanek, M., and Rasson, J.P. (1994). Simulated Annealing: a proof of convergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 652-656.

Kirkpatrick, S., Gelatt, C.D., and Vecchi, M.P. (1983). Optimization by Simulated Annealing. *Science, New Series*, vol. 220, pp. 671-680.

Kmiecik, W., Wójcikowski, M., Koszałka, L., Kasprzak, A. (2010). Task Allocation in Mesh Connected Processors with Local Seach Meta-heuristic Algorithms. *Lecture Notes in Computer Science* vol. 5559, Springer, pp. 215-224.

Koszalka, L., Lisowski, D., and Pozniak-Koszalka, I. (2006). Comparison of Allocation Algorithms for Mesh-Networks with Multistage Experiment. *Lecture Notes in Computer Science*, vol. 3984, Springer, pp. 58-67.

Koszalka, L., Kubiak, M., and Pozniak-Koszalka, I. (2006). Allocation Algorithm for Mesh-Structured Networks, *Proc. of 5th ICN,* IEEE Comp. Society Press, pp. 24.

Laarhoven, J.M., Emile, H., and Aarts, L. (1987). *Simulated Annealing: Theory and Applications*, Springer, Berlin.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Bus Route Optimization: an Experimentation System and Evolution of Algorithms

Krzysztof Golonka*, Leszek Koszałka*, Andrzej Kasprzak*

*Dept. of Systems and Computer Networks, Wroclaw University of Technology,
50-370 Wroclaw, Poland (e-mail: leszek.koszalka@ pwr.wroc.pl).

Abstract: In this paper, we analyze the school bus route optimization problem. It is a crucial social issue that concerns faster and more comfortable transport of students to their schools. Moreover, the route optimization allows to decrease the ticket price, i.e., to maximize the profit of the provider. Since the problem belongs to hard optimization problems, thus, we adapted four meta-heuristic algorithms: Tabu Search, Simulated Annealing, Genetic Algorithm, Complete Overview, and invented by the authors algorithm called Constructor, and additionally Bellman-Ford algorithm used as a helper. In order to measure the efficiency of the considered algorithms we create our own evaluating function called Balance and compare results given by algorithms to the maximum found with Complete Overview. Finally, we designed and implemented an experimentation system to test these algorithms on various problem instances, to emerge the most efficient one.

: Algorithms, meta-heuristic, optimization, experimentation system, transport.

1. INTRODUCTION

Since banking crisis from 2008 many companies were forced to cut expenses and look for more savings. Moreover, nowadays a lot of pressure is put on being green – environmentally - friendly, especially when it comes to industry or transport e.g. Składzien (2008). A lot of effort must be put in analysis and planning process to come across these challenges. This paper focuses on optimization of a school bus route and proposes the direction of searching optimal solutions. The main goal is to find the most profitable route (e.g. the shortest path).

This optimization belongs to non-polynomial problem and has a huge solution space, meaning we can not find the best solution in polynomial time. For small instances it is easy to search through the whole solution space but when instance begins to grow, the required time may become unacceptable. The only one reasonable way to solve this is to use meta-heuristic algorithms that were invented to struggle with such problems. An idea to consider such algorithms based on artificial intelligence like Tabu Search (e.g. described in Glover, 1997), Simulated Annealing (e.g. explained in Laarhoven, 1987) and Genetic Algorithm (e.g. illustrated in Davies, 1987) seems to be promising.

We decided to adapt all three mentioned ideas for implementing algorithms to find an efficient solution to bus route problem. In addition, we tried to invent on our own a new algorithm and we created the algorithm called Constructor which is described in this paper. To determine the shortest path from the starting point to the ending point through all possible bus routes we implemented Bellman-Ford Algorithm.

Moreover, we implemented Complete Overview (CO) algorithm to be able to calculate differences between maximum and optimum found using meta-heuristics (unfortunately, CO may be used only for instances smaller than 20 bus routes because of its complexity and the required time). Finally, we had to introduce an evaluating function as the measure of efficiency for the considered algorithms.

To assess algorithms' efficiency we implemented an experimentation system that allows user to perform series of tests, returns the average values and presents them on plots.

The rest of paper is organized as follows: Section 2 defines the problem to be solved. Section 3 describes the considered algorithms and their roles in optimization process. Section 4 shortly presents the implemented experimentation system. The results of the research appear in Section 5. Last but not least Section 6 provides some remarks and conclusion.

2. PROBLEM STATEMENT

There is given a certain urban area (Fig. 1), that consists of links and Bus Stops (BSs). Lines represent links, numbers next to them represent their lengths and red dots symbolize Bus Stops. The beginning of the route is marked by a green rectangle, but ending point (rectangle) represents the location of the school. It is necessary to determine the most profitable constant for a certain time period route of a school bus to maximize profits of a bus provider. This means that the route may consist only BSs at which the number of pupils waiting for a bus is sufficient not to make loses. Once the route is planned the bus may omit some BSs which are not on this specified route.

The basis for making decision on which BS should stop is the observation of statistics that deliver the number of students (pupils) waiting at a BS on a given time. These values are represented by a matrix, in which each row corresponds to the next hour in bus transit and the columns show the number of pupils (Table 1).

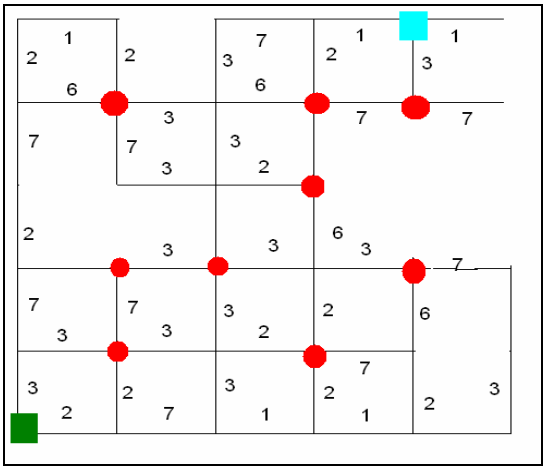


Fig. 1. An example of an instance.

Table 1. Students (pupils) statistics.

Bus Stop / Time	#1	#2	#3	#4
8:00am	5	5	16	9
8:45am	6	1	3	11
9:30am	10	22	4	9
10:15am	0	2	0	0

The input parameters are listed in Table 2.

Table 2. Input parameters.

Sign	Parameter
SBS	Set of potential BSs
(x_i,y_i)	BS_i coordinates
$L_{i,j}$	Link length between i and j BS
$P_{k,j}$	Pupils at k BS at j transit
T	Ticket price
DC	Driver's cost
BC	Bus exploitation cost
J	Number of transits

2.1 *asic Terms*

Bus Stop (BS) is a point on a map where pupils wait for a bus to school. Each BS has its coordinates x and y on a map.

$$S = [x, y] \tag{1}$$

Set of Potential BSs (SBS) is a collection of all BSs on a map, where SBS(i) may be defined by (2).

$$S\ S(i) = \begin{bmatrix} S_1 \\ S_2 \\ \text{M} \\ S_i \end{bmatrix} = \begin{bmatrix} x_1, y_1 \\ x_2, y_2 \\ \text{M} \\ x_i, y_i \end{bmatrix} \tag{2}$$

Link (L) is a matrix defined by (3) that describes lengths of links between BS_i and BS_j.. Some BSs may not be directly linked to others but each BS must have at least one link.

$$(i,\ ) = \begin{bmatrix} l_{1,1} & \Lambda & l_{1,} \\ \text{M} & \text{O} & \text{M} \\ l_{i,1} & \Lambda & l_{i,} \end{bmatrix} \tag{3}$$

$l_{i,j}$ is defined as:

$$l_{i,} = \sqrt{(x_i - x\ )^2 + (y_i - y\ )^2} \tag{4}$$

Route (R) is a path consisting starting point, ending point and going through BSs chosen from SBS. A Route may contain smaller amount of BSs than SBS.

$$R = S\ \ S(k) \qquad \text{where } k \le i \tag{5}$$

Ticket Price (T) informs how much each pupil must pay for taking a bus.

Driver's Cost (DC) equals money paid to a driver for driving one unit of a length of a Route.

Bus Exploitation Cost (BC) equals expenses for fuel used by a bus after driving one unit of a length of a Route.

Number of Transits (J) says how many times the bus is going a route per a day.

Route Length (RL) is a length of a shortest path.

2.2 *Evaluating Function*

The evaluating function introduced by us is called the Balance (6) interpreted as a daily balance – obtained after one day of work. If $Q\ R$ is less than 0 that the provider gets loses, if greater than 0 that the provider gets profits.

$$Q(R) = \sum_{=1} \left(\sum_{k=1}^{K} P_{k,} * T - R\ \ * (DC + \ \ C) \right) \tag{6}$$

Moreover, to make sure that the bus will not go from starting point right to the destination point, we introduce a constraint. The constraint describes the minimal percentage number of all pupils from the statistics (Table 1) that should be delivered to the school (7).

$$100\ \% * \frac{\sum_{=1} \sum_{k=1}^{K} P_{k,}}{\sum_{=1} \sum_{i=1}^{I} P_{k,i}} \ge P_{\%} \tag{7}$$

3. THE ALGORITHMS

3.1. asic Ideas

For experiment purposes we implemented four meta-heuristic algorithms as the main algorithms, including three known algorithms (but specially adopted) : TS - Tabu Search, SA - Simulated Annealing, GA - Genetic Algorithm, and the Constructor (originally proposed by the authors of the paper). The first two of them are described e.g. in Wroblewski (2003), our adaptation of GA as well as the Constructor are explained below. The Route Length is calculated by us using Bellman-Ford algorithm. TS, SA, and GA perform calculations for a certain number of iterations processing on a solution and its neighbourhood.

A solution is defined as a RL (see 2.1) and the neighbourhood is a set of Routes with one different BS, without one BS or with additional one BS.

The performance of each meta-heuristic algorithm is affected by a few factors such as an instance parameters (the size of SBS) and algorithms' inner parameters.

3.2. Simulated Annealing SA

In each iteration the solution is replaced by a new one randomly chosen from the neighbourhood if the new one is better. If the new solution is worse it has 50% of chances to replace the previous one. In this particular implementation we do not have such thing as temperature that changes the probability of replacing solutions. Here probability is constant. This is the only one difference between our SA and the one described in Laarhoven (1987).

3.3. Tabu Search TS

Tabu Search is more complex algorithm than SA because it is searching through the whole neighbourhood of a solution and choosing the best one unlike SA. Moreover TS is more resistant to loops thanks to the taboo list. The best found solution may replace the previous one only if it differs from all the records in a taboo list more than a certain percentage value. The length of the taboo list is limited and when it is full, the old records are overwritten.

3.4. enetic Algorithm A

This algorithm is based on evolutionary mechanisms. The main idea is to create a population of a constant size and observe its evolution meanwhile registering the best ones. The most interesting here is a cross-over process. It requires two individuals and eventually gives two children.
DNA chain is represented in this situation as a single solution. The crossover is described below on an example:

parent no1: 1001|1101
parent no2: 1100|0111

child no1: 1001|0111
child no2: 1100|0111

All the solutions have a chance to hand over gens but the higher the Price function value of the solution, the higher possibility of being picked as a parent. Moreover, each child may mutate with probability 50% that leads to changing only one randomly picked chromosome. As the population size is constant, the newborns must replace the old solutions regardless of their breeding history.

The algorithm implemented in our specified problem is described as follows:

Step 0. Initial population is picked randomly.
Step 1. Pick parents randomly from existing population.
Step 2. Perform breeding process.
Step 3. Choose solutions to extinct.
Step 4. Add children solutions to the rest of solutions in existing population.
Step 5. Check if the optimization function of each solution in current population is not the best global optimum from already explored solution space.
Step 6. Go back to step 1 as many times as the number of iterations.

3.5. The Constructor

The Constructor was invented by us - an inspiration was the idea of searching through a whole neighbourhood while one iteration as in TS. Moreover, we assumed that splitting a big instance to smaller ones, solving them separately and joining all together may come up with quite good results.

At the beginning the Constructor splits the whole instance to smaller instances - containing two BSs, the beginning BS and the last BS. The next step is to search through a neighbourhood of each small instance and modify them. After this operation, the algorithm combines in pairs small instances making them bigger. These operations last until the joining gives back the initial instance.

Step 1. For specified short route that consists of beginning BS, ending BS and temporary amount of BSs find new possible routs picked from the short route's neighborhood.
Step 2. Check the optimization function for solutions, which include new route. Find the best neighbor.
Step 3. Modify current solution by including that best neighbor as a part of the solution.
Step 4. Pick next temporary amount of BSs and go back to step 1, unless all of BSs has already been picked.
Step 5. Double the temporary amount of BSs and go back to step 1, unless the temporary amount of BSs exceeded amount of all BSs.

3. . ellman-Ford F

The well-known (e.g. Wroblewski, 2003) algorithm BF is used to calculate and to determine the shortest path from the beginning right to the destination point through chosen BSs.

4. EXPERIMENTATION SYSTEM

The application was designed, mainly in order to visualize the tested algorithms. The application was created using Visual Studio 2008. The implementation language was C#. Figure 2 shows screenshot of the application window.

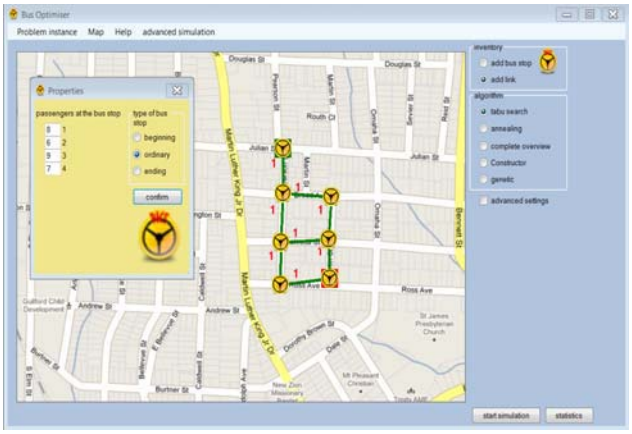


Fig. 2. Application window.

In the beginning the user defines an instance of problem by putting BSs on the map and creating links between them. Next, the beginning and the ending points of the potential route are precised and pupil statistics (by clicking and selecting properties) is determined. Finally, the user selects the considered algorithm and fixes its parameters. After clicking on "start simulation" button the application is searching for an optimal solution. There is a possibility to see an animation of a transit, and some statistics presented also on plots (e.g. served BSs, Balance, the percentage of served pupils) that allows observing results from each iteration.

5. RESEARCH

5.1 *Calibrating algorithms*

The first part of research refers to finding the best parameters' values for two algorithms:

- Tabu Search,
- Genetic Algorithm

For each new set of parameters, a new simulation was made. 10 instances were tested, each instance of problem was tested 10 times. Values in Tables 4 and 5 under all plots are averages from whole test. The parameters of the problem for the considered instances are shown in Table 3.

Table 3. Parameters – set 1.

Bus Stops	15
Test iterations	10
Instances to test	10
[%] passengers	50
driver's cost	1
bus exploatation cost	1
ticket price	3
number of transits	4

5.1.1 *enetic Algorithm*

Tests proved that increasing size of the population as well as increasing number of parents do not have a remarkable influence on improvement of solution (only about 5%). But the number of iterations has vast impact on results (see Table 4), where for 60 iterations it differs from the maximum just of approx. 18%. Because of slight differences in results between 60 and 40 iterations, the optimal value of this parameter was taken as equal to 40 in order to minimize the required estimation time.

Table 4. Results given by GA.

test no.	GA	CO	ΔGA [%]	population	parents	iterations
1	1360	1990	46,32	10	6	20
2	1384	1971	42,41	20	6	20
3	1292	1765	36,61	20	10	20
4	1243	1726	38,86	20	14	20
5	1215	1728	42,22	20	18	20
6	1256	1610	28,18	20	18	30
7	1307	1573	20,35	20	18	40
8	1259	1489	18,27	20	18	60

5.1.2 *Tabu Search*

As shown in Table 5, changing only two parameters make noticeable difference in precision and taboo list length.

Table 5. Results given by TS.

test no.	Tabu Search	CO	ΔTS [%]	Tabu list length	% Precision	Iterations
1	1328	2255	69,80	4	10	20
2	1361	2074	52,39	4	5	20
3	1263	1571	24,39	4	2	20
4	1416	1592	12,43	4	1	20
5	1489	1595	7,12	4	1	30
6	1329	1352	1,73	4	1	50
7	1185	1213	2,36	8	1	50
8	1445	1505	4,15	16	1	50

Decreasing precision to 1% causes improvement of results in comparison to parameters from the first test. Apparently, smaller precision allows TS to make smaller but more frequent steps. This means, that TS explores larger solution space and it is obvious that in this situation the probability of encountering better result is higher. The vital parameter is the number of iterations - along with the same as in genetic algorithm property: the more iterations, the better solution.

5.2 *Comparison – all algorithms*

The next part of research was to compare to CO all three other algorithms with the best inner parameters. Instances of parameters were the same as in previous part (Table 5) apart from minimal percentage passengers served (% passengers) which is variable in this test (Table 6).

Table 6. Parameters – set 2.

TS		SA		GA	
Iteration	50	Iteration	50	Iteration	40
Tabu length	4			Population	20
Precision	1			Parents	18

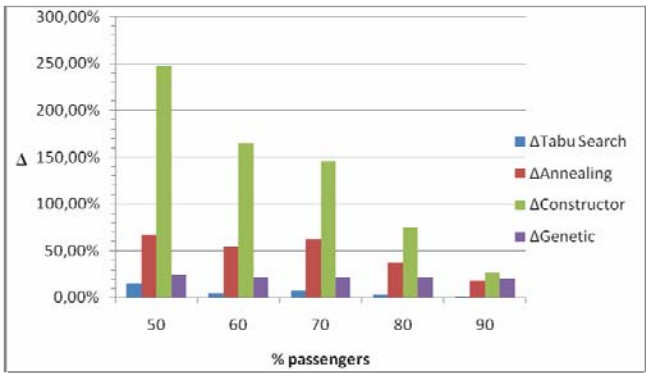


Fig. 3. The average inaccuracy.

According to Fig. 3 the smallest inaccuracy was found for TS - from 10% to less than 1% for 90% passengers. The second efficiency takes GA - from more than 50% to about 10%. Third was SA and the last place for Constructor. Constructor presents the biggest inaccuracy for 50% passengers (almost 250% inaccuracy) but its performance improves when constraint becomes more strict - 90 % passengers. Despite this surprising one result, the rest leaves a lot to wish and disappointed us.

5.3 *TS and SA comparison*

The influence of the number of iterations is shown in Fig. 4.

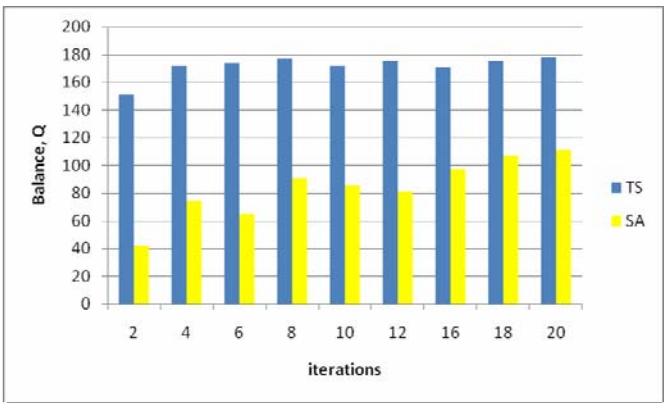


Fig. 4. Comparison of alghoritms: TS vs SA.

This test justifies an observation(rather obvious) that the number of iterations has significant impact on the obtained results. The more iterations, the better results is given by the algorithm.

The main observation is that TS gives better results than SA regardless of the number of iterations, thus TS algorithm may be recommended for searching the optimal route.

5.4 *TS and A confrontation*

The next experiment was related to observing differences between the two metaheuristic algorithms: TS and GA with their best parameters apart from iterations which was a variable. The rest of parameters used were such as specified in Table 3.

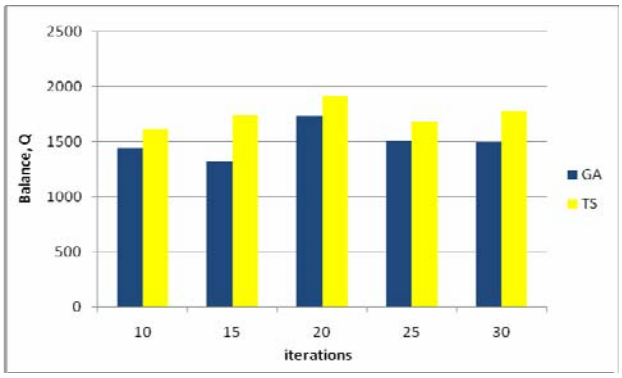


Fig. 5. Comparison of algoritms: GA vs TS.

Similarly to the previous test, TS defeats competitor - GA regardless of iterations number. Although TS wins this competition, the genetic algorithm GA kept pace of TA and the results given by GA were not that bad as in SA case (see Section 5.3).

6. CONCLUSIONS

To sum up, performed research justified the conclusion that TS algorithm gives much better results than SA regardless of defined advanced settings for searching the best solution. SA may give quite good results but much more iterations are needed. The only one algorithm that can compete with TA is GA but the average results of tests show that it would rather never come up with better results than TA. The Constructor algorithm turns out to be the worst and certain improvements are needed to make it somehow useful.

Choosing the best algorithm is half the success, however, setting the most appropriate parameters of such algorithm is a vital issue.

The main goal for solving school bus problem is to provide an opportunity to every single pupil to reach school on time. According to this statement and research presented in this paper, the proposed and recommended algorithm for route planning is Tabu Search (TS).

REFERENCES

Davies, L.D. (1987). *enetic Algorithms and Simulated Annealing*, Morgan Kaufmann Publ.

Gendreau, M. (2003). *An Introduction to Tabu Search*, Universite de Montreal.

Glover, F. (1997). Tabu Search – part I. *ORSA ournal on Computing*, vol. 1, no. 3.

Glover, F. and Kochenberger, G.A. (2002). *andbook of Metaheuristics*, Springer, Heidelberg, New York.

Granville, V., Krivanek, M., and Rasson, J.P. (1994). Simulated Annealing: a proof of convergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 652-656.

Jaszkiewicz, A (1990). *Multiple ob ective metaheuristic algorithms for combinatorial optimi ation*, Poznan.

Kirkpatrick, S., Gelatt, C.D., and Vecchi, M.P. (1983). Optimization by Simulated Annealing. *Science, New Series*, vol. 220, pp. 671-680.

Laarhoven, J.M., Emile, H., and Aarts, L. (1987). *Simulated Annealing: Theory and Applications*, Springer, Berlin.

Skladzien, J. (2008). *Ecological aspects of vehicle transport development*, Opole /in Polish/.

Wroblewski, P. (2003). *Algorithms: data structure and programming technologies*, WNT, Warsaw /in Polish/.

Youssef, H. and Sait, S.M. (1997). *Iterative Computer Algorithms with Applications in Engineering*, Washington.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Routing in Mobile Ad-hoc Networks: an Experimentation System and Evaluation of Algorithms

Maciej Foszczynski*, Marek Adamczyk*, Kamil Musial*, Leszek Koszalka*, Iwona Pozniak-Koszalka*, Andrzej Kasprzak*

*Dept. of Systems and Computer Networks, Wroclaw University of Technology,
Wroclaw, Poland (e-mail:leszek.koszalka@pwr.wroc.pl)

Abstract: The paper concerns the problem of path finding in wireless ad-hoc networks. Several algorithms, including meta-heuristic algorithms, evolutionary algorithm and the created hybrid algorithm, are considered. Algorithms have been implemented into a designed experimentation system. The system allows making simulation experiments along with multistage experiment design. In the paper, the results of some experiments are discussed. Moreover, the comparative analysis of efficiency of algorithms is presented. It may be concluded that the proposed hybrid algorithm seems to be promising.

Keywords: Ad-hoc network, path finding, meta-heuristic algorithms, experimentation system

1. INTRODUCTION

Mobile wireless ad-hoc networks are networks with a short period of life. An ad-hoc network is a wireless network to which mobile devices that can act both as client and access point are connected. The most characteristic feature of the ad-hoc network is the lack of any central control device, and also any device to supervise the operation of this information exchange system. Another important feature is the lack of fixed network infrastructure. Systems with this type of connection, therefore, are characterized by high variability and irregularity, which implies the problems absent, or present to a lesser extent in the standard fixed infrastructure networks, both wired, and wireless. Mobility of devices forming such structure is the cause of irregular construction and is a reason of frequent changes in the network structure. The consequence of these characteristics is high importance of algorithms to find not only the shortest path leading from source to destination node, but also to be able to find it fast, regardless of network structure changes. Performance of the algorithm that solves this problem with a large variation of the network structure is crucial, because the algorithm will have to be used after any change in the network structure.

This paper in its content aims to present and formulate the problem (Section 2), and demonstrate the variety of its synthetic solutions (Section 3). Major emphasis has been made to describe and present the experimentation system created (Section 4), and the results of testing of certain algorithms obtained with this system (Section 5). In the final part of the paper, the matter of prospects for the future is raised, including a summary of the most crucial parts of this paper (Section 6).

2. PROBLEM STATEMENT

To fully realize the problem of pathfinding in a graph of mobile ad-hoc network, one have to imagine a sample network, like the one shown in Fig. 1. It is clear to see, that from a mathematical point of view, this problem can be reduced to find the shortest path between two vertices of a undirected graph.

Fig. 1. Sample structure of ad-hoc network.

Mathematical model symbolizing the entire analysed network is a non directed, weighted graph. Vertices in the graph represent individual devices in the network. Connections between the vertices are the physical representation of the wireless connections between devices. The weight of each of the edges in the form of a specific number, defines the quality of the connection. In order to simplify the mathematical analysis of the problem, it can be assumed that the larger the weight, the worse the connection quality. The final element which is necessary to build a full, abstract representation of the problem is to determine the conditions of existence of the connections between different vertices.

In the proposed model, the possibility to connect two vertices in the graph is defined by their range, which is an abstract representation of the range of wireless devices in real ad-hoc networks. In the mathematical model, it will also be the number given in standardized units, to determine the radius of coverage of the given vertex. Based on the radius, it can be determined which of the neighbouring vertices of a vertex

can connect to it and, therefore, can be connected with an edge, what may represent a real connection.

3. ALGORITHMS

To carry out the simulation, two proactive algorithms and two author's reactive algorithms were implemented during the research. Dijkstra's and A* algorithms' main purpose was to provide comparison to the reactive algorithm in a modified form of Ant Colony Optimization, and one author's hybrid algorithm, which is a combination of modified versions of two of the selected algorithms.

Abstract approach to the subject has allowed to obtain greater flexibility in the implementation of these algorithms.

3.1 *Dijkstra algorithm*

Dijkstra's algorithm is an algorithm that always returns the optimal or close to the optimal route, although it is computationally greedy. In this case, the algorithm has been modified in such way, that after finding the path to the destination node it finishes the pathfinding process.

Necessary condition for the algorithm is to divide the vertices of a graph into two sets e.g. Dijkstra (1959). One set contains the vertices to which paths have been already counted, and the other contains all the nodes which have not yet been processed.

Determination of the path is made iteratively. As the first vertex, the initial, start vertex of the simulation is set. In the

3.2 *A-star algorithm*

A* algorithm, like Dijkstra's algorithm, gives the optimal path between two vertices of the graph, but to calculate the path it uses heuristics e.g. Abolhasan, Wysocki, Dutkiewicz (2003).

The algorithm minimizes the function $f(x) = g(x) + h(x)$ where $g(x)$ is the distance from the start node to the vertex x and $h(x)$ is the path predicted by the heuristic from the vertex x to the destination node. Values of $f(x)$, $g(x)$ and $h(x)$ are stored in three tables e.g. Wirth (1976).

As heuristic functions, we have chosen the „Euclid" function (1), and „Manhattan" function (2).

$$h(x) = \sqrt{(x.X - end.X)^2 + (x.Y - end.Y)^2} \quad (1)$$

$$h(x) = |x.X - end.X| + |x.Y - end.Y| \quad (2)$$

Determination of the path is iterative, as in Dijkstra's algorithm e.g. Marina, Das (2001).

3.3 *ACO algorithm*

The idea of the ant colony optimization is to base the algorithm's work on the behaviour of the colony of ants, seeking a route from their nest to food source and back again e.g. Dorigo, Stützle (2004).

Ants, as they move along the edges of the graph, leave their pheromone to indicate to the other ants that the edge has already been visited e.g. Blum (2005). With time, the concentration of pheromone P_c on the edges of the graph is decreasing with concentration loss factor l, according to (3).

$$P_c = P_c \cdot l \quad (3)$$

Pheromone concentration loss process is continuous and occurs at the beginning of each run of the algorithm's iteration e.g. Dorigo (2007).

Author's modified ACO distinguishes ants into two categories: forward and backward ants. Forward ants' main purpose is to explore the graph and find the destination node. When forward ant reaches the destination, it sends back backward ant and dyes. Backward ants are much more likely to follow the pheromone, because their priority is to consolidate the route and get back to the source node quickly, from where they send forward ants again.

In a classic implementation of this algorithm, routing tables are used to locally memorize the results of the algorithm's work in the network. For the means of an abstract implementation, routing tables have been omitted, as assumed that the subject of the research was the path finding itself, rather than maintaining the route within a given instance of the problem.

Determination of path length in this algorithm is made in an iterative manner. The path which ant chooses for the next step is added to the total value for each ant. Final result is determined as the shortest path of all of the ants.

3.4 *Hybrid algorithm*

Hybrid algorithm is an author's algorithm, which was developed in response to the need to reduce the cost of finding the path, regarding the implementation of the first n steps as quickly as possible, and then, after a quick advancement in path selection in the first stage, further optimization of the path made by using one of specialized algorithms e.g. Michalewicz (1996).

To implement this algorithm, modified version of Ant Colony Optimization was implemented in conjunction with Dijkstra's algorithm e.g. Botea, Muller, Schaeffer (2004). Modification has been made to limit the amount of ants and to modify the way the ant chooses its next vertex in the graph. Algorithm obtained in this way allows for a close to random, but relatively controlled first n steps, which will be made. After completing n steps, the ACO finishes and passes its current vertex as the starting vertex for the next algorithm.

After the calculation of the initial direction, Dijkstra's algorithm is run, which is aimed to find the path to the destination node if it has not been reached yet e.g. Cormen, Leiserson, Rivest, Stein (1990).

Path length in this algorithm is made in an iterative manner, as a sum of path values given by both of the algorithms.

4. EXPERIMENTATION SYSTEM

4.1 *Implementation environment and requirements*

The Windows platform has been chosen as an implementation environment, on which an application in C# programming language has been created. To run the simulator, the workstation must be equiped with Windows 2000/XP/Vista/7 operating system and .NET Framework 3.5.

4.2 *Application features*

The simulator has an interface that allows the user to easily configure all the parameters of the application. Moreover, its construction allows to quickly and easily extend its capabilities, including possible addition of new algorithms.

4.3 *Functional features of the application*

After launching the simulator application, the application main window appears, as shown in Figure 2. The main window is divided into four clearly separated areas.

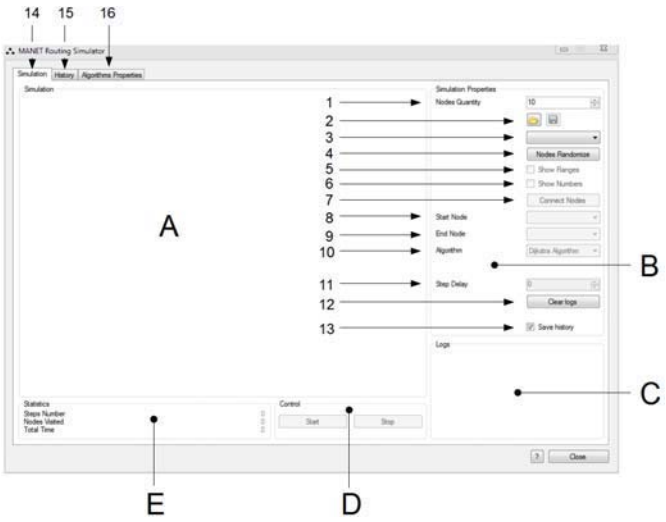


Fig. 2. Application main window.

The largest area of the application window is the area of simulation (A). In this area the graph representing the specific problem, and the effect of the algorithm will be shown. It is also possible to modify a specific instance of the problem before running the algorithm itself.

In the settings area, (B), we see the basic parameters that can be modified in the program. The first one is the parameter determining the number of vertices in the graph, which which is to be generated (1).

Next are two buttons, allowing to save the current graph to a file and load a saved graph to the program (2).

A select form (3) allows to choose a specific instance of the problem saved earlier. To add a graph to the list, save it in a subdirectory called „Graphs" in the root directory of the simulator. After adding the file and restarting the application, saved graph appears on the defined graphs selection list.

Under the selection of defined graphs, we see the graph draw button (4), allowing to generate a random graph, consisting of the number of vertices determined by the parameter (1).

Vertex positions are random according to normal distribution. If the arrangement of the vertices is not satisfying, it is possible to draw another instance by re-clicking on the „Randomize" button, or manually modifying the position of given nodes. Nodes in the simulation area can be moved using drag-and-drop method.

Below are two fields that allow to interfere in the amount of information displayed in the simulation. „Show ranges" select (5), displays the circle around each of the nodes, symbolizing node's range in relation to the other vertices. Selecting „Show numbers" parameter (6) will cause a number to appear next to each node which enables its identification.

Number (7) in the illustration has been assigned to a button that connects all vertices in the graph. Connections are made on the basis of nodes range. The connection between the two vertices a and b may occur if, and only if, the range r of the vertex with less value is less than or equal to the distance d_{ab} between the vertices (4).

$$C(a,b) = \begin{cases} 1 : \max(r_a, r_b) \geq d_{ab} \\ 0 : if\ else \end{cases} \quad (4)$$

The next two fields, (8) and (9), allow the selection of the source and destination node in the graph. Algorithms will find the shortest path between the initial and final vertex, using only the available connections. There is a possibility that it will be impossible to find any path between two selected vertices.

After selecting the initial and final vertex, an algorithm that will look for the shortest path between them can be chosen. Selection of the algorithm takes place by selecting from the drop-down list (10).

If the algorithm supports additional parameters for its operation, before the start of the simulation it is possible to configure the parameters in „Algorithm Properties" tab (16).

The last parameter that can be set is the „Step delay" (11). Here the number of milliseconds that the simulator will wait after each step of the algorithm can be specified. Note that due to the large variety of algorithms, this parameter is purely indicative.

Additional button „Clear logs" (12), is used to delete the exported results of the algorithm run.

The last option available in the main settings area is a field which allows to enable or disable algorithm run history (13).

When this option is enabled, step-by-step algorithm history analyse is possible in the „History" tab (15).

Algorithm results field (C) is located under the main settings area. Basic results of algorithm run are shown in this field.

Below the simulation area two buttons marked „Start" and „Stop" are located (D). These buttons allow to start and stop the simulation.

Current algorithm run information are shown in the live statistics area (E). These statistics are updated with every step of the algorithm, so if the delay of the algorithm iteration was set, it will be possible to analyse statistics during the run of the algorithm.

4.4 *Realization of the research*

Implementation environment allows for testing of the algorithms in several aspects. The key parameter, which is the subjected of the tests, is the overall quality of the path d_i, which is obtained as a result of the algorithm run and it is the target function (5) that has to be minimized by each of the algorithms.

$$F_c = \sum_i d_i \tag{5}$$

At the same time, the algorithm must visit the least amount of vertices possible, and take the smallest amount of time for its action. Number of vertices visited by the algorithm and the time of his realization are associated with its actual demand for resources and traffic generated by the algorithms in the network, therefore the quality of these parameters is not left without a meaning to the estimation of the quality of functioning of the algorithms.

Remaining at the level of abstract simulation of the behaviour of algorithms for searching paths in the graph, the quality of paths and quantity of visited vertices is taken into account and in this respect, the algorithms are compared.

5. INVESTIGATIONS

5.1 *Research thesis*

It is estimated that Dijkstra's algorithm provides an optimal, or very close to the optimal solution, but obtains it at great expense of calculation, which should result in relatively long run time. In the real network environment, the additional disadvantage of this algorithm is the need to process the entire graph each time a request to find the appropriate path is sent.

A* algorithm, based on the heuristic methodology, as a result of its action finds the optimal solution to the problem, using relatively large amount of resources to obtain it, so it predictably is to visit a large number of nodes in the graph.

Another approach to the problem is presented by the Ant Colony Optimization which in contrast to the other algorithms can run in the network for a long period of time, gradually improving the result and adapting to various network structure changes. In its abstract implementation, this algorithm should not show up in finding the optimal path, since the run time has been limited. Noteworthy, in the real implementation of the algorithm it exhibits a high degree of flexibility to adapt to rapidly changing network topology.

Experimental implementation of the hybrid algorithm is an interesting subject of research. It is difficult to accurately predict the algorithm behaviour and possible results, but according to the assumptions, the algorithm is to provide relatively satisfactory outcome in the short period of time, while showing a small number of visited vertices.

5.2 *Experiment design*

Each algorithm will be tested for five different numbers of vertices in the graph. Instances of graphs with 20, 30, 50, 70 and 100 vertices were selected, and saved in order to provide the same test environment for each of the algorithms.

For each of the numbers of vertices in the graph and the values of parameters of each algorithm, ten measurements will be made, which will allow to objectively asset the quality of the results, thus calculating the average results for each of the algorithms.

Summary of planned research is presented in Table 2. All experiments will be made in a research environment described earlier.

Table 2. Experiment Design.

Algorithm	Parameter	Vertices quantity				
Dijkstra	-	20	30	50	70	100
A*	Euclid	20	30	50	70	100
A*	Manhattan	20	30	50	70	100
ACO	P_c = 0,0004	20	30	50	70	100
ACO	P_c = 0,0016	20	30	50	70	100
ACO	P_c = 0,0064	20	30	50	70	100
ACO	P_c = 0,0128	20	30	50	70	100
Hybrid	n = 5	20	30	50	70	100
Hybrid	n = 10	20	30	50	70	100
Hybrid	n = 20	20	30	50	70	100

5.3 *Results and discussion*

To address the results obtained during the experiments to raised earlier research thesis in the best possible way, especially relevant results have been chosen to confirm the thesis.

In the first place, the first of the thesis, concerning the efficiency of Dijkstra's algorithm, was put under question. Performed simulations of the algorithm run time for 100 vertices, shown in Fig. 3, confirm the assumption that the

algorithm is characterized by a relatively low efficiency, needing a lot of time to process all the data.

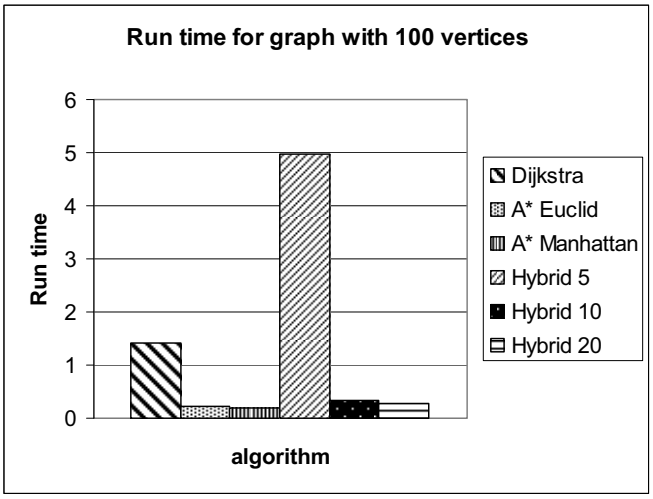


Fig. 3. Algorithms run time for 100 vertices.

It is worth noting that a high processing time has also been obtained for the hybrid algorithm, which greater part for the graph of 100 vertices is Dijkstra's algorithm, which further confirms the truth of stated thesis.

A* search algorithm, due to the complex structure of the implementation using the heuristic methods, has proved to visit the largest number of vertices, which confirms the related thesis. Example of the number of visited nodes for the graph of 30 vertices, shown in Fig. 4, classifies it right after the Ant Colony Optimization, which in the actual implementation is intended to work without the time limitation.

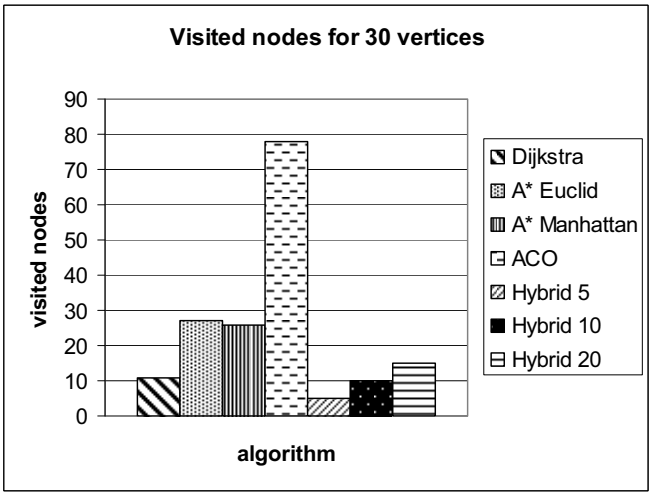


Fig. 4. Visited nodes for 30 vertices.

A noteworthy fact is that irrespective of the type of used heuristic function, A* algorithm, according to the thesis, is characterized by a large number of visited vertices, and so, in fact, a large number of generated connections, but generating the optimal solution of the pathfinding problem.

According to the thesis set for the Ant Colony Optimization, it did not provide optimal results, however, it is able to adapt to the network structure. Figure 5 shows how the path quality obtained by the ACO differs from the quality of paths developed by other algorithms in adequate run time. Clearly, author's ACO algorithm is able to find very good quality path and is further characterized by very high flexibility of action.

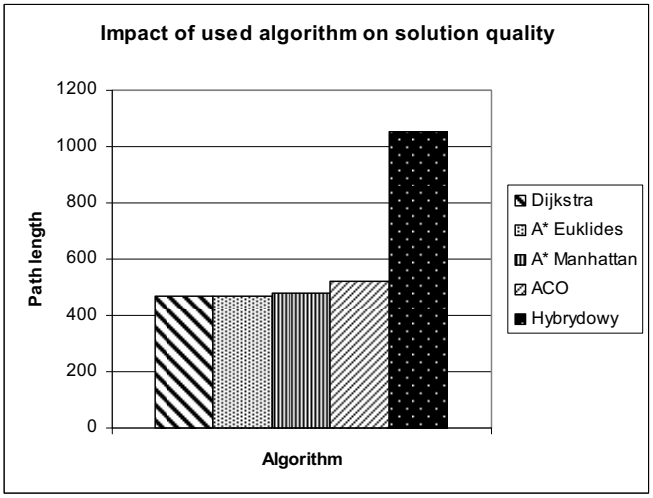


Fig. 5. Path length for 30 vertices.

It is worth to note, that the quality of path obtained by the ACO changes with the pheromone concentration loss factor. Fig. 6 shows, that properly chosen pheromone loss factor can help to make the algorithm even more effective.

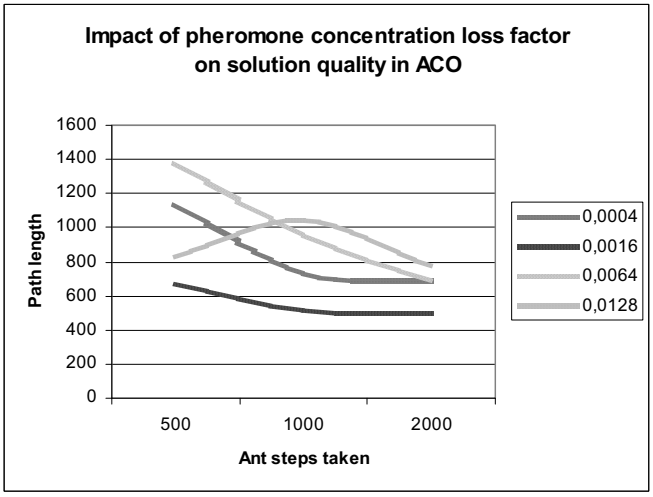


Fig. 6. Pheromone loss impact on solution quality in ACO.

The results of an experimental hybrid algorithm proved to be a confirmation of assumptions of its possible behaviour. With the increase in the contribution of modified Ant Colony Optimization, which means increasing the importance of the pseudo-random part of the algorithm, hybrid algorithm significantly increased the speed of its operation.

As shown in Fig. 7, the implementation of the first 10 steps using the modified ACO resulted in a drastic reduction of the

algorithm run time, at the cost of decreasing the quality of the solution.

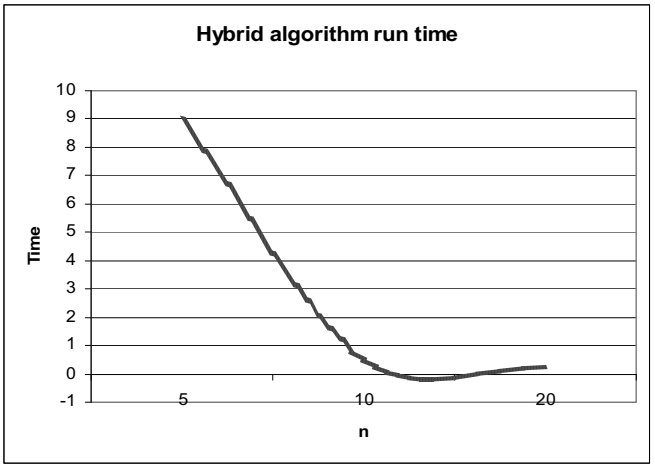


Fig. 7. Hybrid algorithm run time.

With the increase of the *n* parameter, the number of steps taken by the algorithm has significantly decreased. The dependence is shown on Figure 8. Number of visited vertices remained more or less stable, which further emphasizes the importance of pseudo-random part of the algorithm to reduce the amount of the calculation.

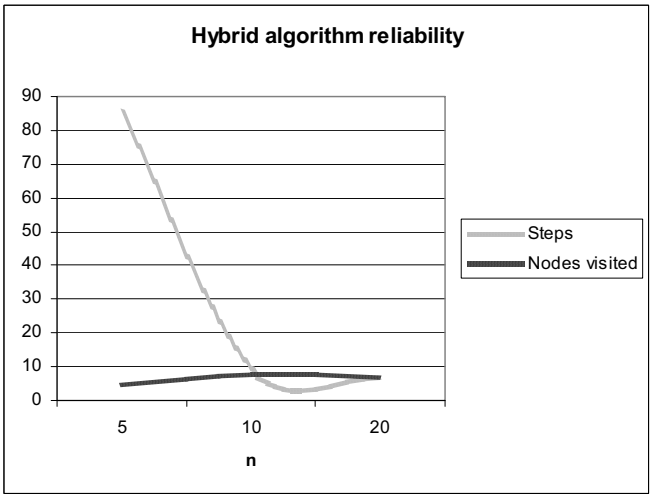


Fig. 8. Hybrid algorithm reliability.

Close to random nature of the hybrid algorithm is stressed by the fact that for the *n* parameter value equal to 20, the number of performed steps has slightly increased, which is caused by too much involvement of the random part of the algorithm. An appropriately balanced algorithm parameters can improve the overall quality of obtained results and the algorithm itself provides promissing results and a solid basis for further research and development.

6. CONCLUSIONS

Research carried out under the project allowed to draw far-reaching proposals for the design of systems based on the idea of finding a path in wireless ad-hoc networks.

Diversity of the algorithms realizing the routing in wireless ad-hoc networks available to implement requires to clarify and clearly specify the system requirements. When it is known that the system must be resistant to changes in network and rapid adaptations to new conditions, it is advised to use algorithms that provide the desired flexibility, for example, Ant Colony Optimization algorithm. If the key is to obtain a satisfactory solution to the problem in the shortest time possible and subject minimize the consumption of resources, a good solution could be a hybrid algorithm, similar to the algorithm implemented for this project, which can combine the best features from selected algorithms while maintaining an appropriate balance between their drawbacks.

In the future implementation of similar project, the right direction would be to develop the idea to closer simulate the reality, gradually moving away from abstract approaches. This would enable more specific implementation of the algorithms for selected problems and to conduct more in-depth research. Nodes could use the parameters of the actual nodes of ad-hoc network, which combined with assigning more details to the connection between two nodes would increase the level of realism, which would help to carry out further tests, developing more accurate reflection of reality.

Program providing the role of the simulation environment was designed with a possibility to expand it with additional modules. Increasing the functionality and reducing the level of abstraction can provide a solid basis for future research in this topic.

REFERENCES

Abolhasan, M., Wysocki, T., and Dutkiewicz, E. (2003). *A review of routing protocols for mobile ad hoc networks*, University of Wollongong.

Marina, M.K. and Das, S.R. (2001). *On-Demand Multipath Distance Vector Routing in Ad Hoc Networks*, University of Cincinnati.

Botea, A., Muller, M. and Schaeffer, J. (2004). Near Optimal Hierarchical Path-Finding, *Journal of Game Development*.

Dijkstra, E.W. (1959). A Note on Two Problems in Connexion with Graphs, *Numerische Mathematik*.

Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C. (1990). Dijkstra's algorithm. *Introduction to Algorithms*, Section 24.3, MIT Press.

Dorigo, M and Stützle, T. (2004). *Ant Colony Optimization,* MIT Press.

Blum, C. (2005). *Ant colony optimization: Introduction and recent trends*, Physics of Life Reviews.

Dorigo, M. (2007). *Ant Colony Optimization.* Scholarpedia.

Wirth, N. (1976). *Algorithms + Data Structures = Programs,* Prentice Hall.

Michalewicz, Z. (1996). Genetic Algorithms + Data Structures = Evolution Programs, Springer.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Testing SQL queries: an experimentation system and efficiency evaluation

Michał Hans*, Paweł Kmiecik*, Iwona Pozniak-Koszalka*, Andrzej Kasprzak

**Dept. of Systems and Computer Networks, roclaw niversity of Technology, 5 -3 roclaw, Poland e-mail: les ek.kos alka pwr.wroc.pl .*

Abstract: This paper's main goal is to discuss useful optimizing methods of database queries based on PHP, MyS L and PostgreS L examples. The research was done on the specially prepared environment: computer workstation with Apache, PHP, MyS L and PostgreS L installed on it. The databases storing different amount of data were prepared. Several aspects of optimization were researched, including: Influence of using cache on processing time while querying on the example of DATA field; Researching queries that use SELECT * structure; Influence of adding LIMIT 1 condition on processing time when searching for unique line; Influence of field indexing on processing time; Comparing ENUM and VARCHAR fields; Researching different methods of querying for a random line.

Keywords: Database, sql query, optimization methods, php, mysql, postgresql.

1. INTRODUCTION

Imagine the situation when one stands in front of the exclusive buffet with countless amount of delicious courses. The task is to try them all, but first you have to think through: in which order. Which flavors in combination with others flavors would give the maximum of the pleasure

uite similar, but less enjoyable and subjective, are the problems that databases programmers have to face. While designing a database queries one has to remember that there are many different ways in which the DBMS can execute tasks and acquire the answers. Of course, all the methods give you the same results, but the processing times will differ. The objective of this paper is to describe and to present the results of the tests made using the query optimization methods. All the tests were made on Apache Server and two databases: PostgreS L and MyS L. All the scripts that contain queries were written in PHP.

The rest of the paper is organized as follows. Six aspects of query optimization is presented in Section 2, and in Section 3 the used approaches are described. The designed and implemented experimentation system is presented in Section 4. Sections 5 contains results of investigations. Final remarks appear in Section 6.

2. ADDRESSING THE PROBLEM

This paper will closely focus on 6 aspects of query optimization:

1. The influence of using cache while querying on the example of DATA field: The most of the database servers have a built-in and turned-on cache option (Hernandez, 2003). It is considered as the most effective method of reducing the time needed for executing the query. When the same query is to be processed many times, the result is kept in the cache memory which is the fastest memory available. The problem is that there are many conditions under which the cache memory is blocked and not used, as in example: `$r = mysql_query("SELECT Name FROM Workers WHERE HireDate >= CURDATE()");` For this query cache memory will not be used, because the result of `CURDATE()` function is not known so, the query cannot be compared with previous ones.
2. Researching queries that use SELECT * structure.
3. Looking at queries that check if the database contains at least one line that fulfill given conditions. In that case, it happens that programmers do not add LIMIT 1 condition, for example:
   `SELECT SQL_NO_CACHE * FROM 'Workers_100k' WHERE HireDate <= '2010-01-24';`
4. The impact of field indexing on processing time: How can we improve the query performance using indexing (O'Neil, 1997)
   `SELECT SQL_NO_CACHE Surname FROM 'Workers_100k' WHERE Surname LIKE 'a%'`
5. Comparison between *ENUM* and *VARCHAR* fields: How to improve the processing time with queries that ask for lines containing *VARCHAR* field
6. Looking at different methods of querying for a situation contained queries for random rows. uery to optimize:
   `SELECT Name FROM Workers WHERE HireDate <= '2010-01-24'`

3. DESCRIPTION OF APPROACHES

For the problems presented in Section 2, the following approaches were considered:

Ad.1. The reason why cache was disabled is using the function `CURDATE()`. This situation takes place every time for non-deterministic function like `NOW()` or `RAND()`. This kind of functions can return unique results every time. To solve the speed problem with the current date function we should try to get current date in PHP and build a query which contains that date. In this way we can cache our query for whole day until midnight, as in example:

```
SELECT Name FROM Workers WHERE HireDate <= '2010-01-24';
```

Ad.2. More data to search always mean more time is required to get the results. Looking at the most of web applications scripts which use databases connections we can find many examples when programmers get all of the fields from the row (by using select *). When we launch this kind of queries we can access the every field by filtering the array in PHP. This means that if we want to read only one field from whole row, the rest fields are just wasting server memory and they are not used. This situation is very important when the database server location is far away from PHP server. To solve that problem we should always point to the database which fields are interesting for us. For example, when we want to read only the name of the worker that has *id* no we should use:

```
SELECT Surname FROM Workers_100k Where id = 666;
```

Ad.3. Solution for this kind of problem is to add condition `LIMIT 1` at the end of the query. In this way the database engine stops after first fitted row instead of looking for the other that matches in the database.

```
SELECT Surname FROM Workers_100k Where id = 666 LIMIT 1;
```

Ad.4. To improve the speed of getting the results, we should create a columns index for the things that we are searching in. This solution is a very good method when we are looking for a common searching field. For example when we are looking for workers surnames:

```
CREATE INDEX IndexName ON 'users' (last_name);
SELECT SURNAME FROM 'users' WHERE last_name LIKE 'a%';
```

Ad.5. `ENUM` columns are very fast to search. Even though they are store at database as `TINYINT` they can contain a string. When we have a `VARCHAR` field that contains many similar values (for example status as active, suspend and locked) we should use `ENUM` to save a lot of memory.

Ad.6. To find the better way to get random row from the table in a database we compare two of them together. The first one is to count all of the rows in a table, and then to generate random number in PHP which will be the number of our random row. The second way of approaching this problem is to use a `RAND()` function in S L query.

```
$polecenie1 = "SELECT count(*) FROM workers";
$w1 = mysql_query($polecenie1);
$d = mysql_fetch_row($w1);
$rand = rand(0,$d[0] - 1);
$polecenie2 = "SELECT $pole FROM workers LIMIT $rand, 1";
$w2 = mysql_query($polecenie2);
```

4. E PERIMENTATION SYSTEM

We were testing the queries before and after modification in our web application. Application was based on PHP and could run with two S L databases (MyS L and PostgreS L). We decided to work on these servers because there are open sourced and there are free of charge for everyone. The testing environment involved an ASUS M50VN with dual core CPU (2,26 GHz) and 3GB of RAM. We were working using Windows Vista Home Premium with Service Pack 1 32bit system with WAMP server which contains Apache Server, MyS L and PHP. Additionally, we also integrated a PostgreS L with it. In both databases (Table 1 and Table 2) we created the same tables. Remark: In Table 2 the xxx is a number of rows.

Table 1. Structure of data pwr database.

Table name	rows	size
Departments	10	2,2 KB
Workers_1k	1 000	93,5 KB
Workers_10k	10 000	918,1 KB
Workers_100k	100 000	8,9 MB

Table 2. Structure of Workers xxx table.

Field	Type	Info
Id	Int (11)	Auto icrement
Name	Varchar (255)	
Surname	Varchar (255)	
City	Varchar (255)	
Street	Varchar (255)	
HireDate	Date	
DepartmentID	Int (11)	
Status	Varchar (20)	
Status2	Enum (active', locked', suspend')	

The application runs using AJA to send queries and to show the results. Using the application it is possible to identified the specific problems, including the following:

TEST1 – Influence of using cache while querying on the example of `DATA` field
TEST2 – Researching queries that use `SELECT *` structure
TEST3 – Researching queries that check if the database is containing at least one line that fulfill given conditions.
TEST4 – Influence of the field indexing on the processing time.

TEST5 – Comparison between fields for *EN M* and *VARCHAR*.

TEST6 – Looking at different methods of querying for a random row.

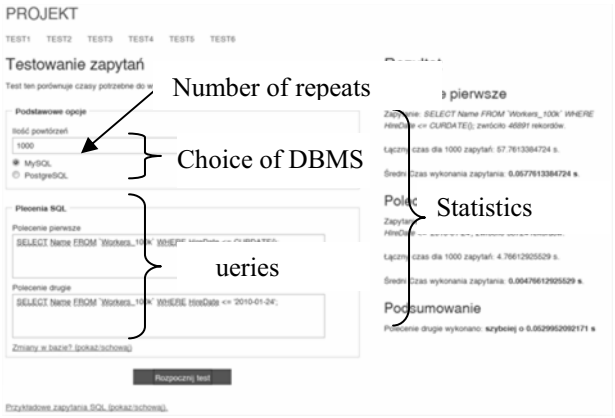


Fig. 1. Application interface.

Our application is able to test any two queries typed in by the user. The user also is able to choose the database engine and the number of repeats. On the right side of screen after running the test we are able to see the statistic for our test.

5. RESULTS OF INVESTIGATIONS

In Tables 1 – 6, we present the results of our research obtained during six tests corresponded to the approaches discussed in the previous sections. The results are also shown in convenient way on Figures 2 – 7. In figures the vertical axis represents productivity growth in percent (for example, 200 means two times faster run). The horizontal axis represents the number of repeats.

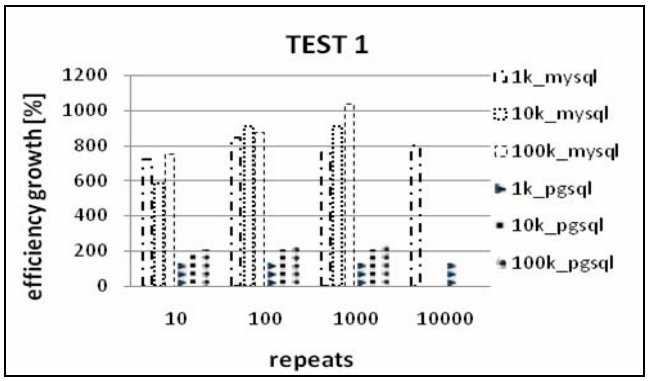


Fig. 2. Test 1 - comparison.

Table 3. Results of Test 1.

Repeats	1k_mysql	10k_mysql	100k_mysql	1k_pgsql	10k_pgsql	100k_pgsql
10	721	584	752	125	189	201
100	842	911	878	129	205	217
1000	762	906	1035	123	205	232
10000	799	-	-	124	-	-

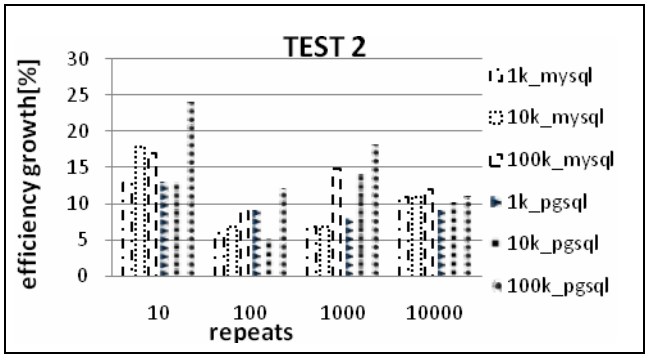


Fig. 3. Test 2 - comparison.

Table 4. Results of Test 2.

Repeats	1k_mysql	10k_mysql	100k_mysql	1k_pgsql	10k_pgsql	100k_pgsql
10	13	18	17	13	13	24
100	6	7	9	9	5	12
1000	7	7	15	8	14	18
10000	11	11	12	9	10	11

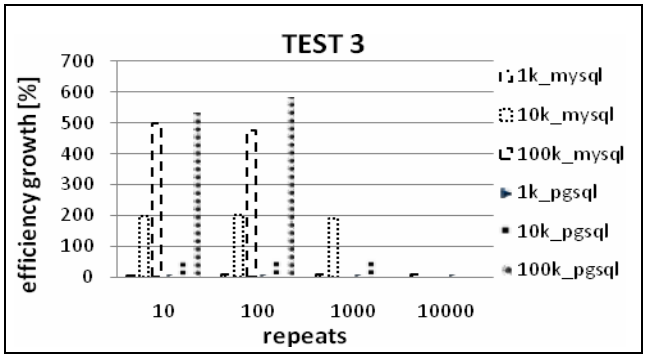


Fig. 4. Test 3 - comparison.

Table 5. Results of Test 3.

Repeats	1k_mysql	10k_mysql	100k_mysql	1k_pgsql	10k_pgsql	100k_pgsql
10	5	199	500	5	53	530
100	6	202	475	5	54	578
1000	6	189	-	5	54	-
10000	6	-	-	5	-	-

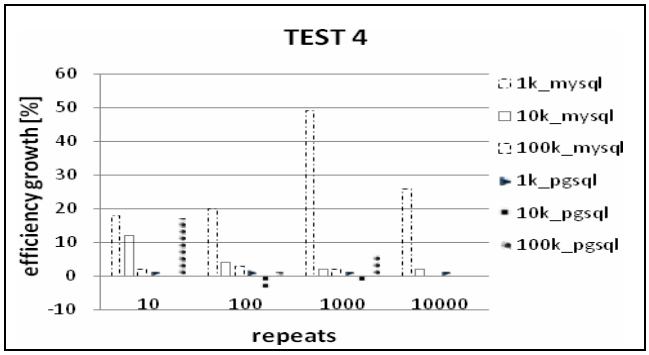


Fig. 5. Test 4 - comparison.

Table 6. Results of Test 4.

Repeats	1k_mysql	10k_mysql	100k_mysql	1k_pgsql	10k_pgsql	100k_pgsql
10	18	12	2	1	0	17
100	20	4	3	2	-4	1
1000	49	2	2	1	-2	6
10000	26	2	-	1	0	-

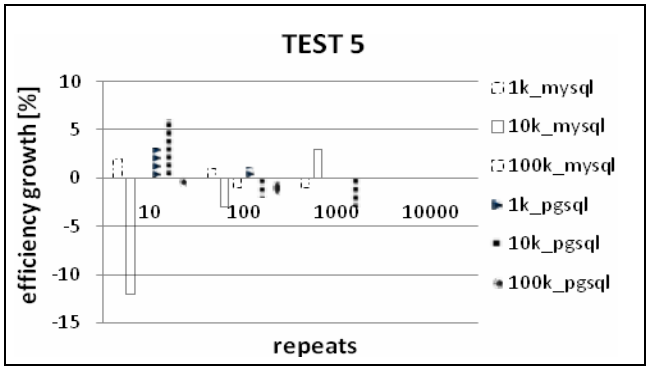


Fig. 6. Test 5 - comparison.

Table 7. Results of Test 5.

Repeats	1k_mysql	10k_mysql	100k_mysql	1k_pgsql	10k_pgsql	100k_pgsql
10	2	-12	0	3	6	-1
100	1	-3	-1	1	-2	-2
1000	1	3	-	0	-3	-
10000	0	-	-	0	-	-

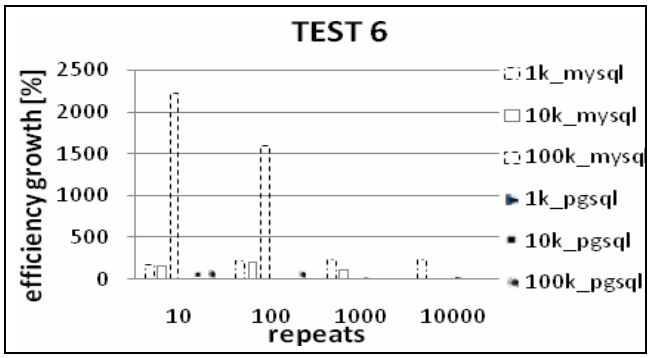


Fig. 7. Test 6 - comparison.

Table 8. Results of Test 6.

Repeats	1k_mysql	10k_mysql	100k_mysql	1k_pgsql	10k_pgsql	100k_pgsql
10	166	159	2224	3	82	116
100	216	201	1602	4	10	109
1000	227	112	-	6	1	-
10000	223	-	-	17	-	-

6. SUMMARY

Looking at the obtained results (shown in Section 5) we can observe that most of our modifications increase the productivity.

The designed and implemented application, considered in this paper, can be easily modified in the future, in order to work with the other databases and more potential test scenarios. The application has an open structure, so we can use it to compare times of run for any two queries. We can run the first query, and after that modify table and automatically run the second query. In this manner we can compare a lot of different scenarios.

Unfortunately, our workstation was not capable enough to store the entire test in some cases e.g. when the number of rows was close to 100 000. Because of that limit, we are thinking about moving the application to other – more efficient server.

REFFERENCES

Fritchey, G. and Dam, S. (2009). *SQ Server 2 Query Performance Tuning Distilled,* USA

Gulutzan, P. and Pelzer, T. (2002). *SQ Performance Tuning*, USA

Hernandez, M. J. (2003). *Database Design for Mere Mortals : A ands-On uide to Relational Database Design, Second Edition*, USA

Ioannidis, Y. (1997). *Query Optimi ation*, Computer Sciences Department University of Wisconsin Madison, WI53706

Ioannidis, Y. and Kang, Y. (1990). *Randomi ed algorithms for optimi ing large oin ueries,* USA

O'Neil, P. and uass, D. (1997). *Improved uery performance with variant indexes*, USA

Schwartz, B. Zaitsev, P. Tkachenko, V. Zawodny, J. Lentz A., and Balling D. (2008) *igh performance MySQ* , USA

Widenius, M. Axmark, D. (2002). *Mys l Reference Manual,* 1st edition, USA

Worsley, J. C. Drake, J. D. (2002). *Practical PostgreSQ* , USA.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Properties of NCGPC applied to nonlinear SISO systems with a relative degree one or two

M. Dabo. N. Langlois. H. Chafouk

IRSEEM, Technopole du Madrillet, 76810 St Etienne du Rouvray, France
(Tel: +33 32 91 58 58 ; e-mail: nicolas.langlois@esigelec.fr).

Abstract: In this paper, we present some properties of the nonlinear continuous-time generalized predictive control (NCGPC) when this latter is applied to nonlinear single-input single-output (SISO) systems with a relative degree equal to one or two. From a simple change of coordinates, the resulting closed-loop linear system has, among others, the following properties: for a relative degree one, its equivalent time response is double the prediction horizon time; for a relative degree two, its overshoot is constant and equal to 0.685 with a natural frequency of 1.83 times the inverse of the prediction horizon time. The control law is applied to two academic examples. Some simulation results are shown to highlight these properties.

Keywords: Nonlinear predictive control, continuous time systems, dynamic properties, linear analysis.

1. INTRODUCTION

Predictive control was introduced by Richalet et al. (1978) as a heuristic predictive model for the control of industrial processes. Long-range predictive controllers based on predictive strategy as PCA (predictive control algorithm) and DMC (dynamic matrix control) have also been used in Bruijn et al. (1980) and Cutler et al. (1980) respectively. Clarke et al. (1987a, b) give a more general approach of this method known as generalized predictive control (GPC) for discrete time systems. Demircioglu et al. (1991, 1992) introduced, respectively, continuous time generalized predictive control (CGPC) and multivariable CGPC, namely MCGPC, finding that it is more natural to solve problems of control in the continuous time domain. Chen (2001) and Chen et al. (2003) propose a control design method based on predictive control for nonlinear systems, for which prediction is based on expansion in Taylor series. A major result of this control design method is that closed-loop stability is guaranteed when the relative degree of the considered nonlinear system is less than or equal to four. In this paper, we focus our study on the properties of the closed-loop linear systems resulting from NCGPC control law when this latter is applied to nonlinear systems with relative degree one or two. Interesting properties rise from this study: 1) a SISO nonlinear system of dimension one equal to its relative degree, has a time constant and a time response, respectively equal to 1.5 times and 2 times the prediction horizon time, via NCGPC control law; 2) a SISO nonlinear system of dimension two equal to its relative degree, has a constant damping ratio equal to $\xi = 0.685$ for any given value of the prediction horizon time T and an undamped natural frequency equal to $\omega_n = 1.83 / T$.

The paper is outlined as follows: section II presents unconstrained NCGPC while section III highlights the properties of the considered nonlinear system of dimension one or two. In section IV, applications are presented through two academic examples.

2. UNCONSTRAINED NCGPC

2.1 System considered

Consider a nonlinear SISO system of the form

$$\begin{cases} \dot{x}(t) = f(x(t)) + g(x(t))u(t) \\ y(t) = h(x(t)) \end{cases} \quad (1)$$

where $x \in X \subset \Re^n$, $y \in Y \subset \Re$ and $u \in U \subset \Re$. The goal is to find a control law so that the output $y(t)$ of (1) tracks asymptotically a given reference signal $\omega(t)$. Unconstrained predictive control consists in deriving a control law by minimizing a receding horizon performance index (or criterion), in a finite prediction horizon time without taking into account constraints on the vector state, the input and on the output.

2.2 Relative Degree

To simplify the exposition, the standard geometric notation for Lie derivatives is used in this paper. For a real-valued function h on $\Re^n$ and a vector field f on $\Re^n$, the Lie derivative of h along f at $x \in \Re^n$ is given by:

$$L_f h(x) = \sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(x) f_i(x) \tag{2}$$

The nonlinear SISO system (1) is said to have a relative degree ρ around x^0 if

(i) $L_g L_f^k h(x) = 0$ for all x in a neighbourhood of x^0 and all $k < \rho\text{-}1$,

(ii) $L_g L_f^{\rho-1} h(x^0) \neq 0$, regarding Isidori (1995).

Regarding Chen (2001), the relative degree ρ of (1) is said to be well-defined if (1) has the relative degree ρ at all points in an operating set.

2.3 Zero Dynamics

Zero dynamics corresponds to the dynamics describing the internal behaviour of the system when input and initial conditions have been chosen in such a way that the output remains identically zero, Isidori (1995).

2.3 Error Prediction

The way to predict the output is based on the expansion in Taylor series. An approximation of the reference signal is done in the same way. The expansion in Taylor series of output y up to an order equal to relative degree ρ is

$$\hat{y}(t+\tau) = \sum_{k=0}^{\rho} y^{(k)}(t) \frac{\tau^k}{k!} + R(\tau^\rho) \tag{3}$$

where t is the present instant, $t + \tau$ the moment for which the prediction is made. $R(\tau^\rho)$ which represents high order terms of the Taylor series expansion of the output is neglected in the following. From this,

$$\hat{y}(t+\tau) \approx \begin{bmatrix} 1 & \tau & \cdots & \dfrac{\tau^\rho}{\rho!} \end{bmatrix} \begin{bmatrix} y(t) \\ \dot{y}(t) \\ \vdots \\ y^{(\rho)}(t) \end{bmatrix} \tag{4}$$

with

$$\begin{cases} y(t) = h(x(t)) \\ \dot{y}(t) = L_f h(x(t)) \\ \vdots \\ y^{(\rho)}(t) = L_f^\rho h(x(t)) + L_g L_f^{\rho-1} h(x(t)) u(x(t)) \end{cases} \tag{5}$$

An expression of the reference signal ω can be obtained in the same way. As our goal is to find a control law so that the output y asymptotically tracks the reference signal ω, let us define the error $e(t) = y(t) - \omega(t)$. Therefore, with an appropriate control law, the error is equal to zero in a finite time if and only if the output is equal to the reference signal. The error prediction can then be defined as

$$\hat{e}(t+\tau) = \hat{y}(t+\tau) - \hat{\omega}(t+\tau) \tag{6}$$

where $\hat{y}(t+\tau)$ and $\hat{w}(t+\tau)$ denote, respectively, the approximated predicted output (4) and the approximation by Taylor series of the reference signal for any given instant τ. Let

$$\Lambda(\tau) = \begin{bmatrix} 1 & \tau & \cdots & \dfrac{\tau^\rho}{\rho!} \end{bmatrix} \tag{7}$$

and

$$Y(t) = \begin{bmatrix} y(t) \\ \dot{y}(t) \\ \vdots \\ y^{(\rho)}(t) \end{bmatrix} \text{ and } \Omega(t) = \begin{bmatrix} \omega(t) \\ \dot{\omega}(t) \\ \vdots \\ \omega^{(\rho)}(t) \end{bmatrix} \tag{8}$$

Rewriting (6) in the matrix form yields

$$\hat{e}(t+\tau) = \Lambda(\tau) E(t) \tag{9}$$

where $E(t) = Y(t) - \Omega(t)$.

2.4 Control Law

The control law will be derived under the assumptions in Chen (2001) and Chen et al. (2003). Consider the receding horizon performance index

$$J = \frac{1}{2} \int_0^T [\hat{e}(t+\tau)]^2 d\tau \tag{10}$$

where $T \in \Re_*^+$ is the prediction horizon time and τ a given instant belonging to interval [t, t + T]. Plugging equation (9) into (10) yields:

$$J = \frac{1}{2} E^t(t) \left[\int_0^T \Lambda^t(\tau) \Lambda(\tau) d\tau \right] E(t) \tag{11}$$

For practical reasons, let us define the prediction matrix as

$$\Pi(T, \rho) = \int_0^T \Lambda^t(\tau) \Lambda(\tau) d\tau \tag{12}$$

where Π(T, ρ) is of dimensions (ρ + 1)×(ρ + 1). Thus to derive the control law, we need to minimize the criterion with respect to control u. This yields:

$$\left(\frac{\partial E(t)}{\partial u(t)} \right)^t \Pi(T, \rho) E(t) = 0 \tag{13}$$

Vector E can be written as follows:

$$E = \begin{bmatrix} y(t) - \omega(t) \\ \dot{y}(t) - \dot{\omega}(t) \\ \vdots \\ y^{(\rho)}(t) - \omega^{(\rho)}(t) \end{bmatrix} \tag{14}$$

Separating terms which contain u from those that do not, yields:

$$E = \begin{bmatrix} h-\omega \\ L_f h - \dot{\omega} \\ \vdots \\ L_f^{\rho} h - \omega^{(\rho)} \end{bmatrix} + \begin{bmatrix} 0_{\rho\times 1} \\ uL_g L_f^{\rho-1} h \end{bmatrix} \quad (15)$$

Therefore

$$\left(\frac{\partial E(t)}{\partial u(t)}\right)^t = \begin{bmatrix} 0_{1\times\rho} & L_g L_f^{\rho-1} h(x(t)) \end{bmatrix} \quad (16)$$

For the sake of simplicity let us define

$$D(x(t)) = L_g L_f^{\rho-1} h(x(t)) \quad (17)$$

Thus substituting expression $L_g L_f^{\rho-1} h(x(t))$ by $D(t)$ in equation (16) and the resulting equation in (13) yields:

$$\begin{bmatrix} 0_{1\times\rho} & D \end{bmatrix} \Pi \begin{bmatrix} h-\omega \\ \vdots \\ L_f^{\rho} h - \omega^{(\rho)} + Du \end{bmatrix} = 0 \quad (18)$$

After simplifications, we obtain:

$$D\Pi_s \begin{bmatrix} h-\omega \\ \vdots \\ L_f^{\rho} h - \omega^{(\rho)} + Du \end{bmatrix} = 0 \quad (19)$$

where Π_s, of dimensions $1 \times (\rho + 1)$, corresponds to the last row of the prediction matrix. As the relative degree is supposed well-defined, D cannot vanish for all $x \in X$. Therefore, separating (19) into two parts, one with the control law u on the left-hand side and other without it on the right-hand side, yields:

$$\Pi_s \begin{bmatrix} 0_{\rho\times 1} \\ Du \end{bmatrix} = \Pi_s \begin{bmatrix} \omega - h \\ \vdots \\ \omega^{(\rho)} - L_f^{\rho} h \end{bmatrix} \quad (20)$$

By simplifying the left-hand side of (20), we have the following equation:

$$\Pi_{ss} Du = \Pi_s \begin{bmatrix} \omega - h \\ \vdots \\ \omega^{(\rho)} - L_f^{\rho} h \end{bmatrix} \quad (21)$$

where Π_{ss}, of dimensions 1×1, corresponds to the last element of vector Π_s. The control law is then given by

$$u = D^{-1}\Pi_{ss}^{-1}\Pi_s \begin{bmatrix} \omega - h \\ \vdots \\ \omega^{(\rho)} - L_f^{\rho} h \end{bmatrix} \quad (22)$$

Let $K = \Pi_{ss}^{-1}\Pi_s$. The computation of K from (12) yields:

$$K(T,\rho) = \begin{bmatrix} \frac{\rho!}{T^{\rho}} \frac{2\rho+1}{\rho+1} & \cdots & \frac{\rho!}{T^{\rho-1}} \frac{2\rho+1}{(l)!(\rho+l+1)} & \cdots & 1 \end{bmatrix} \quad (23)$$

Define $K_{\rho l}$ as the component corresponding to the $(l+1)^{th}$ column of matrix $K(T, \rho)$. It is equal to

$$K_{\rho l} = \frac{\rho!}{l!} \frac{2\rho+1}{(\rho+l+1)T^{\rho-1}} \quad (24)$$

for any integer l so that $0 \le l \le \rho$. This yields finally the following vector control law

$$u(x(t)) = \frac{-\sum_{l=0}^{\rho} K_{\rho l}(T,\rho)\left[L_f^l h(x(t)) - \omega^{(l)}(t)\right]}{L_g L_f^{\rho-1} h(x(t))} \quad (25)$$

As in Dabo et al. (2009), let consider the change of coordinates:

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_{\rho} \end{bmatrix} = \begin{bmatrix} y - \omega \\ \dot{y} - \dot{\omega} \\ \vdots \\ y^{(\rho-1)} - \omega^{(\rho-1)} \end{bmatrix} \quad (26)$$

These new coordinates yield the nonlinear system

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \vdots \\ \dot{z}_{\rho} = L_f^{\rho} h - \omega^{(\rho)} + uL_g L_f^{\rho-1} h \end{cases} \quad (27)$$

Replacing u by (25) in (27) yields the following closed-loop linear and controllable system

$$\begin{cases} \dot{Z} = AZ \\ O = CZ \end{cases} \quad (28)$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -K_{\rho 0} & -K_{\rho 1} & -K_{\rho 2} & \cdots & -K_{\rho(\rho-1)} \end{bmatrix} \quad (29)$$

The characteristic polynomial $P_{\rho}(\lambda)$ of (29) is given by

$$P_{\rho}(\lambda) = K_{\rho 0} + K_{\rho 1}\lambda + \cdots \lambda^{p} = 0 \quad (30)$$

where Kρl is given by (24).

3. SOME PROPERTIES OF NCGPC

Our goal in this study is to highlight some properties of NCGPC when this latter is applied to nonlinear systems with a relative degree equal to one or two. For the sake of simplicity, we will consider in this study that: 1) the dimension of the nonlinear system is equal to the relative degree; 2) the reference signal ω is a constant.

3.1 Case of relative degree ρ = 1

Consider a nonlinear SISO system (1) of dimension one equal to the relative degree ρ. The resulting closed-loop linear system has the characteristic polynomial

$$P_1(\lambda) = K_{10} + \lambda \tag{31}$$

This polynomial is equivalent to the denominator $D_1(p)$ of a first order transfer function $H_1(p)$ given by:

$$H_1(p) = \frac{G_1}{1+\theta p} \tag{32}$$

where G_1 is the static gain and θ the time constant. From this and by analogy, the pole of the characteristic polynomial $P_1(\lambda)$ and that of the transfer function $H_1(p)$ are equivalent. Therefore

$$K_{10} = \frac{1}{\theta} \tag{33}$$

where K_{10} is the first element of (23) and is equal to

$$K_{10} = \frac{\rho!}{T^\rho}\frac{2\rho+1}{\rho+1} \tag{34}$$

Hence, we have

$$\frac{1}{\theta} = \frac{\rho!}{T^\rho}\frac{2\rho+1}{\rho+1} \tag{35}$$

and as ρ = 1, $\theta = \frac{2T}{3}$ and $\lambda = -\omega_c = -\frac{3}{2T}$. From this, we can deduce that the time response $t_r = 3\theta = 2T$ and the cut-off frequency $\omega_c = 1/\theta = 3/2T$ are functions of the prediction horizon time.

Theorem 3.1: The application of NCGPC to SISO nonlinear system of dimension one equal to its relative degree, leads, in the right space of coordinates, to a linear 1st order system with transfer function defined by a time constant $\theta = 2T/3$ and a static gain G_1 equal to the reference signal ω_1.

3.2 Case of relative degree ρ = 2

Consider a nonlinear SISO system (1) of dimension two equal to the relative degree ρ. The resulting closed-loop linear system has the characteristic polynomial

$$P_2(\lambda) = K_{20} + K_{21}\lambda + \lambda^2 \tag{36}$$

This polynomial is equivalent to the denominator $D_2(p)$ of a linear transfer function $H_2(p)$ of order two given by:

$$H_2(p) = \frac{G_2}{p^2 + 2\xi\omega_n p + \omega_n^2} \tag{37}$$

From (37), we can deduce the damping ratio ξ and the natural pulsation ωn of $P_2(\lambda)$. Hence:

$$\begin{cases} K_{20} = \omega_n^2 \\ K_{21} = 2\xi\omega_n \end{cases} \tag{38}$$

From equation (24), we have:

$$\begin{cases} \frac{\rho!}{T^\rho}\frac{2\rho+1}{\rho+1} = \omega_n^2 \\ \frac{\rho!}{T^{\rho-1}}\frac{2\rho+1}{\rho+2} = 2\xi\omega_n \end{cases} \tag{39}$$

or, equivalently,

$$\begin{cases} \frac{\rho!}{T^\rho}\frac{2\rho+1}{\rho+1} = \omega_n^2 \\ \frac{1}{2\xi}\frac{\rho!}{T^{\rho-1}}\frac{2\rho+1}{\rho+2} = \omega_n \end{cases} \tag{40}$$

Putting the second equation of the above system to the power two yields:

$$\begin{cases} \frac{\rho!}{T^\rho}\frac{2\rho+1}{\rho+1} = \omega_n^2 \\ \left(\frac{1}{2\xi}\frac{\rho!}{T^{\rho-1}}\frac{2\rho+1}{\rho+2}\right)^2 = \omega_n^2 \end{cases} \tag{41}$$

Therefore, we have:

$$\omega_n = \sqrt{\frac{\rho!}{T^\rho}\frac{2\rho+1}{\rho+1}} \tag{42}$$

and

$$\xi(T,\rho) = \frac{1}{2}\frac{\rho!}{T^{\rho-2}}\sqrt{\frac{(\rho+1)(2\rho+1)}{(\rho+2)^2}} \tag{43}$$

Finally, for $\rho = 2$, $\omega_n \approx 1.83/T$ and $\xi \approx 0.685$. It is interesting to note that the damping ratio ξ is always lower than $\sqrt{2}/2$. From (36), we can deduce the complex conjugate poles of $P_2(\lambda)$:

$$\lambda_{1,2}(T,\rho) = -\frac{K_{21}(T,\rho)}{2} \pm \frac{j}{2}\sqrt{4K_{20}(T,\rho) - K_{21}^2(T,\rho)} \tag{44}$$

Replacing K_{20} and K_{21} by their numerical values yields:

$$\lambda_{1,2}(T) = -\frac{1}{T}(1.25 \pm 1.33j) \tag{45}$$

From (42) and (43) the dynamics properties of 2nd order systems usually considered in the time domain such as rise-time, time-to-peak and settling times can be then easily written as functions of T. In the frequency domain it is interesting to note that the percent overshoot (P.O.) and the maximum magnitude ($M_{p\omega}$) are constant since they are functions of the damping ratio only: $M_{p\omega} \approx 1$ and $P.O. \approx 5.21$.

Theorem 3.2: The application of NCGPC to SISO nonlinear system of dimension two equal to its relative degree, leads, in the right space of coordinates, to a 2nd order linear transfer function with a constant damping ratio $\xi \approx 0.685$ and a natural frequency $\omega_n \approx 1.83/T$.

4. APPLICATIONS

In this section we will consider two academic applications. All simulations are derived via Matlab Simulink version 7.0.1. Zooms of some figures are given to show important details such as time response (relative degree one) or overshoot (relative degree two). For both academic applications of dimensions one and two, all step sizes (Max, Min and Initial) and the "Absolute tolerance" are on "auto". Only the "Relative tolerance" is kept equal to 10^{-3} for the first system and on "auto" for the second one. The solvers used are Ode45 (Dormand-Prince) and Ode23 (Bogacki-Shampine), respectively, for the first and second academic applications.

1) Nonlinear system with relative degree one. Consider the following nonlinear SISO system of dimension one:

$$\begin{cases} \dot{x}(t) = 3x^2(t) + u(t) \\ y(t) = x(t) \end{cases} \tag{46}$$

a) Analysis and application of NCGPC control law: As our goal is to track a desired reference signal ω(t), let us define the error e(t) between the output y(t)and the desired reference signal above ω(t):

$$e(t) = y(t) - \omega(t) \tag{47}$$

This yields a relative degree one that is equal to the dimension of (46) and hence we have no zero dynamics. In order to apply NCGPC control law, consider a new nonlinear system

$$\dot{z}(t) = L_f h(x(t)) - \dot{\omega}(t) + L_g e(x(t))u(t) \tag{48}$$

resulting from the following change of coordinates

$$z(t) = h(x(t)) - \omega(t) \tag{49}$$

Hence we apply NCGPC control law

$$u(x(t)) = \frac{-\sum_{l=0}^{1} K_{1l}(T,\rho)\left[L_f^l h(x(t)) - \omega^{(l)}(t)\right]}{L_g e(t)} \tag{50}$$

where $K_1 = [K_{10} \;\; K_{11}]$ with $K_{11} = 1$. This control law guarantees closed-loop stability through system

$$\dot{z}(t) = -K_{10} z(t) \tag{51}$$

because the corresponding relative degree is one which is less than or equal to four, Chen et al. (2003).

b) Simulation results: Gain matrix K_1 is given in Table 1. as a function of the prediction horizon time T. Fig. 1. shows simulation results.

Table1. Gain matrix K_1

T (s)	K_1
1	[1.5 1]
2	[0.75 1]
3	[0.5 1]
4	[0.375 1]
5	[0.3 1]

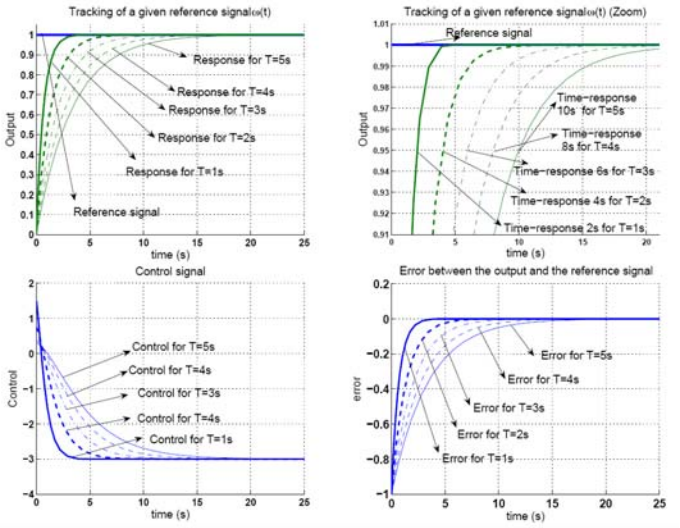


Fig. 1. Time responses, error and control for different values of the prediction horizon time T (1 to 5 seconds).

2) Nonlinear system with relative degree two. Consider the following nonlinear SISO system of dimension two:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = 2x_1^2(t) - 3u(t) \\ y(t) = x_1(t) \end{cases} \tag{52}$$

a) Analysis and application of NCGPC control law: we define an error as (47) with the following change of coordinates

$$\begin{cases} z_1(t) = h(x(t)) - \omega(t) \\ \dot{z}_2 = L_f h(x(t)) - \dot{\omega}(t) \end{cases} \tag{53}$$

and hence we have the following nonlinear system

$$\begin{cases} \dot{z}_1(t) = z_2(t) \\ \dot{z}_2 = L_f^2 h(xt)) - \ddot{\omega}(t) + L_g L_f h(x(t))u(t) \end{cases} \tag{54}$$

Applying control law u to system (54) yields the linear and stable closed-loop system

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -K_{20} & -K_{21} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad (55)$$

where K_{20} and K_{21} are first and second components of the gain matrix K_2 given by $K_2 = [K_{20}\, K_{21}\, K_{22}]$ with $K_{22} = 1$.

b) Simulation results: Gain matrix K_2 is given in Table 2. as a function of the prediction horizon time T

Table 2. Gain matrix K_2

T (s)	K_2
1	[3.33 2.5 1]
2	[0.83 1.25 1]
3	[0.37 0.83 1]
4	[0.21 0.63 1]
5	[0.13 0.5 1]

Fig. 2. presents the behaviour of the stable closed-loop linear system for different values of T. Notice that the tracking of $\omega(t)$ is correct and the percent overshoot (P.O.) is the expected one, indeed, approximately 5.21.

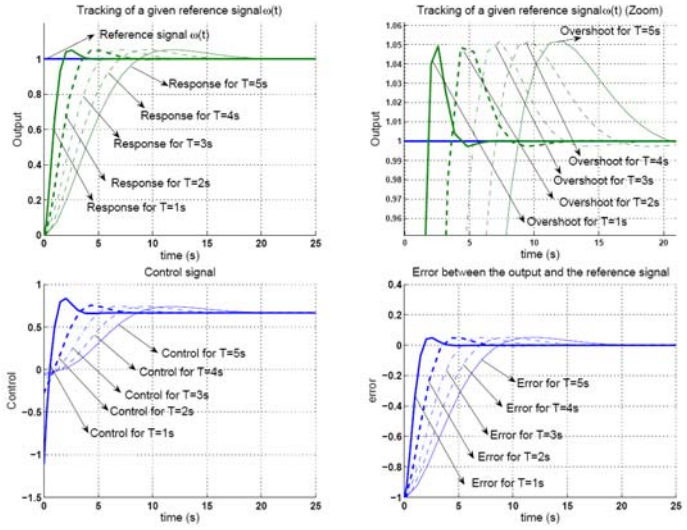


Fig. 2. Overshoot, control and error for different values of the prediction horizon time T (1 to 5 seconds).

6. CONCLUSIONS

In this paper, important properties of NCGPC are highlighted when this latter is applied to nonlinear SISO systems with a relative degree equal to their dimension and lower than three. Future work will investigate the case of nonlinear systems with a relative degree equal or greater than three. We propose to split the resulting characteristic polynomial of (relative) degree greater than two into a product of polynomials of degree one or two.

ACKNOWLEDGMENTS

Prof. N. Langlois gratefully thanks “Région Haute Normandie”, OSEO and FEDER for financially supporting this work and its forthcoming application to diesel engine control in the frame of the ORIANNE (Outils numéRIques pour le mAquettage de foNctions coNtrôle motEur) project labelled by the French competitiveness clusters Mov’eo and Aerospace valley.

REFERENCES

Bruijn, P. M., Bootsma, L. J., and Verbruggen, H. B. (1980). Predictive control using impulse response models. IFAC symposium on digital computer applications to process control, Dusseldorf.

Chen, W. H., Balance, D. J., and Gawthrop, P. J. (2003). Optimal control of nonlinear systems: a predictive control approach. *Automatica*, volume (39), 633-641.

Chen, W. H. (2001). Analytic predictive controllers for nonlinear systems with ill-defined relative degree. *Proc.-Control Theory Appl.*, volume (148), N1.

Clarke, D. W., Mohtadi, C., and Tuffs, P. S. (1987). Generalized Predictive Control-Part I, The Basic Algorithm. *Automatica*, volume (23), 137-148.

Clarke, D. W., Mohtadi, C., and Tuffs, P. S. (1987). Generalized Predictive Control-Part II, Extensions and Interpretation. *Automatica*, volume (23), 149-160.

Cutler, C. R. and Ramaker, B. L. (1980). Dynamic matrix control: a computer control algorithm. JACC, San Francisco.

Dabo, M., Chafouk, H. and Langlois, N. (2009). Unconstrained NCGPC with a guaranteed closed-loop stability: Case of nonlinear SISO systems with the relative degree greater than four. CDC’09, Shanghai.

Demircioglu, H. and Gawthrop, P. J. (1991). Continuous-time Generalized Predictive Control (CGPC). *Automatica*, volume (27), 55-74.

Demircioglu, H. and Gawthrop, P. J. (1992). Multivariable Continuous-time Generalized Predictive Control (MCGPC). *Automatica*, volume (28), 697-713.

Granjon, Y. (2003). *Systèmes linéaires, non linéaires, à temps continu, à temps discrets, représentation d’état*, Dunod, Belgique.

Isidori, A. (1995). *Nonlinear control systems*, Springer Verlag, Englewood Cliffs, New York.

Richalet, J., Rault, A. , Testud, J. L. and Papon, J. (1978). Model predictive heuristic control: application to industrial processes. *Automatica*, volume (14), 413-428.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Improvement of the decoupling feature of decentralized predictive functional control

K. Zabet, R. Haber

Department of Process Engineering and Plant Design, Laboratory of Process Automation,
Cologne University of Applied Science, D-50679 Köln, Betzdorfer Str. 2, Germany
fax: +49-221-8275-2836 and e-mail:robert.haber@ fh-koeln.de, khaled.zabet@smail.fh-koeln.de

Abstract: Two simple decoupling techniques are presented for decentralized PFC (Predictive Functional Control) control of TITO (Two-Input, Two-Output) processes. Both techniques are based on situation or signal dependent adaptation of the controller parameters. By means of first one the desired settling time is tuned in synchronization to a reference signal change. According to the second one the desired settling time is set dependent on the actual control error. The second method makes the synchronization to a set value change superfluous and its realization is therefore very easy.

Keywords: Predictive functional control, settling time, controller adaptation

1. INTRODUCTION

Decoupling in multivariable processes is an important issue. It is desired that one manipulated variable would affect only one controlled variable, while the others would keep their previous values. MIMO (Multi-Input, Multi-Output) controllers can handle this problem using manually designed decoupling controllers or MIMO predictive controller which performs the decoupling automatically.

Multivariable processes are often controlled by SISO controllers, because these are easier to realize than MIMO controllers. The question arises how the decoupling can be improved without complicated multivariable controller design. Maurath et al. (1986) recommended some partly complicated methods for improved decoupling. In this paper two different methods are recommended for decentralized multivariable control. The SISO controller is realized by PFC (Predictive Functional Control) (Richalet and O'Donavan, 2009), as PFC is a very effective SISO controller which can also handle constraints.

The paper is structured as follows. In Section 2 the SISO PFC algorithm is shown. In section 3 a TITO process is controlled with fixed decentralized controller parameters. In Sections 4 and 5 two different methods are shown how the controller parameters can be adapted to decrease the coupling effect.

2. PREDICTIVE FUNCTIONAL CONTROL

The principle of PFC is that the controlled variable y achieves the reference trajectory at the target point (or points) using one change (or minimal number of changes) in the manipulated variable u. The desired change in the controlled variable y during the prediction horizon n_p (from the actual time k) is calculated from the desired change of the reference trajectory and the predicted change of the model output y_m. The manipulated variable u can be calculated easily from the change of the reference trajectory and the predicted change of the model output in the prediction point, see Fig.1.

A PT1 (proportional, first-order) process without dead time (chosen for simplicity) is described in discrete-time as

$$y(k) = -a\, y(k-1) + K_p (1+a) u(k-1) \tag{1}$$

where y is the process output, u is the process input, a is the discrete-time process parameter and K_p is the static gain of the process.

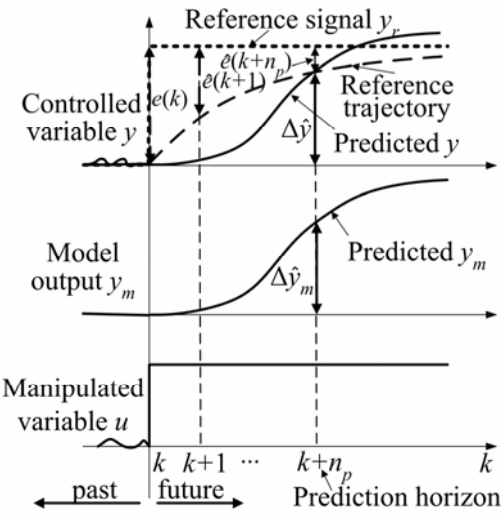


Fig.1. PFC principle

The desired changes in the controlled output y during n_p can be defined supposing that y reaches the reference trajectory at the target point (n_p step ahead) as follows:

$$\hat{y}(k+n_p \mid k) - y(k) = e(k) - \hat{e}(k+n_p \mid k) \tag{2}$$

where $e(k) = y_r(k) - y(k)$ and y_r is the constant reference signal.

The reference trajectory can be chosen an exponential function for simplicity. Then the control error is decreasing monotonously:

$$\begin{aligned}\hat{e}(k+1|k) &= \lambda_r e(k)\\ \hat{e}(k+2|k) &= \lambda_r \hat{e}(k+1|k) = \lambda_r^2 e(k)\\ &\dots\\ \hat{e}(k+n_p|k) &= \lambda_r^{n_p} e(k)\end{aligned} \quad (3)$$

where λ_r is the reduction ratio of the control error.

The reference trajectory provides the settling time $t_{95\%}=T_c$ for the closed loop control system if $\lambda_r = \exp(-3\Delta t/T_c)$, where Δt is the sampling time.

From (2) and (3), the desired change in y is defined as follows:

$$\Delta\hat{y}(k+n_p|k) = \hat{y}(k+n_p|k) - y(k) = (1-\lambda_r^{n_p})e(k) \quad (4)$$

The changes of y can be predicted also using the process model equation:

$$y_m(k) = -a_m\, y_m(k-1) + K_m(1+a_m)u(k-1) \quad (5)$$

where y_m is the process model output, a_m is the discrete process model parameter and K_m is the static gain of the process model.

Supposing that the actual input signal u is kept constant during the prediction horizon, the predicted model output becomes after n_p steps:

$$\begin{aligned}\hat{y}_m(k+1|k) &= -a_m\, y_m(k) + K_m(1+a_m)u(k)\\ \hat{y}_m(k+2|k) &= -a_m\, y_m(k+1|k) + K_m(1+a_m)u(k)\\ &= (-a_m)^2 y_m(k) + (-a_m+1)K_m(1+a_m)u(k)\\ &= (-a_m)^2 y_m(k) + K_m[1-(-a_m)^2]u(k)\\ &\dots\\ \hat{y}_m(k+n_p|k) &= (-a_m)^{n_p} y_m(k) + K_m[1-(-a_m)^{n_p}]u(k)\end{aligned}$$

Then, the predicted change in y_m becomes:

$$\begin{aligned}\Delta\hat{y}_m(k+n_p|k) &= \hat{y}_m(k+n_p|k) - y_m(k)\\ &= [1-(-a_m)^{n_p}][K_m u(k) - y_m(k)]\end{aligned} \quad (6)$$

Simple comparison between the predicted change of the reference trajectory in (4) and the predicted change of y_m in (6) results in the manipulated variable:

$$u(k) = k_0[y_r - y(k)] + k_1 y_m(k) \quad (7a)$$

where:

$$k_0 = \frac{1-\lambda_r^{n_p}}{K_m[1-(-a_m)^{n_p}]}, \qquad k_1 = \frac{1}{K_m} \quad (7b)$$

If the process has dead time $d = d_m$ then $y(k)$ in (7a) has to be replaced by $\hat{y}(k+d_m|k)$

$$\hat{y}(k+d_m|k) = y(k) + [y_m(k) - y_m(k-d_m)] \quad (8)$$

In case of higher-order aperiodic processes the transfer function can be partitioned in parallel connection of first-order processes

$$\hat{y}_{m,i}(k) = -a_{m,i}\hat{y}_{m,i}(k-1) + K_{m,i}(1+a_{m,i})u(k-1) \quad (9)$$

with the corresponding parameters $K_{m,i}$ and $a_{m,i}$ of the i-th sub-process. (If the process has multiple poles then different but very similar poles have to be assigned to each multiple pole.)

The basic algorithm can be easily extended for this case, as well (Khadir and Ringwood, 2008):

$$u(k) = k_0[y_r - y(k)] + \sum_{i=1}^{n} k_i y_{m,i}(k) \quad (10a)$$

where n is the order of the process,

$$k_0 = \frac{1-\lambda_r^{n_p}}{\sum_{j=1}^{n} K_{m,j}[1-(-a_{m,j})^{n_p}]},\ k_i = \frac{1-(-a_{m,i})^{n_p}}{\sum_{j=1}^{n} K_{m,j}[1-(-a_{m,j})^{n_p}]} \quad (10b)$$

3. DECENTRALIZED CONTROL WITH FIXED CONTROLLER PARAMETERS

In order to illustrate the problem of coupling a TITO process model (Fig.2) is considered. The two controlled variables (y_1, y_2) are controlled by the two manipulated variables (u_1, u_2).

The sub-models are aperiodic processes with different static gains K_{pij}, time constants T_{ij}, and dead times T_{dij}. All processes have some (n_{ij}) equal time constants:

- P_{11}: K_{p11}=1.5, T_{11}=1.0 min, n_{11}=2, T_{d11}=0.1 min
- P_{12}: K_{p12}=0.5, T_{12}=0.5 min, n_{12}=4, T_{d12}=0.5 min
- P_{21}: K_{p21}=0.75, T_{21}=0.5 min, n_{21}=3, T_{d21}=0.8 min
- P_{22}: K_{p22}=1.0, T_{22}=2.0 min, n_{22}=1, T_{d22}=0.2 min

The block diagram of the process is shown in Fig.2.

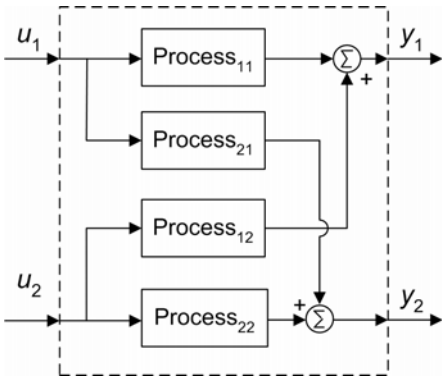


Fig. 2. TITO process model

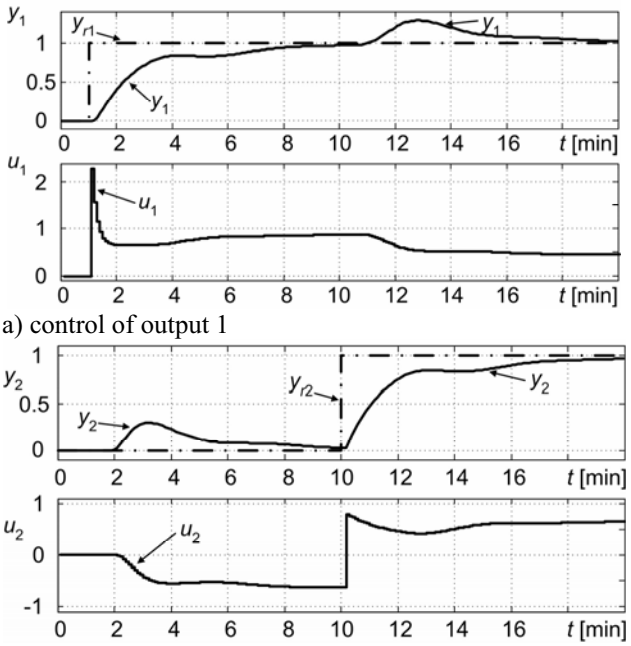
Decentralized TITO control was used with two independent SISO controllers with fixed parameters. The sampling time was Δt=0.1 min and the controller parameters of the two SISO controllers were:

- prediction horizons: n_{p1}=5 and n_{p2}=2,
- desired settling times: $T_{c1}=T_{c2}$=4 min.

The control scenario was:

- at t=1 min stepwise increase of the reference signal of y_1 from 0 to 1,
- at t=10 min stepwise increase of the reference signal of y_2 from 0 to 1.

Fig. 3 shows the decentralized TITO predictive control with the above parameters.



a) control of output 1

b) control of output 2

Fig.3. Decentralized TITO control with constant controller parameters

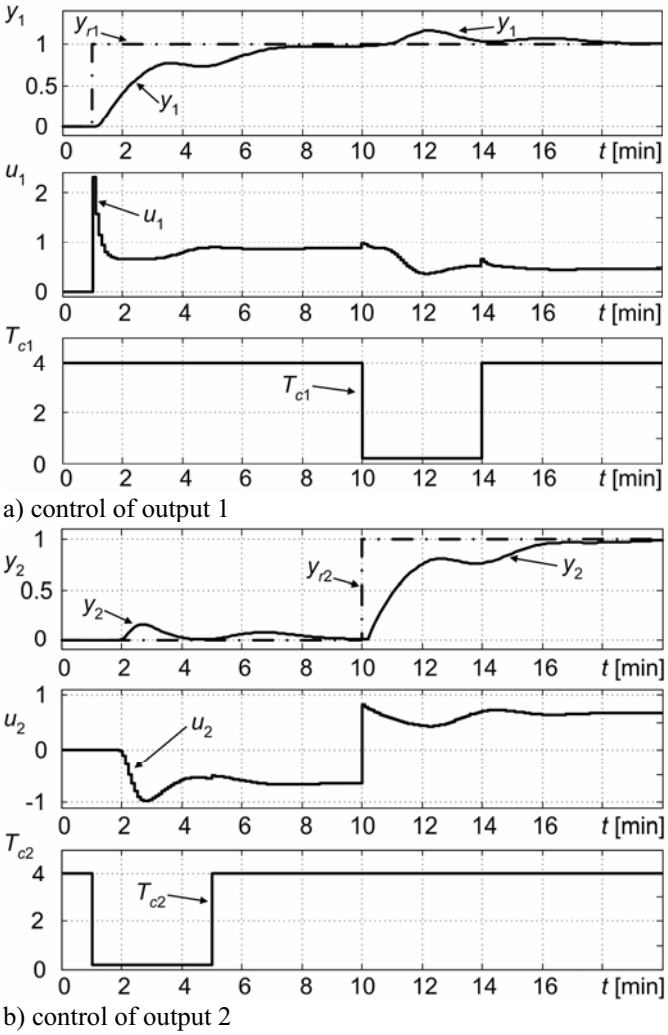
The control of the reference signal changes is slow and aperiodic (settling time $t_{95\%}$≈7.2 min). There are changes of about 30% (related to the reference signal changes) with settling time $t_{95\%}$≈7 min in the controlled variables whose reference signal was kept constant.

4. REFERENCE SIGNAL CHANGE-DEPENDENT CONTROLLER PARAMETER ADAPTION

The main controller parameter with PFC is the desired settling time T_c.. The decoupling ability with a TITO process can be improved by tuning the settling times (T_{c1} and T_{c2}).

Decreasing of the desired settling time of the controlled variable whose reference signal was kept constant accelerates the control and hence reduces the control error in this controlled variable. Fig. 4 illustrates this case for reference signal changes. The controller parameters (desired settling time) of both controlled variables were changed from T_{c1}=T_{c2}=4 min to T_{c1}=T_{c2}=0.2 min for that controlled variable whose reference signal was not changed in the moment of the change of the other reference signal. The duration of the change was 4 min which is equal to the desired longer settling time. The control of the reference signal changes is faster than before ($t_{95\%}$≈5.8 min) and the control error is about 16% (related to the reference signal changes) with $t_{95\%}$ about 1.9 to 2.8 min in the controlled variables whose reference signal was kept constant. The plots show that the two processes are better decoupled than in Fig. 3 where the controller parameters were kept constant.

The critical point of this method is the detection of the reference signal change. Sometimes this time point is known by the technology in advance. Otherwise a reference signal change can be detected with methods of signal analysis. Nevertheless a method which does not have to care about the time point of the reference signal change would be preferable.



a) control of output 1

b) control of output 2

Fig.4. Decentralized TITO control with changing of the desired settling time at reference signal steps

5. CONTROL ERROR-DEPENDENT CONTROLLER PARAMETER ADAPTION

The synchronization at the reference signal change can be performed automatically if the desired settling times are decentralized functions of the control errors. With a stepwise change of the reference signal the control error of the related controlled variable is increased faster than the control error of the other controlled variable whose reference signal was kept constant. Consequently, if the settling time is set proportional to the control error for both controlled variables then after a

stepwise change of a reference signal the settling time of the controlled variable whose reference signal was changed will be higher than the settling time of the controlled variable whose reference signal was not changed. Consequently the controlled variable whose reference signal was not changed will be controlled faster that acts as a forced decoupling.

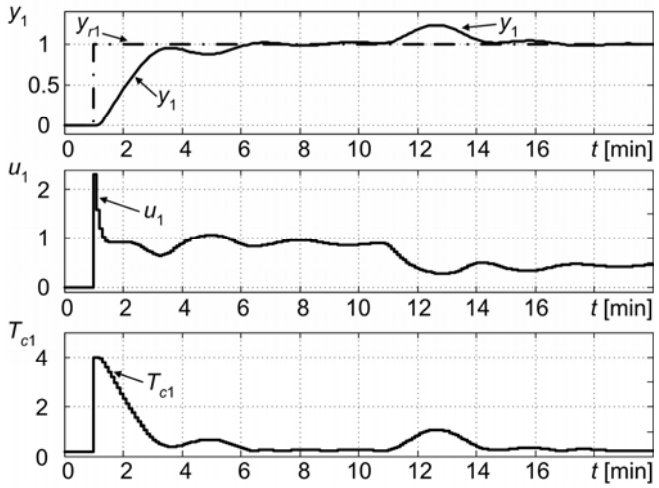
The following linear dependence of the desired settling times on the control error were applied in the simulation:

$$T_{ci} = T_{ci,\min} + \left(T_{ci,\max} - T_{ci,\min}\right) \cdot \left|e_i(k)\right| \tag{11}$$

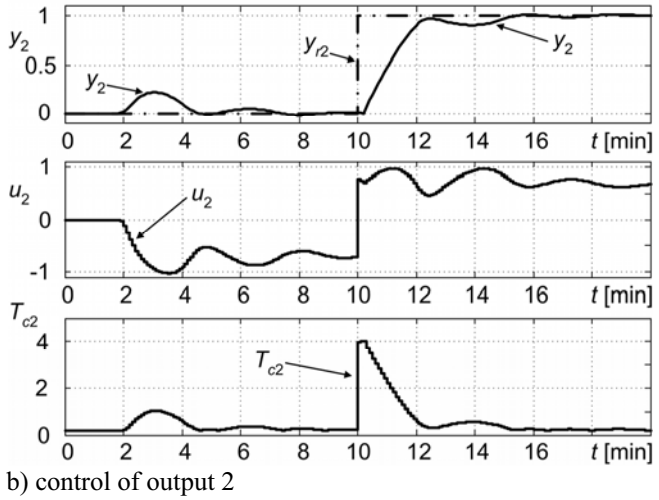
with $T_{c1,\max}= T_{c2,\max}=4$ min and $T_{c1,\min}= T_{c2,\min}=0.2$ min.

Eq. (11) shows that the control is fast if there is no control error (in steady-state) or with small control error.

Fig. 5 shows that the control of the variables whose set value was changed is slightly faster than with the first method ($t_{95\%}\approx 4.7$ min) and the control error is about 22% (related to the reference signal changes) with $t_{95\%}$ of about 2.4 to 3.2 min in the controlled variables whose reference signal was kept constant).



a) control of output 1



b) control of output 2

Fig.5. Decentralized TITO control with the control error-dependent desired settling time

The plot shows that the second method is not as good as the first method but the decoupling effect became much better in comparison with Fig. 3, where no controller parameter was changed. As mentioned already the realization of this control error dependent adaptation is easier than detecting changes in a set value. It is interesting that if using the control error dependent adaptation of the controller parameters not only the control of those controlled variables became faster whose reference signal remained constant but also of those whose set value was changed.

6. CONCLUSION

New methods were presented for reducing the decoupling effect of decentralized TITO PFC control with proper adaptation of the controller parameters. The two methods (1) reference signal change-dependent settling time, and (2) control error-dependent settling time were presented and simulated. Both methods have shown improved decoupling effects: less control error and faster control of the controlled signal whose reference signal was not changed. With the first method the control error was a bit smaller and achieved its final value a bit faster than with the second method. With the second method the controlled variable whose set value was changed achieved its final value faster than with the first method or without any adaptation of the controller parameters. The only disadvantage was the more turbulent manipulated signal which is now subject of further investigations. In addition it is very easy to realize the second method in practice. The presented idea can also be extended for processes with more than 2 controlled signals.

Similar method has been successfully applied by Schmitz et al. (2007) with GPC (Generalized Predictive Control).

The next step will be the application of the new controller tuning method for (not decentralized) TITO PFC control, when even better decoupling ability is expected than here.

7. ACKNOWLEDMENTS

The first author gratefully acknowledges the scholarship of the General People's Committee for Higher Education, Great Socialist People's Libyan Arab Jamahiriya.

REFERENCES

Khadir, M.T., Ringwood, J.V. (2008). Extension of first order predictive functional controllers to handle higher order internal models, Int. J. Appl. Math. Comput. Sci., Vol. 18, No. 2, 229–239.

Maurath, P. R., D. E. Seborg and D. A. Mellichamp (1986). Achieving Decoupling with Predictive Controllers. Proc. Amer. Control Conf., Seattle, 1372-1377.

Richalet, J., O'Donavan, D. (2009). Elementary Predictive Functional Control. Springer Verlag, Berlin, ISBN: 978-1-84882-492-8.

Schmitz U., Haber R., Arousi F., Bars R. (2007). Decoupling predictive control by error dependent tuning of the weighting factors, AT&P Journal PLUS, 2, Bratislava, 131 – 140.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Equality Constraints in Sensor Faults Reconfigurable Control Design

D. Krokavec, A. Filasová

Department of Cybernetics and Artificial Intelligence,
Technical University of Košice, Košice, Slovak Republic
(e-mail: dusan.krokavec@tuke.sk,anna.filasova@tuke.sk)

Abstract: The paper presents the design method based on the memory-less feedback control for stabilization of discrete-time systems with a sensor fault, where the fault is described by an equality constraint given on the state variable associated with the faulty sensor. The design conditions are presented in the form of linear matrix inequalities. The validity of the proposed method is demonstrated by a numerical example with an equality constraint setting on the faulty state variable sensor.

Keywords: Reconfigurable control, equality constraints, linear matrix inequalities, state feedback, singular systems.

1. INTRODUCTION

Modern technological processes rely on sophisticated control systems to meet increased performance and safety requirements. A conventional control for a complex system may result in an unsatisfactory performance, or even instability, in the event of malfunctions in actuators, sensors or other system components. To overcome such weaknesses, new approaches to control system design have to be developed in order to tolerate component malfunctions while maintaining stability and acceptable performance properties. These types of closed-loop control systems are known as fault-tolerant control systems (FTCS) having the ability to accommodate component failures automatically. Bibliographical reviews can be found in Jiang (2005); Patton (1997); Zhang and Jiang (2003), new developments in fault-tolerant control methods are presented e.g. in Benítéz-Pérez and García-Nocetti (2005); Blanke et al. (2003); Krokavec and Filasová (2007); Noura *et al.* (2009); Simani et al. (2003).

In the last years many significant results have spurred interest in the problem of determining control laws for the systems with constraints. For the typical case where a system state reflects a certain physical entities this class of constraints rises because of physical limits and these ones usually keep the system state in a region of the technological conditions. Subsequently this problem can be formulated using technique dealing with the system state constraints directly, where it can be coped with efficiently using linear system techniques (Ko and Bitmead (2007)). Therefore, a special form of the constrained problems can be so formulated with the goal to optimize the reconfigurable control structure while the system state variables satisfy the equality constraints Krokavec and Filasová (2008, 2009). This design task is specified as a singular one and associated methods have to be used to design the controller parameters.

A number of problems that arise in the state feedback control, possibly formulated using Lyapunov function, bounded real lemma, etc. can be reduced to a handful of standard convex and quasi-convex problems that involve matrix inequalities. It is known that the optimal solution can be computed by using interior point methods (Nesterov and Nemirovsky (1994)) which converge in polynomial time with respect to the problem size, and efficient interior point algorithms have recently been developed for and further development of algorithms for these standard problems is an area of active research. Some progres review in this research field one can find in Boyd *et al.* (1994); Skelton et al. (1998), and the references therein.

This paper is concerned with the problem of reconfigurable control design while the state variable associated with the faulty sensor is described by an equality constraint. Based on the discrete-time state description the attention is focused on the memory-less feedback control parameter optimization. It is assumed that the system is free of the actuator faults, and according to the performance of active FTCS it is supposed that the state variable sensor faults detection and isolation schemes are available. Controller switching is taking into account since such different faulty system representations is known, and stabilizing controllers are pre-computed off-line.

The developed design method starts with adaptation of the methodology given in Ko and Bitmead (2007) and can be noted as an extension method to the pseudo-inverse methods (PIM) considered e.g. in Staroswiecki (2005), as well as to the degraded reference models used in Zhang and Jiang (2003). Using simple regularization the eigenstructure assignment method is adapted, and, in addition to a single sensor fault, the optimized control law parameters design conditions are derived. Finally the numerical example is shown to demonstrate the role of such equality constraints in the design procedure.

⋆ The work presented in this paper was supported by VEGA, Grant Agency of Ministry of Education and Academy of Science of Slovak Republic under Grant No. 1/0328/08. This support is very gratefully acknowledged.

2. PROBLEM FORMULATION

The systems under consideration are understood as the multi-input and multi-output linear (MIMO) dynamic systems with single sensor faults. Without lose of generality this class of the discrete-time linear dynamic system can be represented in the state-space form as

$$\boldsymbol{q}(i+1) = \boldsymbol{F}\boldsymbol{q}(i) + \boldsymbol{G}\boldsymbol{u}(i) \tag{1}$$

$$\boldsymbol{y}(i) = \boldsymbol{C}\boldsymbol{q}(i) + \boldsymbol{f}_h(i) = \boldsymbol{C}_h\boldsymbol{q}(i) \tag{2}$$

where $\boldsymbol{q}(i) \in I\!R^n$, $\boldsymbol{u}(i) \in I\!R^r$, and $\boldsymbol{y}(i) \in I\!R^m$ are vectors of the state, input and objective output variables, respectively, and matrices $\boldsymbol{F} \in I\!R^{n\times n}$, $\boldsymbol{G} \in I\!R^{n\times r}$, and $\boldsymbol{C} \in I\!R^{n\times n}$ are real matrices. It is supposed that the system is without actuator faults and a monitored sensor fault is modelled by an additive vector $\boldsymbol{f}_h(i) \in I\!R^r$.

The standard linear memory-less state feedback controllers of the form

$$\boldsymbol{u}(i) = -\boldsymbol{K}\boldsymbol{q}(i) \tag{3}$$

are used in control reconfiguration structure, and throughout the paper it is assumed that the couple $(\boldsymbol{F}, \boldsymbol{G})$ is controllable and all state variables are measurable.

3. PRELIMINARIES

Proposition 1. (e.g. see Skelton et al. (1998)) Let $\boldsymbol{\Lambda}$ is a matrix variable and $\boldsymbol{A}$, $\boldsymbol{B}$ are known non-square matrices of appropriate dimensions such the equality

$$\boldsymbol{B}\boldsymbol{\Lambda} = \boldsymbol{A} \tag{4}$$

can be set. Then all solutions to $\boldsymbol{\Lambda}$ are

$$\boldsymbol{\Lambda} = \boldsymbol{B}^{\ominus 1}\boldsymbol{A} + (\boldsymbol{I} - \boldsymbol{B}^{\ominus 1}\boldsymbol{B})\boldsymbol{\Lambda}^{\circ} \tag{5}$$

where

$$\boldsymbol{B}^{\ominus 1} = \boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1} \tag{6}$$

is Moore-Penrose pseudoinverse of $\boldsymbol{B}$, and $\boldsymbol{\Lambda}^{\circ}$ is an arbitrary matrix of appropriate dimension.

Proof. Supposing that the product $\boldsymbol{B}\boldsymbol{B}^T$ is a regular matrix, then pre-multiplying left-hand side of (4) by the identity matrix gives

$$\boldsymbol{B}\boldsymbol{\Lambda} = \boldsymbol{B}\boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1}\boldsymbol{A} \tag{7}$$

and using notation (6) it yields

$$\boldsymbol{\Lambda} = \boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1}\boldsymbol{A} = \boldsymbol{B}^{\ominus 1}\boldsymbol{A} \tag{8}$$

Let $\boldsymbol{\Lambda}^{\circ}$ is a matrix of appropriate dimension such that substituting in (7) results in

$$\boldsymbol{B}\boldsymbol{\Lambda}^{\circ} = \boldsymbol{B}\boldsymbol{B}^{\ominus 1}\boldsymbol{A} = \boldsymbol{B}\boldsymbol{B}^{\ominus 1}\boldsymbol{B}\boldsymbol{\Lambda}^{\circ} \tag{9}$$

Thus

$$\boldsymbol{B}(\boldsymbol{I} - \boldsymbol{B}^{\ominus 1}\boldsymbol{B})\boldsymbol{\Lambda}^{\circ} \doteq \boldsymbol{0} \tag{10}$$

$$(\boldsymbol{I} - \boldsymbol{B}^{\ominus 1}\boldsymbol{B})\boldsymbol{\Lambda}^{\circ} \doteq \boldsymbol{0} \tag{11}$$

respectively, where $\boldsymbol{I}$ is the identity matrix of appropriate dimension. Therefore, for an arbitrary $\boldsymbol{\Lambda}^{\circ}$ of appropriate dimension (8), (11) implies (5). □

Note, matrix pseudoinverse is generalized for a singular matrix $\boldsymbol{B}\boldsymbol{B}^T$.

Proposition 2. Let $\boldsymbol{E} \in I\!R^{n\times n}$ is a real square matrix with non-repeated eigenvalues, satisfying the equality constraint

$$\boldsymbol{d}^T\boldsymbol{E} = 0 \tag{12}$$

Then one from its eigenvalues is zero, and (normalized) $\boldsymbol{d}^T$ is the left raw eigenvector of $\boldsymbol{E}$ associated with this zero eigenvalue.

Proof. (e.g. see Krokavec and Filasová (2007)) If $\boldsymbol{E} \in \in I\!R^{n\times n}$ is a real square matrix having non-repeated eigenvalues the eigenvalue decomposition of $\boldsymbol{E}$ takes the form

$$\boldsymbol{E} = \boldsymbol{N}\boldsymbol{Z}\boldsymbol{M}^T \tag{13}$$

$$\boldsymbol{N} = [\boldsymbol{n}_1 \ \cdots \ \boldsymbol{n}_n], \ \boldsymbol{M} = [\boldsymbol{m}_1 \ \cdots \ \boldsymbol{m}_n], \ \boldsymbol{M}^T\boldsymbol{N} = \boldsymbol{I} \tag{14}$$

$$\boldsymbol{Z} = \operatorname{diag}[z_1 \ \cdots \ z_n] \tag{15}$$

where $\boldsymbol{n}_l$ is right eigenvector, and $\boldsymbol{m}_l^T$ is left eigenvector associated with the eigenvalue z_l of $\boldsymbol{E}$, $l = 1, 2, \ldots n$. Then (12) can be rewritten as

$$\boldsymbol{d}^T[\boldsymbol{n}_1 \cdots \boldsymbol{n}_h \cdots \boldsymbol{n}_n] \operatorname{diag}[z_1 \cdots z_h \cdots z_n]\boldsymbol{M}^T = 0 \tag{16}$$

If $\boldsymbol{d}^T = \boldsymbol{m}_h^T$ then the orthogonal property (14) implies

$$[0_1 \ \cdots \ 1_h \ \cdots 0_n] \operatorname{diag}[z_1 \ \cdots \ z_h \ \cdots \ z_n]\,\boldsymbol{M}^T = 0 \tag{17}$$

and it is evident that (17) be satisfied only if $z_h = 0$.

Note this can be easily proven for a square matrix with one zero eigenvalue and some repeated eigenvalues. □

4. CONSTRAINED CONTROL

Using the control law of the form (3) the equilibrium control equations take the forms

$$\boldsymbol{q}(i+1) = (\boldsymbol{F} - \boldsymbol{G}\boldsymbol{K})\boldsymbol{q}(i) \tag{18}$$

$$\boldsymbol{y}(i) = \boldsymbol{C}\boldsymbol{q}(i) \tag{19}$$

and prescribed by a matrix $\boldsymbol{D} \in I\!R^{k\times n}$, $\operatorname{rank}(\boldsymbol{D}) = k < r$ there it is considered the design constraint

$$\boldsymbol{q}(i) \in \mathcal{N}_{\boldsymbol{D}} = \{\boldsymbol{q} : \boldsymbol{D}\boldsymbol{q} = \boldsymbol{0}\} \tag{20}$$

where the state vectors have to satisfy equalities

$$\boldsymbol{D}\boldsymbol{q}(i+1) = \boldsymbol{D}(\boldsymbol{F} - \boldsymbol{G}\boldsymbol{K})\boldsymbol{q}(i) = \boldsymbol{0} \tag{21}$$

for $i = 1, 2, \ldots$. It is supposed the matrix $\boldsymbol{D}$ is chosen by such way that

$$\boldsymbol{D}(\boldsymbol{F} - \boldsymbol{G}\boldsymbol{K}) = \boldsymbol{0} \ \Rightarrow \ \boldsymbol{D}\boldsymbol{F} = \boldsymbol{D}\boldsymbol{G}\boldsymbol{K} \tag{22}$$

respectively, as well as that the closed-loop system matrix $\boldsymbol{F}_c = (\boldsymbol{F} - \boldsymbol{G}\boldsymbol{K})$ is stable (all its eigenvalues lie in the unit circle in the complex plane $\mathcal{Z}$). Therefore, $\mathcal{N}_{\boldsymbol{D}}$ be the constrain subspace, and state be constrained in this subspace (the null space of $\boldsymbol{D}$). Under these conditions the system state stays within the subspace, i.e. $\boldsymbol{q}(i)$, $\boldsymbol{F}\boldsymbol{q}(i) \in \mathcal{N}_{\boldsymbol{D}}$, respectively.

Solving (22) with respect to $\boldsymbol{K}$ then (5) implies all solutions of $\boldsymbol{K}$ as follows

$$\boldsymbol{K} = (\boldsymbol{D}\boldsymbol{G})^{\ominus 1}\boldsymbol{D}\boldsymbol{F} + (\boldsymbol{I} - (\boldsymbol{D}\boldsymbol{G})^{\ominus 1}\boldsymbol{D}\boldsymbol{G})\boldsymbol{K}^{\circ} \tag{23}$$

where $\boldsymbol{K}^{\circ}$ is a design parameter matrix. Thus, it is possible to express (23) as

$$\boldsymbol{K} = \boldsymbol{J} + \boldsymbol{L}\boldsymbol{K}^{\circ} \tag{24}$$

with

$$\boldsymbol{J} = (\boldsymbol{D}\boldsymbol{G})^{\ominus 1}\boldsymbol{D}\boldsymbol{F}, \quad \boldsymbol{L} = \boldsymbol{I} - (\boldsymbol{D}\boldsymbol{G})^{\ominus 1}\boldsymbol{D}\boldsymbol{G} \tag{25}$$

where $\boldsymbol{L}$ is the projection matrix (the orthogonal projector onto the null space $\mathcal{N}_{\boldsymbol{D}\boldsymbol{G}}$ of $\boldsymbol{D}\boldsymbol{G}$).

Fact 1. Seeking a control policy of the form

$$\boldsymbol{u}(i) = -\boldsymbol{K}\boldsymbol{q}(i) + \boldsymbol{N}_w\boldsymbol{w}(i) \tag{26}$$

where $\boldsymbol{N}_w \in I\!R^{m\times r}$, $\boldsymbol{w}(i) \in I\!R^r$, then (21) implies

$$\begin{aligned} &\boldsymbol{D}\boldsymbol{q}(i+1) = \\ &= \boldsymbol{D}(\boldsymbol{F}-\boldsymbol{G}\boldsymbol{K})\boldsymbol{q}(i) + \boldsymbol{D}\boldsymbol{G}\boldsymbol{N}_w\boldsymbol{w}(i) = \boldsymbol{D}\boldsymbol{G}\boldsymbol{N}_w\boldsymbol{w}(i) \end{aligned} \tag{27}$$

and it is evident that the system steady state is not zero, but proportional to steady state of $\boldsymbol{w}$.

5. RECONFIGURABLE CONTROL

To design the reconfigurable control law it is supposed that the system (1), (2) is without actuator faults and a single fault (the h-th sensor fault) is interpreted by the constraint (e.g. see Krokavec and Filasová (2008))

$$\boldsymbol{q}(i) \in \mathcal{N}_{\boldsymbol{d}_h^T} = \{\boldsymbol{q} : \boldsymbol{d}_h^T\boldsymbol{q} = 0\} \tag{28}$$

$$\boldsymbol{d}_h^T = [\, 0 \ \cdots \ 0 \ \ 1_h \ \ 0 \ \cdots \ 0 \,] \tag{29}$$

where $h \in \{1, 2, \ldots, n\}$. This interpretation means that the h-th sensor output is stuck at zero because of the h-th sensor malfunction.

The above defined constraints set modifies (25) as follows

$$\boldsymbol{J}_h = (\boldsymbol{d}_h^T\boldsymbol{G})^{\ominus 1}\boldsymbol{d}_h^T\boldsymbol{F} \tag{30}$$

$$\boldsymbol{L}_h = \boldsymbol{I} - (\boldsymbol{d}_h^T\boldsymbol{G})^T\left(\boldsymbol{d}_h^T\boldsymbol{G}(\boldsymbol{d}_h^T\boldsymbol{G})^T\right)^{-1}\boldsymbol{d}_h^T\boldsymbol{G} \tag{31}$$

and

$$\boldsymbol{K}_h = \boldsymbol{J}_h + \boldsymbol{L}_h\boldsymbol{K}_h^{\circ} \tag{32}$$

The proof of the following theorem can be found in Krokavec and Filasová (2009), and implies from properties of the Proposition 2.

Theorem 1. Designing with respect to $\boldsymbol{d}_h^T$ having structure (29) the characteristic polynomial of the closed-loop system be

$$P(z) = \det(z\boldsymbol{I}_n - \boldsymbol{F}_{ch}) = z\det\left(z\boldsymbol{I}_{n-1} - \boldsymbol{W}_h\right) \tag{33}$$

where $\boldsymbol{F}_{ch} = \boldsymbol{F} - \boldsymbol{G}\boldsymbol{K}_h$ is the closed-loop system matrix, and $\boldsymbol{W}_h$ is its h-th principal minor.

6. RECONFIGURABLE CONTROL DESIGN

Theorem 2. For the system (1), (2) the sufficient condition for a stable reconfigurable control is that there exist a positive definite symmetric matrix $\boldsymbol{Y}_h > 0$, $\boldsymbol{Y}_h \in I\!R^{n\times n}$, and a matrix $\boldsymbol{Z}_h \in I\!R^{r\times n}$ such that

$$\boldsymbol{Y}_h = \boldsymbol{Y}_h^T > 0 \tag{34}$$

$$\begin{bmatrix} -\boldsymbol{Y}_h & \boldsymbol{Y}_h(\boldsymbol{F}-\boldsymbol{G}\boldsymbol{J}_h)^T - \boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{G}^T \\ * & -\boldsymbol{Y}_h \end{bmatrix} < 0 \tag{35}$$

where $\boldsymbol{J}_h$, $\boldsymbol{L}_h$ are defined in (30), (31), respectively.

Thus, $\boldsymbol{K}_h^{\circ}$ can be computed as

$$\boldsymbol{K}_h^{\circ} = \boldsymbol{Z}_h\boldsymbol{Y}_h^{-1} \tag{36}$$

and the control law gain matrix $\boldsymbol{K}_h$ is given as in (32).

Proof. Defining Lyapunov function as follows

$$v(\boldsymbol{q}(i)) = \boldsymbol{q}^T(i)\boldsymbol{P}_h\boldsymbol{q}(i) > 0 \tag{37}$$

where $\boldsymbol{P}_h = \boldsymbol{P}_h^T > 0$, $\boldsymbol{P}_h \in I\!R^{n\times n}$, then the forward difference along a solution of (1) is

$$\Delta v(\boldsymbol{q}(i)) = \boldsymbol{q}^T(i+1)\boldsymbol{P}_h\boldsymbol{q}(i+1) - \boldsymbol{q}^T(i)\boldsymbol{P}_h\boldsymbol{q}(i) < 0 \tag{38}$$

$$\Delta v(\boldsymbol{q}(i)) = \boldsymbol{q}^T(i)(\boldsymbol{F}_{ch}^T\boldsymbol{P}_h\boldsymbol{F}_{ch} - \boldsymbol{P}_h)\boldsymbol{q}(i) < 0 \tag{39}$$

respectively, where

$$\boldsymbol{F}_{ch} = \boldsymbol{F} - \boldsymbol{G}\boldsymbol{J}_h - \boldsymbol{G}\boldsymbol{L}_h\boldsymbol{K}_h^{\circ} \tag{40}$$

and (39) implies

$$\boldsymbol{P}_{ch}^{\circ} = \boldsymbol{F}_{ch}^T\boldsymbol{P}_h\boldsymbol{F}_{ch} - \boldsymbol{P}_h < 0 \tag{41}$$

Therefore, using Schur complement property it yields

$$\boldsymbol{P}_{ch}^{\circ} = \begin{bmatrix} -\boldsymbol{P}_h & (\boldsymbol{F}-\boldsymbol{G}\boldsymbol{J}_h-\boldsymbol{G}\boldsymbol{L}_h\boldsymbol{K}_h^{\circ})^T \\ * & -\boldsymbol{P}_h^{-1} \end{bmatrix} < 0 \tag{42}$$

Defining the congruence transform matrix

$$\boldsymbol{T}_c = \mathrm{diag}\left[\, \boldsymbol{P}_h^{-1} \ \ \boldsymbol{I}_n \,\right] \tag{43}$$

and multiplying right-hand and left-hand side of (42) by $\boldsymbol{T}_c$ it can be obtained

$$\begin{bmatrix} -\boldsymbol{P}_h^{-1} & \boldsymbol{P}_h^{-1}(\boldsymbol{F}-\boldsymbol{G}\boldsymbol{J}_h-\boldsymbol{G}\boldsymbol{L}_h\boldsymbol{K}_h^{\circ})^T \\ * & -\boldsymbol{P}_h^{-1} \end{bmatrix} < 0 \tag{44}$$

and with notation

$$\boldsymbol{P}_h^{-1} = \boldsymbol{Y}_h > 0, \qquad \boldsymbol{K}_h^{\circ}\boldsymbol{P}_h^{-1} = \boldsymbol{Z}_h \tag{45}$$

(44) implies (35). □

7. LIMITS IN GAIN NORM

Theorem 3. For the system (1), (2) a stable reconfigurable control with norm bounded $\boldsymbol{K}_h$ exists if there exist a positive definite symmetric matrix $\boldsymbol{Y}_h > 0$, $\boldsymbol{Y}_h \in I\!R^{n\times n}$, a matrix $\boldsymbol{Z}_h \in I\!R^{r\times n}$ and scalars $\mu_{h1} > 0$, $\mu_{h2} > 0$, $\mu_1, \mu_2 \in I\!R$ such that

$$\boldsymbol{Y}_h = \boldsymbol{Y}_h^T > 0, \quad \mu_{h1} > 0, \quad \mu_{h2} > 0 \tag{46}$$

$$\begin{bmatrix} -\boldsymbol{Y}_h & \boldsymbol{Y}_h(\boldsymbol{F}-\boldsymbol{G}\boldsymbol{J}_h)^T - \boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{G}^T \\ * & -\boldsymbol{Y}_h \end{bmatrix} < 0 \tag{47}$$

$$\begin{bmatrix} -\mu_{h1}\boldsymbol{I}_n & \boldsymbol{J}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h \\ * & -\boldsymbol{Y}_h \end{bmatrix} < 0 \tag{48}$$

$$\begin{bmatrix} -\boldsymbol{Y}_h & \boldsymbol{Z}_h^T\boldsymbol{L}_h^T \\ * & -\mu_{h2}\boldsymbol{I}_m \end{bmatrix} < 0 \tag{49}$$

where $\boldsymbol{J}_h$, $\boldsymbol{L}_h$ are defined in (30), (31), respectively.

Thus, $\boldsymbol{K}_h^{\circ}$ can be computed using (36) and the control law gain matrix $\boldsymbol{K}_h$ is given as in (32).

Proof. Considering (45) it is possible to write

$$\|\boldsymbol{u}(i)\|^2 = \|\boldsymbol{K}_h\boldsymbol{q}(i)\|^2 = \|(\boldsymbol{J}_h + \boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h)\boldsymbol{q}(t)\|^2 \quad (50)$$

where $\|\cdot\|$ denotes any vector norm. Then using Frobenius norm (50) implies

$$\boldsymbol{q}^T(t)(\boldsymbol{J}_h + \boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h)^T(\boldsymbol{J}_h + \boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h)\boldsymbol{q}(t) = m_h \quad (51)$$

$$\begin{aligned}\boldsymbol{q}^T(t)(\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{J}_h + \boldsymbol{J}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{P}_h^{\frac{1}{2}})\boldsymbol{q}(t)+\\ +\boldsymbol{q}^T(t)(\boldsymbol{J}_h^T\boldsymbol{J}_h + \boldsymbol{P}_h\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h)\boldsymbol{q}(t) = m_h\end{aligned} \quad (52)$$

respectively, where $m_h = \|\boldsymbol{u}(i)\|^2$. Using property

$$\boldsymbol{A}\boldsymbol{B}^T + \boldsymbol{B}\boldsymbol{A}^T \leq \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{B}\boldsymbol{B}^T \quad (53)$$

to the elements of (52) in the first brackets there exists an upper-bound

$$\begin{aligned}\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{J}_h + \boldsymbol{J}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{P}_h^{\frac{1}{2}} \leq\\ \leq \boldsymbol{P}_h + \boldsymbol{J}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{J}_h\end{aligned} \quad (54)$$

and (52) can be rewritten as

$$\begin{aligned}\boldsymbol{q}^T(t)\boldsymbol{J}_h^T(\boldsymbol{I}_m + \boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h\boldsymbol{Z}_h^T\boldsymbol{L}_h^T)\boldsymbol{J}_h)\boldsymbol{q}(t)+\\ +\boldsymbol{q}^T(t)\boldsymbol{P}_h\boldsymbol{q}(t) + \boldsymbol{q}^T(t)\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h)\boldsymbol{P}_h\boldsymbol{q}(t) \geq m_h\end{aligned} \quad (55)$$

It can be possible to consider

$$\boldsymbol{J}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h\boldsymbol{P}_h\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{J}_h < \mu_{h1}\boldsymbol{I}_n \quad (56)$$

and with $\boldsymbol{Y}_h = \boldsymbol{P}_h^{-1}$ (56) implies (46)

With respect to the last element of (55) it can be considered

$$\boldsymbol{P}_h^{\frac{1}{2}}\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h)\boldsymbol{P}_h^{\frac{1}{2}} < \mu_{h2}\boldsymbol{I}_n \quad (57)$$

Thus, with $\boldsymbol{Y}_h = \boldsymbol{P}_h^{-1}$ it yields

$$\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h < \mu_{h2}\boldsymbol{Y}_h \quad (58)$$

$$-\boldsymbol{Y}_h + \mu_{h2}^{-1}\boldsymbol{Z}_h^T\boldsymbol{L}_h^T\boldsymbol{L}_h\boldsymbol{Z}_h < 0 \quad (59)$$

respectively. Then, using Schur complement property, (59) implies (49).

Therefore, inserting (56), (57) into (55) yields

$$\boldsymbol{q}^T(t)(\boldsymbol{J}_h^T\boldsymbol{J}_h + \mu_{h1}\boldsymbol{I}_n + (1+\mu_{h2})\boldsymbol{P}_h)\boldsymbol{q}(t) \geq m_h \quad (60)$$

This concludes the proof. □

8. ILLUSTRATIVE EXAMPLE

To demonstrate properties of the proposed approach, the system with two-inputs and two-outputs is used in the example. The parameters of this system were

$$\boldsymbol{F} = \begin{bmatrix} 0.9993 & 0.0987 & 0.0042\\ -0.0212 & 0.9612 & 0.0775\\ -0.3875 & -0.7187 & 0.5737\end{bmatrix}$$

$$\boldsymbol{G} = \begin{bmatrix}0.0051 & 0.0050\\ 0.1029 & 0.0987\\ 0.0387 & -0.0388\end{bmatrix},\ \boldsymbol{C} = \boldsymbol{I}_3$$

respectively. Considering the third sensor fault described by the equality constrain

$$d_3^T = [0\ 0\ 1]$$

the feedback gain matrix parameters were obtained as follows

$$\boldsymbol{J}_3 = \begin{bmatrix}-4.9935 & -9.2616 & 7.3930\\ 5.0064 & 9.2855 & -7.4121\end{bmatrix}$$

$$\boldsymbol{L}_3 = \begin{bmatrix}0.5013 & 0.5000\\ 0.5000 & 0.4987\end{bmatrix}$$

Solving (34), (35) for LMI matrix variables $\boldsymbol{Y}_3$ and $\boldsymbol{Z}_3$ using Self–Dual–Minimization (SeDuMi) package for Matlab (Peaucelle et al. (2002)), the feedback gain matrix design problem in the reconfigurable control was solved as feasible with the matrices

$$\boldsymbol{Y}_3 = \begin{bmatrix}0.7497 & -0.2504 & -0.0015\\ -0.2504 & 0.7351 & -0.0172\\ -0.0015 & -0.0172 & 0.8356\end{bmatrix}$$

$$\boldsymbol{Z}_3 = \begin{bmatrix}0.1802 & 0.8298 & 0.0382\\ 0.1797 & 0.8276 & 0.0381\end{bmatrix}$$

Inserting $\boldsymbol{Y}_3$ and $\boldsymbol{Z}_3$ into (36) there was computed the additive feedback gain matrix

$$\boldsymbol{K}_3^\circ = \begin{bmatrix}0.6974 & 1.3681 & 0.0752\\ 0.6956 & 1.3646 & 0.0750\end{bmatrix}$$

$$\boldsymbol{K}_3 = \begin{bmatrix}-4.2961 & -7.8935 & 7.4683\\ 5.7021 & 10.6501 & -7.3371\end{bmatrix}$$

The closed-loop system matrix $\boldsymbol{F}_{c3}$ was a stable matrix having structure and eigenvalue spectrum as follows

$$\boldsymbol{F}_{c3} = \begin{bmatrix}0.9927 & 0.0857 & 0.0028\\ -0.1419 & 0.7223 & 0.0332\\ 0.0000 & 0.0000 & 0.0000\end{bmatrix}$$

$$\Omega(\boldsymbol{F}_{c3}) = [0.9357\ 0.7793\ 0.0000]$$

Solving (46)–(48) with respect to LMI variables $\boldsymbol{Y}_3$, $\boldsymbol{Z}_3$, μ_{31}, μ_{32} gives

$$\mu_{31} = 0.8572, \quad \mu_{32} = 1.0347$$

$$\boldsymbol{Y}_3 = \begin{bmatrix}0.8694 & -0.2370 & -0.0035\\ -0.2370 & 0.7695 & -0.0236\\ -0.0035 & -0.0236 & 0.8767\end{bmatrix}$$

$$\boldsymbol{Z}_3 = \begin{bmatrix}0.1550 & 0.3028 & 0.0073\\ 0.1546 & 0.3021 & 0.0073\end{bmatrix}$$

which results in

$$\boldsymbol{K}_3^\circ = \begin{bmatrix}0.3120 & 0.4903 & 0.0228\\ 0.3112 & 0.4891 & 0.0227\end{bmatrix}$$

$$\boldsymbol{K}_3 = \begin{bmatrix}-4.6815 & -8.7712 & 7.4158\\ 5.3176 & 9.7746 & -7.3894\end{bmatrix}$$

Thus, the final results were

$$\boldsymbol{F}_{c3} = \begin{bmatrix}0.9966 & 0.0946 & 0.0033\\ -0.0643 & 0.8990 & 0.0437\\ 0.0000 & 0.0000 & 0.0000\end{bmatrix}$$

$$\rho(\boldsymbol{F}_{c3}) = [0.9478 \pm 0.0608\,\mathrm{i}\ 0.0000]$$

In order to asses the performance of the proposed design method here are now presented simulation results to demonstrate the effect of reconfigurable control law to the control system responses with respect to the

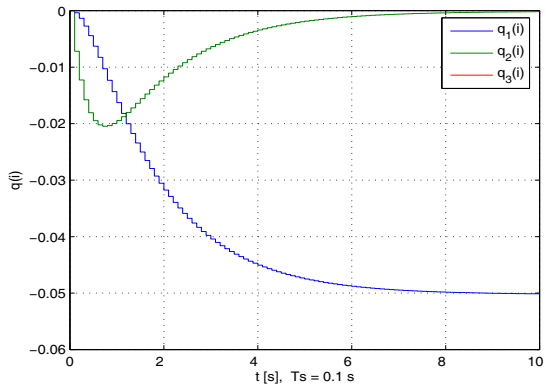


Fig. 1. State variable response of the closed-loop system (Controller design based on Lyapunov inequality)

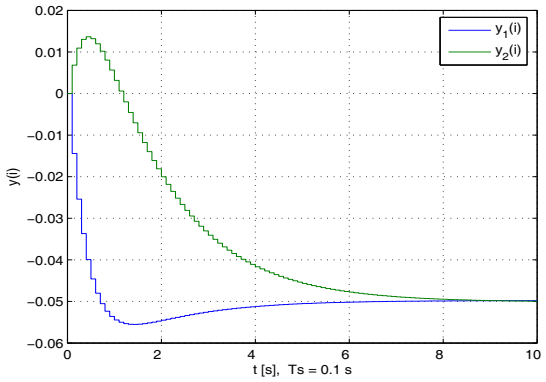


Fig. 2. Objective output response of the closed-loop system (Controller design based on Lyapunov inequality)

third faulty sensor, starting from this faulty initial condition. Simulated states and objective outputs shown in Fig. 1, Fig. 2 demonstrates the effect using a controller designed by Lyapunov inequality (35), and in Fig. 3, Fig. 4 using a controller designed by (47)–(48). All simulations reflects the control policy (26) with $\boldsymbol{N}_w$ computed in such way that the static decoupling was obtained (Wang (2003)), and with $\boldsymbol{w}^T = = [-0.05\ -0.05]^T$. As seen in the figures the short as well as the long run behavior of the control system is quite acceptable.

9. CONCLUDING REMARKS

In this paper the constructive design method based on the classical memory-less feedback control for the stabilization of discrete-time systems with one faulty sensor is presented. The required state feedback gain can be obtained by solving a linear matrix inequality (LMI) feasibility problem. This ensures that the closed-loop system control law gain matrix is optimized while the optimal solution is founded since LMI still form the basis of these algorithms.

Computational methods for determining feedback gains based on the equality constrains as mentioned above are discussed in Ko and Bitmead (2007). Adaptations to reconfigurable control design in Krokavec and Filasová (2008, 2009) are derived using linear quadratic control (LQR) principle, which appears to be applicable if the criterion cost matrices are changed in dependency on considered single sensor fault. In contrast, the above presented design principle is based on asymptotical stability of the closed-loop systems. Using Lyapunov inequality based design principle no free matrix design parameter can be used to tune the system dynamic properties. To overlap this design principle with limits in the gain norm is introduced to obtain a positive impact in reducing the control action activity.

Of course, there may exist some sensors in any system, which staying faulty the system cannot be stabilized using presented methods owing to numerical instability problem.

The validity of the proposed methods is demonstrated by a numerical example with an equality constraint tying a faulty sensor. In the example it was shown that a sensor fault model (in an output error sense) can be used if a equality constraint is set on a faulty sensor output. This suggest that a more theoretical and complete investigation of this principle may be worthwhile. We intend to continue studying the applications of similar techniques to understanding issues in reconfigurable control system design. In particular, we hope that analogous techniques to those developed in this paper will be useful for studying other classes of faults.

REFERENCES

H. Benítéz - Pérez and F. García - Nocetti. *Reconfigurable Distributed Control*. Springer–Verlag, London, 2005.

M. Blanke, M. Kinnaert, J. Lunze and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*. Springer, Berlin, 2003.

D. Boyd, L. El Ghaoui, E. Peron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM Society for Industrial and Applied Mathematics, Philadelphia, 1994.

S.X. Ding. *Model-based Fault Diagnosis Techniques. Design Schemes, Algorithms, and Tools*. Springer–Verlag, Berlin, 2008.

G.J.J. Ducard. *Fault-tolerant Flight Control and Guidance Systems. Practical Methods for Small Unmanned Aerial Vehicles*. Springer–Verlag, London, 2009.

A. Filasová and D. Krokavec. Observer state feedback control of discrete-time systems with state equality constraints. *Archives of Control Sciences*, 20(3):253-266, 2010.

J. Jiang. Fault-tolerant Control Systems. An Introductory Overview. *Acta Automatica Sinica*, 31(1):161-174, 2005.

S. Ko and R.R. Bitmead. Optimal control of linear systems with state equality constraints. In *Proceedings of the 16th IFAC World Congress 2005*, Prag, [CD-ROM], 2005.

S. Ko and R.R. Bitmead. Optimal control for linear systems with state equality constraints. *Automatica*, 43: 1573-1582, 2007.

D. Krokavec and A. Filasová. *Dynamic System Fault Diagnosis*. Elfa, Košice, Slovakia, 2007. (in Slovak)

D. Krokavec and A. Filasová. Performance of reconfiguration structures based on the constrained control. In *Proceedings of the 17th IFAC World Congress 2008*, Seoul, 1243–1248, 2008.

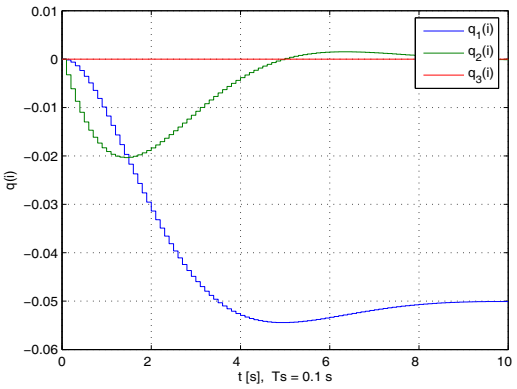


Fig. 3. State variable response of the closed-loop system (Controller design based on the set of inequalities)

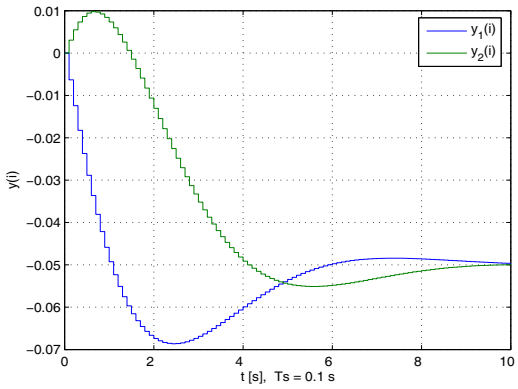


Fig. 4. Objective output response of the closed-loop system (Controller design based on the set of inequalities))

D. Krokavec and A. Filasová. Control reconfiguration based on the constrained LQ control algorithms. In *Preprints of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes Safeprocess 2009*, Barcelona, 686–691, 2009.

Y. Nesterov and A. Nemirovsky. *Interior Point Polynomial Methods in Convex Programming. Theory and Applications.* SIAM Society for Industrial and Applied Mathematics, Philadelphia, 1994.

H. Noura, D. Theilliol, J.C. Ponsart and A. Chamseddine. *Fault-tolerant Control Systems. Design and Practical Applications.* Springer–Verlag, London, 2009.

R.J. Patton. Fault-tolerant control. The 1997 situation. In *Proceedings of the 3rd IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes Safeprocess97, Vol. 2*, Hull, 1033–1054, 1997.

D. Peaucelle, D. Henrion, Y. Labit, and K. Taitz. *User's Guide for SeDuMi Interface 1.04.* LAAS-CNRS, Toulouse, 2002.

S. Simani, C. Fantuzzi, and R.J. Patton. *Model-Based Fault Diagnosis in Dynamic Systems Using Identification Techniques.* Springer, London, 2003.

R.E. Skelton, T. Ivasaki and K. Grigoriadis. *A Unified Algebraic Approach to Linear Control Design.* Taylor & Francis, London, 1998.

M. Staroswiecki. Fault tolerant control. The pseudo-inverse method revisited. In *Proceedings of the 16th IFAC World Congress 2005*, Prag, [CD-ROM], 2005.

D. Theilliol, C. Join, and Y.M. Zhang. Actuator fault tolerant control design based on a reconfigurable reference input. *International Journal of Applied Mathematics and Computer Science*, 18(4):553-560, 2008.

Q.G. Wang. *Decoupling Control*, Springer, Berlin, 2003.

T.J. Yu, C.F. Lin, and P.C. Müller: Design of LQ regulator for linear systems with algebraic-equation constraints. *Proceedings of the 35th Conference on Decision and Control.* Kobe, 4146-4151, 1996.

Y.M. Zhang and J. Jiang. Fault tollerant control system design with explicit consideration of performance degradation. *IEEE Transactions on Aerospace and Electronic Systems*, 9(3):838–848, 2003.

Y.M. Zhang and J. Jiang. Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control*, 32:229-252, 2008.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Set-point reconfiguration in case of severe actuator fault

B. Boussaid [*] **, C. Aubrun** [*] **and N. Abdelkrim** [**]

[*] *Centre de Recherche en Automatique de Nancy (CRAN), University of Nancy, CNRS, BP 239, 54506 Vandoeuvre Cedex, France (e-mail: {boussaid.boumedyen,christophe.aubrun}@cran.uhp-nancy.fr)*
[**] *Modlisation, Analyse et Commande des Systèmes (MACS), University of Gabès, Omar Ibn Khattab Road, 6029 Gabès, Tunisia (e-mail: naceur.abdelkrim@enig.rnu.tn)*

Abstract: In this paper, we propose a set-point reconfiguration approach based on reference-offset governor device and reconfigurable Linear Quadratic Regulator (LQR). Limited by constraints, the nominal performances of any system may be degraded according to fault occurrence. The idea is to modify the nominal set-point where the constraints are threatened to be violated especially after a severe actuator fault appearance. The effectiveness of the proposed method is illustrated by an aircraft numerical example affected by actuator faults and subject to constraints on the actuator dynamic ranges.

Keywords: Fault tolerant control systems, set-point reconfiguration, reference governor, LQ Controller, actuator fault.

1. INTRODUCTION

In industrial processes, systems to be controlled are becoming more and more complex. One of these complexities is due to the necessity of satisfying input/state constraints which is dictated by physical limitations of the actuators. Meanwhile, some plant variables must be kept within safe limits. In recent years, several feedback control techniques of dynamic systems have been developed which are able to handle input and/or state-related constraints (See Angeli et al. [2001]; Bemporad et al. [1997]; Casavola et al. [2000]; Gilbert et al. [1995]; Gilbert and Tan [1991]). In general, these methods are based on predictive approaches which are used to synthesize Command or Reference Governor. In Kolmanovsky and Sun [2006] a Parameter Governor unit is proposed which enforces pointwise-in-time constraints on the evolutions of relevant system variables. Later, both Reference Governor and Parameter Governor actions are integrated in a single unit as Reference-Offset Governor (ROG) in Casavola et al. [2007], which adds many advantages especially in enlarging the set of feasible evolutions of the system. The function of ROG device is to modify, whenever necessary, the reference and add an offset to the nominal control action in order to enforce pointwise-in-time constraints and to improve the overall system transient performance (See Casavola et al. [2006, 2007]).

Regarding Fault-Tolerant Control (FTC) design, the post-fault system should recover the original performance but sometimes (See Jiang and Zhang [2006]; Zhang and Jiang [2003]), it is considered that the system can operate under degraded performance. Besides that, the degree of the system redundancy and the available actuator capabilities can be significantly reduced according to the magnitude of the fault. Moreover, the FTC may cause damage to the system, and even result in loss of system stability (See Jiang and Zhang [2002]).

In this paper, a new reconfiguration approach based on set-point modification using Reference-Offset Governor device is proposed to act in severe actuator faults. In fact, when the magnitude of fault is important, the probability of actuator saturation and performance degradation is high. This situation requires to switch to degraded mode with degraded performance. In our case, the performance degradation is achieved by changing the nominal set-points according to the importance of the actuator fault magnitude.

The main contribution of this paper is to add a Reference-Offset Governor unit to a classic reconfigurable system in order to improve the system transient performance and to reduce the system performance degradation after severe actuator fault occurrence. The formulation of the problem is presented in section 2. Section 3 is reserved to the design of the FTC system with the ROG unit and the LQ controller. Section 4 is dedicated to illustrate the idea with an example of flight control followed by simulation results. Finally, the paper is ended by a conclusion.

2. PROBLEM FORMULATION

Let us consider the following Linear Time-Invariant (LTI) system in discrete time

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) + G_d d(t) \\ y(t) = Cx(t) \end{cases} \quad (1)$$

Where $x(k) \in \mathscr{R}^n$ is the state vector, $u(k) \in \mathscr{R}^m$ is the input vector, $y(k) \in \mathscr{R}^p$ is the output vector, $d(t) \in \mathscr{R}^{n_d}$ is an exogenous bounded disturbance and (A,B,C,G_d) represents the system dynamics.

Let rank $(C)=p$ and rank $(B)=m \geq p$. Assume that the full-state x is available. By solving the Linear Quadratic Regulation problem (See Staroswiecki [2003]; Harkegard and Glad [2005]), the optimal control law is given by :

$$u(t) = -Kx(t) + K_r r(t) \tag{2}$$

with

$$K = R^{-1}B^T P \tag{3}$$

$$K_r = R^{-\frac{1}{2}}(C(BK-A)^{-1}BR^{-\frac{1}{2}})^+ \tag{4}$$

where Q is a positive semi-definite matrix and R is a positive definite matrix. Q and R are preselected by the designer to achieve the nominal performance. P is a unique positive semi-definite and symmetric solution of the Algebraic Riccati Equation (ARE)

$$A^T P + PA + Q - PBR^{-1}B^T P = 0 \tag{5}$$

Let us consider the global system including the ROG unit and the feedback controller, as depicted in Fig. 1.

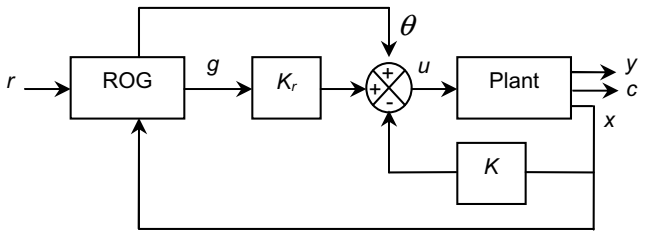


Fig. 1. Global system diagram including Controller and ROG blocs

According to Fig. 1, the control input can be written as :

$$\begin{aligned} u(t) &= -Kx(t) + K_r g(t) + \theta(t) \\ &= -Kx(t) + K_z z(t) \end{aligned} \tag{6}$$

where $K_z = [K_r \quad I_m]$ and $z(t) = [g(t) \quad \theta(t)]^T$.
One replaces (6) in (1), one gets

$$\begin{aligned} x(t+1) &= (A-BK)x(t) + BK_z z(t) + G_d d(t) \\ &= \Phi x(t) + Gz(t) + G_d d(t) \end{aligned} \tag{7}$$

Where $\Phi = (A-BK)$ and $G = BK_z$.
Besides, if one considers only the control input constraints, and one puts $H_c = -K$ and $L = K_z$

$$c(t) = H_c x(t) + Lz(t) + L_d d(t) \tag{8}$$

So, the LTI system in (1) becomes

$$\begin{cases} x(t+1) = \Phi x(t) + Gz(t) + G_d d(t) \\ y(t) = H_y x(t) \\ c(t) = H_c x(t) + Lz(t) + L_d d(t) \end{cases} \tag{9}$$

with $x(t) \in \mathscr{R}^n$ the state vector which includes the controller states; $g(t) \in \mathscr{R}^p$ the manipulable reference which would essentially coincide with the reference $r(t) \in \mathscr{R}^p$; $\theta(t) \in \mathscr{R}^m$ an adjustable offset on the nominal control law which is assumed to be selected from a given convex and compact set Θ, with $0_m \in int\Theta$; $d(t) \in \mathscr{R}^{n_d}$ an exogenous bounded disturbance satisfying $d(t) \in \mathscr{D}; \forall t \in \mathscr{Z}_+$ with $\mathscr{D}$ a specified convex and compact set such that $0_{n_d} \in \mathscr{D}$; $y(t) \in \mathscr{R}^p$ the output, viz. a performance related signal; $c(t) \in \mathscr{R}^{n_c}$ the constraints vector, $c(t) \in \mathscr{C}; \forall t \in \mathscr{Z}_+$; with $\mathscr{C} \subset \mathscr{R}^{n_c}$ a prescribed constrained set. It is assumed that:

A.1) Φ is a stable matrix;

A.2) System (9) is offset-free w.r.t. $g(t)$ i.e.
$H_y(I_n - \Phi)^{-1}G_g = I_p$

where $z(t) = [g(t) \quad \theta(t)]^T \in \mathscr{R}^{p+m}$, is the ROG output and the following matrices are defined $G = [G_g \quad G_\theta]$, $L = [L_g \quad L_\theta]$.

The ROG design problem consists of generating, at each time t; the command input $z(t)$ as an algebraic function of the current state $x(t)$ and reference $r(t)$

$$z(t) := \bar{z}(x(t), r(t)) \tag{10}$$

The ROG output is based on the minimization of a cost function subject to prescribed constraints. The cost function has the following form

$$J(x(t), z(t), r) = \|g(t) - r\|^2_{\Psi_g} + \|\theta(t)\|^2_{\Psi_\theta} \tag{11}$$

where $\Psi_g = \Psi_g^T > 0_m$, $\Psi_\theta = \Psi_\theta^T > 0_m$ and $\|v\|^2_\Psi := v^T \Psi v$. Thus, at each time $t \in \mathscr{Z}_+$, the ROG output is chosen according to the solution of the following constrained optimization problem Casavola et al. [2007]

$$z(t) := arg \min_{z \in \mathscr{V}(x(t))} J(x(t), z(t), r) \tag{12}$$

3. FTC SYSTEM DESIGN

Let us consider the following proposed scheme in FTC solution (See Fig. 2).

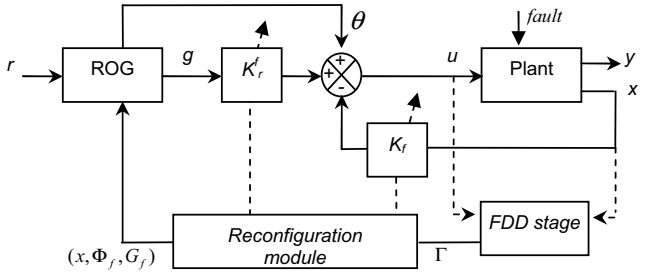


Fig. 2. Proposed scheme of FTC system design

3.1 Controller Reconfiguration

Assume that, after fault occurrence, the new dynamic system behavior of the LTI model is :

$$\begin{cases} x(t+1) = A_f x(t) + B_f u(t) \\ y(t) = Cx(t) \end{cases} \tag{13}$$

where the matrices A_f and B_f are detectable and the system post-fault is controllable. Assuming that (A_f,B_f) is still stabilizable and from the Bellman's optimality principle, the optimal reconfigurable strategy (Staroswiecki [2003]) consists of applying a new optimal control to system (13)

$$u_f(t) = -K_f x(t) + K_r^f r(t) \tag{14}$$

with

$$K_f = R^{-1}B_f^T P_f \tag{15}$$

$$K_r^f = R^{-\frac{1}{2}}(C(B_f K_f - A_f)^{-1}B_f R^{-\frac{1}{2}})^+ \tag{16}$$

where P_f is a unique positive semi-definite and symmetric solution of the Algebraic Riccati Equation (ARE)

$$A_f^T P_f + P_f A_f + Q - P_f B_f R^{-1} B_f^T P_f = 0 \tag{17}$$

3.2 ROG design in faulty case

We propose here an extension of the ROG principle (See Casavola et al. [2007]) in the faulty case. First of all, it is assumed that the matrices A_f and B_f are detectable and the system after fault is controllable. Besides, the controller is reconfigurable and the new feedback and feed-forward gains are denoted K_f and K_r^f. Also, it is considered hereafter the actuator faults only which implies that

$$A_f = A \tag{18}$$
$$B_f = B(I_m - \Gamma) \tag{19}$$

where Γ is the fault distribution matrix;

$$\Gamma = diag(\gamma_i) \tag{20}$$

with $\gamma_i \in [0;1]$ for $i \in \{1,...,m\}$ and m denotes the number of actuators.

In closed-loop scheme, the system state representation is given by :

$$\begin{aligned} x(t+1) &= (A - B_f K_f)x(t) + B_f K_z^f z(t) + G_d d(t) \\ &= (A - BK_f + B\Gamma K_f)x(t) + (BK_z^f - B\Gamma K_z^f)z(t) + G_d d(t) \\ &= (\Phi + B\Gamma K_f)x(t) + (G - B\Gamma K_z^f)z(t) + G_d d(t) \\ &= \Phi_f x(t) + G_f z(t) + G_d d(t) \end{aligned} \tag{21}$$

where Φ_f and G_f represented the global system dynamics after fault occurrence :

$$\Phi_f = \Phi + B\Gamma K_f \tag{22}$$
$$G_f = G - B\Gamma K_z^f \tag{23}$$

Thus, the state description of the plant becomes :

$$\begin{cases} x(t+1) = \Phi_f x(t) + G_f z(t) + G_d d(t) \\ y(t) = H_y x(t) \\ c(t) = H_c^f x(t) + L_f z(t) + L_d d(t) \end{cases} \tag{24}$$

Considering the ROG unit in the faulty case as shown in Fig. 2, and we assume that :

B.1) Φ_f is a stable matrix;
B.2) system (24) is offset-free w.r.t. $g(t)$ i.e.
$H_y(I_n - \Phi_f)^{-1} G_g^f = I_p$

The solution of the cost function (11) is :

$$z(t) := arg \min_{z \in \mathscr{V}_f(x(t))} J(x(t), z(t), r) \tag{25}$$

with $\mathscr{V}_f(x(t))$ is the set of the disturbance-free virtual evolution of the constraints vector $\bar{c}_f(k, x(t), z)$ after fault occurrence

$$\mathscr{V}_f(x(t)) = \{z \in \mathscr{W}_\delta^f : \bar{c}_f(k, x(t), z) \in \mathscr{C}_k^f, \forall k \in \mathscr{Z}_+\} \tag{26}$$

where $\bar{c}_f(k, x(t), z)$ is given by :

$$\bar{c}_f(k, x(t), z) = H_c^f \left(\Phi_f^k x(t) + \sum_{i=0}^{k-1} \Phi_f^{k-i-1} G_f z \right) + L_f z \tag{27}$$

and the two sets; $\mathscr{W}_\delta^f$ and $\mathscr{C}_f^\delta$; are given by :

$$\mathscr{W}_\delta^f := \{z \in \mathscr{R}^{p+m} : \bar{c}_z \in \mathscr{C}_f^\delta\} \tag{28}$$
$$\mathscr{C}_f^\delta := \mathscr{C}_\infty^f \sim \mathscr{B}_\delta \tag{29}$$

Note that $\mathscr{C}_\infty^f$ is constructed from recursive sets $\mathscr{C}_k^f$

$$\mathscr{C}_\infty^f := \bigcap_{k=0}^{k_0^f} \mathscr{C}_k^f \tag{30}$$

where the sets $\mathscr{C}_k^f$ are defined from $k \in \{0, 1, \ldots, k_0^f\}$ as

$$\begin{aligned} \mathscr{C}_0^f &:= \mathscr{C}^f \sim L_d \mathscr{D} \\ &\vdots \\ \mathscr{C}_k^f &:= \mathscr{C}_{k-1}^f \sim H_c^f \Phi_f^{k-1} G_d \mathscr{D} \end{aligned} \tag{31}$$

with $\mathscr{C}^f$ is the prescribed constrained set after fault occurrence. The following properties hold true for the above described ROG in faulty case.

Theorem 1. Let assumptions (B.1) be fulfilled. Consider system ((24)) along with the ROG selection rule (25), and let $\mathscr{V}_f(x(0))$ be non-empty. Then:
1. The minimizer in (25) uniquely exists at each $t \in \mathscr{Z}_+$ and can be obtained by solving a convex constrained optimization problem, viz. $\mathscr{V}_f(x(0)) = \mathscr{V}(x(t_f))$ non-empty implies $\mathscr{V}_f(x(t))$ non-empty along the trajectories generated by the ROG command (24). Such the time of fault occurrence t_f is detrmined by the FDD stage.
2. The set $\mathscr{V}_f(x(t)), \forall x(t) \in \mathscr{R}^n$, is finitely determined, viz. there exists an integer k_0^f such that if $\bar{c}_f(k, x(t), z) \in \mathscr{C}_k^f, k \in \{0, 1, \ldots, k_0^f\}$, then $\bar{c}_f(k, x(t), z) \in \mathscr{C}_k^f \forall k \in \mathscr{Z}_+$. Such a constraint horizon k_0^f can be determined off-line.
3. The constraints are fulfilled for all $t \in \mathscr{Z}_+$.
4. The overall system is asymptotically stable; in particular, whenever $r(t) \equiv r, \lim_{t \to \infty} \theta(t) = 0_m$, and $g(t)$ converges either to r or to its best steady-state admissible approximation $\hat{r}$, with

$$\hat{z}(t) := [\hat{r} \quad 0_m]^T := arg \min_{z \in \mathscr{V}(x(t))} J(x(t), z(t), r) \tag{32}$$

Consequently, by the offset-free condition (B.2), $\lim_{t \to \infty} \bar{y}(t) = \hat{r}$, where $\bar{y}$ is the disturbance-free component of y.

Proof. The proof is similar to that presented in Casavola et al. [2007].

4. NUMERICAL EXAMPLE

The example of study concerns a flight control example for delta-canard configuration describing small single engine fighter, presented in (Harkegard and Glad [2005]). In this example, the controlled variables are the angle of attack, α, the sideslip angle, β and the roll rate, p. The system state includes; besides the manipulated variables; the pitch rate, q, and the yaw rate, r. The control surface vector contains the position of the canard wings, δ_c, the right and left elevons, δ_{re} and δ_{le}, and the rudder, δ_r.
For this considered flight case, we consider a low speed, Mach 0.22, and altitude 3000m. Besides, we suppose that the actuator dynamics are neglected and the actuator position constraints are

$$\delta_{min} = (-55^\circ \quad -30^\circ \quad -30^\circ \quad -30^\circ)^T$$
$$\delta_{max} = (25^\circ \quad 30^\circ \quad 30^\circ \quad 30^\circ)^T$$

Consider the following vectors for output, state and control variables

$$y = (\alpha \quad \beta \quad p)^T$$
$$x = (\alpha \quad \beta \quad p \quad q \quad r)^T$$
$$\delta = (\delta_c \quad \delta_{re} \quad \delta_{le} \quad \delta_r)^T$$

The nominal set-points of regulated outputs are; 30° for the angle of attack, α, and 10° for the sideslip angle, β, and $70^\circ/s$ for the roll rate, p.

For sampling time of 0.5s, the linearized discrete-time model is

$$\begin{cases} x(t+1) = Ax(t) + B\delta(t) \\ y(t) = Cx(t) \end{cases} \tag{33}$$

where the numerical values of the system matrices are

$$A = \begin{pmatrix} 1.0214 & 0.0054 & 0.0003 & 0.4176 & -0.0013 \\ 0 & 0.6307 & 0.0821 & 0 & -0.3792 \\ 0 & -3.4485 & 0.3979 & 0 & 1.1569 \\ 1.1199 & 0.0024 & 0.0001 & 1.0374 & -0.0003 \\ 0 & 0.3802 & -0.0156 & 0 & 0.8062 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.1823 & -0.1798 & -0.1795 & 0.0008 \\ 0 & -0.0639 & 0.0639 & 0.1397 \\ 0 & -1.5840 & 1.5840 & 0.2936 \\ 0.8075 & -0.6456 & -0.6456 & 0.0013 \\ 0 & -0.1005 & 0.1005 & -0.4114 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

For the LQ controller, we fix Q and R to

$$Q = diag(1,2,1,1,3)$$
$$R = diag(5,20,12,2)$$

we obtain the following nominal controller feedback and feed-forward gains:

$$K = \begin{pmatrix} 1.0610 & 0.0059 & 0.0003 & 0.6264 & -0.0012 \\ -0.9294 & 0.2306 & -0.2042 & -0.5348 & -0.1089 \\ -0.9275 & -0.2409 & 0.2037 & -0.5338 & 0.1110 \\ 0.0040 & 0.4245 & 0.1228 & 0.0021 & -1.1279 \end{pmatrix}$$

$$K_r = \begin{pmatrix} 0.8393 & -0.0049 & 0.0001 \\ -0.7818 & -0.7399 & -0.3450 \\ -0.7793 & 0.7481 & 0.3448 \\ 0.0054 & 1.9978 & -0.2161 \end{pmatrix}$$

For the global system representation as in equation (9), the numerical values are

$$\Phi = \begin{pmatrix} 0.4944 & 0.0022 & -0.0000 & 0.1115 & 0.0001 \\ -0.0007 & 0.6015 & 0.0388 & -0.0004 & -0.2357 \\ -0.0041 & -2.8261 & -0.2844 & -0.0022 & 1.1397 \\ -0.9357 & -0.0096 & -0.0007 & -0.1583 & 0.0035 \\ 0.0014 & 0.6022 & -0.0060 & 0.0008 & 0.3201 \end{pmatrix}$$

and $G = [G_g \quad G_\theta]$ with

$$G_g = \begin{pmatrix} 0.4335 & -0.0006 & -0.0000 \\ 0.0009 & 0.3742 & 0.0139 \\ 0.0055 & 2.9437 & 1.0293 \\ 1.6856 & -0.0066 & -0.0001 \\ -0.0020 & -0.6723 & 0.1582 \end{pmatrix}$$

$$G_\theta = \begin{pmatrix} 0.1823 & -0.1798 & -0.1795 & 0.0008 \\ 0 & -0.0639 & 0.0639 & 0.1397 \\ 0 & -1.5840 & 1.5840 & 0.2936 \\ 0.8075 & -0.6456 & -0.6456 & 0.0013 \\ 0 & -0.1005 & 0.1005 & -0.4114 \end{pmatrix}$$

The following sub-figures in figure Fig. 3 show the set-points tracking (See Fig. 3 (a)-(b)-(c)) and the corresponding controls signals (See Fig. 3 (d)-(e)-(f)-(g)) in fault-free case. In Fig. 3 (f), the control input is near the upper limit but does not reached it. As shown on sub-figures (d)-(e)-(g), the control signals are far from upper and lower control limits which means that all contraints are respected for all actuators. Thus, the reference offset-governor unit does no action which implies that the continuous line and dotted line in Fig. 3 (a)-(b)-(c) are superimposed and there is no modification of nominal references.

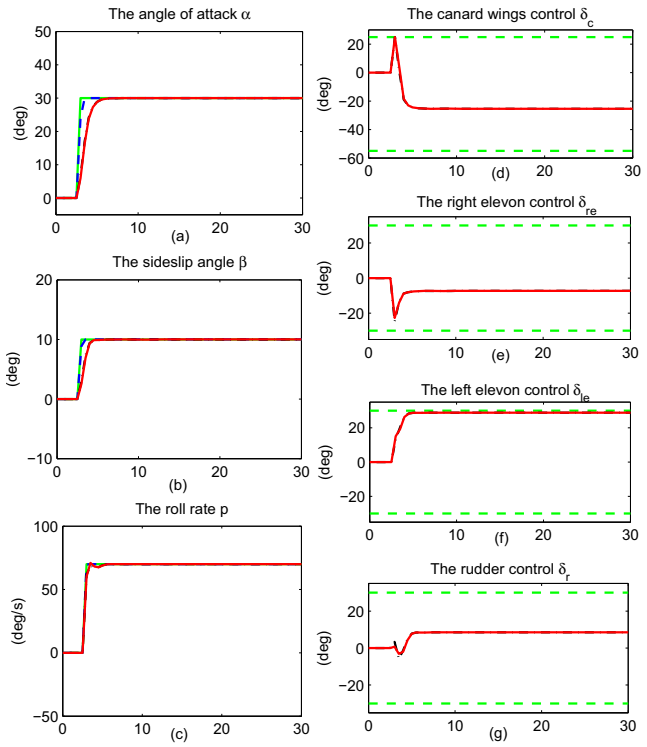


Fig. 3. Fault-free system (a)-(b)-(c) System responses: dotted line - without ROG bloc, continuous line - with ROG bloc, and References : continuous line - nominal, dashed line - modified, (d)-(e)-(f)-(g) Control signals : continuous line - with ROG bloc, dotted line - without ROG bloc, dashed line - upper and lower control limits.

5. SIMULATION RESULTS AND INTERPRETATIONS

To illustrate our approach, a lost of 100% of actuator effectiveness is considered. The occurrence of the fault is at time 10s

and concerns either the actuator 2 or the actuator 3. Note that in both cases, only one actuator is broken.

- *case 1 : 100% lost of actuator 2 effectiveness*

In this case, the new feedback and feed-forward gains of the controller after reconfiguration are

$$K_f = \begin{pmatrix} 1.7344 & -0.3156 & 0.0339 & 1.1212 & 0.1815 \\ 0 & 0 & 0 & 0 & 0 \\ -0.5373 & -0.1796 & 0.0920 & -0.3455 & 0.0648 \\ -0.2883 & 1.4736 & 0.1858 & -0.1734 & -2.0529 \end{pmatrix}$$

$$K_r^f = \begin{pmatrix} 0.9554 & 1.4063 & 0.2814 \\ 0 & 0 & 0 \\ -0.7327 & 1.7443 & 0.3325 \\ -0.3864 & 3.1164 & -0.3635 \end{pmatrix}$$

The results of simulation are shown in Fig. 4.

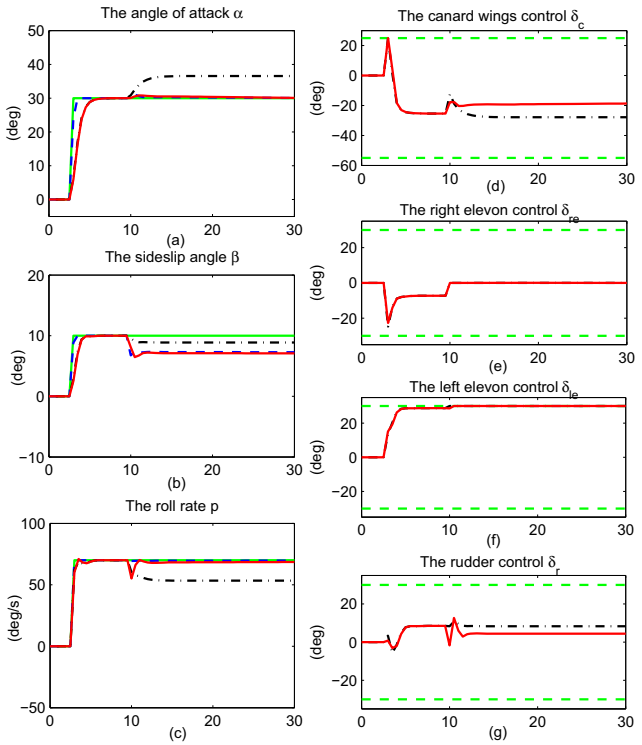


Fig. 4. Faulty-system in case 1: (a)-(b)-(c) System responses: dotted line- without ROG bloc, continuous line- with ROG bloc, and References : continuous line- nominal, dashed line- modified, (d)-(e)-(f)-(g) Control signals : continuous line- with ROG bloc, dotted line- without ROG bloc, dashed line- upper and lower control limits.

In this figure (See Fig. 4), the continuous and dotted lines in (a)-(b)-(c) are not superimposed after time 10s; time of fault occurrence. This observation means that one or more actuators are saturated. This point is confirmed by graphes at Fig. 4 (f). The left elevon control, δ_{re}, reaches the upper limit few seconds after the fault occurrence which induces the saturation of the actuator 3 and the performance degradations in three outputs. However, the solution with ROG unit improves enormously the different outputs by small set-point changing in the sideslip angle, *beta* (See continuous line in Fig. 4 (b)). Note that the system track perfectly the new generated references.

- *case 2 : 100% lost of actuator 3 effectiveness*

In this case of study, the reconfiguration module generate the new following controller feedback and feed-forward gains :

$$K_f = \begin{pmatrix} 1.8540 & 0.2370 & -0.0227 & 1.1950 & -0.1272 \\ -0.3631 & 0.1360 & -0.0594 & -0.2325 & -0.0552 \\ 0 & 0 & 0 & 0 & 0 \\ 0.2216 & 1.5692 & 0.1857 & 0.1336 & -2.1151 \end{pmatrix}$$

$$K_r^f = \begin{pmatrix} 1.1167 & -1.5459 & -0.2592 \\ -0.4946 & -1.7990 & -0.3000 \\ 0 & 0 & 0 \\ 0.2971 & 3.2205 & -0.3767 \end{pmatrix}$$

The results of simulation are shown in Fig. 5.

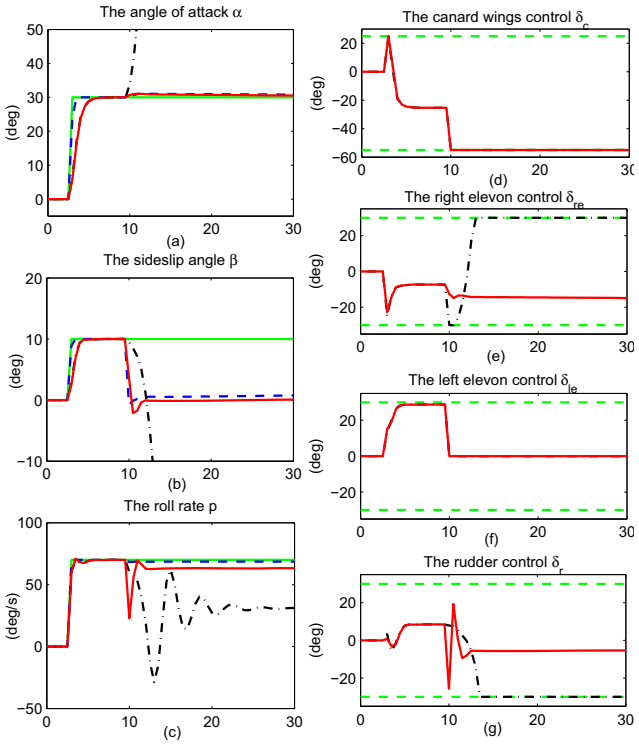


Fig. 5. Faulty-system in case 2: (a)-(b)-(c) System responses: dotted line- without ROG bloc, continuous line- with ROG bloc, and References : continuous line- nominal, dashed line- modified, (d)-(e)-(f)-(g) Control signals : continuous line- with ROG bloc, dotted line- without ROG bloc, dashed line- upper and lower control limits.

In this case of fault (See Fig. 5), the degradation of the system performance in the standard scheme; with reconfigurable bloc only; is quite obvious (See dotted line in Fig. 5 (a)-(b)-(c)). Nevertheless, our proposed FTC scheme with ROG bloc reduces the performance degradation of the post-fault system and ensures the system stability (See continuous lines in Fig. 5 (a)-(b)-(c)). Besides, the modified references (See dashed lines in Fig. 5 (b)-(c)) make the constraints on control not violated (See dotted line in Fig. 5 (d)-(e)-(f)-(g)). Only the canard wings control, δ_c reach th lower limit which explains the important degradation in sideslip angle, *beta* (See dashed and continuous lines in Fig. 5 (b)). Finally, we can see the total lost of effectiveness of actuator 3 after the occurrence of fault at time 10s (See continuous line in Fig. 5 (f)).

6. CONCLUSION

This paper presented a new approach for fault accommodation by modifying the nominal set-points in order to reduce the system performance degradation and ensure the system stability, in the case of severe actuator fault occurrence. This approach is based on reconfigurable controller and Reference-Offset Governor unit which basic function is to avoid the constraints violation. The simulation results of a numerical example proof that this technics might improve the system performance and ensure a safe plant functioning.

REFERENCES

Angeli, D., Casavola, A., and Mosca, E. (2001). On feasible set-membership state estimators in constrained command governor control. *Automatica*, 37, 151–156.

Bemporad, A., Casavola, A., and Mosca, E. (1997). Nonlinear control of constrained linear systems via predictive reference management. *IEEE Transactions on Automatic Control*, 42(3), 340–349.

Casavola, A., Franze, G., and Sorbara, M. (2007). Reference-offset governor approach for the supervision of constrained networked dynamical systems. In *in Proc. of the European Control Conference*, pp. 7–14.

Casavola, A., Mosca, E., and Angeli, D. (2000). Robust command governors for constrained linear systems, robust command governors for constrained linear systems. *IEEE Transactions on Automatic Control*, vol. 45, pp. 2071–2077.

Casavola, A., Papini, M., and Franz, G. (2006). Supervision of networked dynamical systems under coordination constraints. *IEEE Transaction on Automatic Control*, Vol. 51, No. 3, 421–437.

Gilbert, E., Kolmanovsky, I., and Tan, K. (1995). Discrete-time reference governors and the nonlinear control of systems with state and control constraints. *Int. J. on Robust and Nonlinear Control*, vol. 5, pp. 487–504.

Gilbert, E. and Tan, K. (1991). Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Transactions on Automatic Control*, 36(9), 1008–1020.

Harkegard, O. and Glad, S. (2005). Resolving actuator redundancyoptimal control vs. control allocation. *Automatica*, vol. 41, pp. 137 144.

Jiang, J. and Zhang, Y. (2002). Graceful performance degradation in active fault tolerant control systems. In *In Proc. of the 15th IFAC World Congress b'02, Barcelona, Spain*.

Jiang, J. and Zhang, Y. (2006). Accepting performance degradation in fault-tolerant control system design. *IEEE Transactions on Control Systems Technology*, vol. 14, 284–292.

Kolmanovsky, I. and Sun, J. (2006). Parameter governors for discrete-time nonlinear systems with pointwise-in-time state and control constraints. *Automatica*, 42, 841–848.

Staroswiecki, M. (2003). Actuator faults and the linear quadratic control problem. In *in Proc. of the 42th IEEE Conference on Decision and Control, Maui, Hawaii, USA*, pp. 959–965. Maui, Hawaii USA.

Zhang, Y. and Jiang, J. (2003). Fault tolerant control system design with explicit consideration of performance degradation. *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 39(NO. 3), pp. 838–848.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Connections of Functional States for Automaton Identification: Application in a Steam Generator Monitoring

J. F. Botia[1], H. Sarmiento[1,2], C. Isaza[1]

[1]Electronic Engineering Department – GEPAR Research Group, University of Antioquia, Medellin, Colombia, (e-mail: {javier.botia, phosm984}@jaibana.udea.edu.co and cisaza@udea.edu.co).
[2]Polytechnic Jaime Isaza Cadavid, Medellín, Colombia.

Abstract: An automaton fixes the connections among functional states in a process. If the functional states are obtained by clustering, non all classes or functional states are defined by the human operator, consequently, a method to suggest the connections among classes is necessary. The proposed method allows estimating automatically the automaton from the membership degrees obtained by a fuzzy clustering method. The method was used to find a Fuzzy Automaton of a steam generator process. In order to verify the independence of the clustering method, three different clustering techniques were used. The connection method proposed gives the same results for these cases.

Keywords: System Monitoring, Fuzzy Clustering, Fuzzy Automata, Complex Systems, Functional States.

1. INTRODUCTION

In the industrial environment, the human operator has the knowledge of the process. This expert can perform multivariable analysis to identify states and to track the ongoing process of the system, which he anticipates a failure state (Isermann, 2006). The supervision systems and fault diagnosis, based on detection of the *functional states* allow estimating the current behavior of a process. It constitutes an aid to the human operator (Lamrini, et al., 2005). In a complex process, the operator knows the system and some functional states, but he may not know all connections among functional states (Aguilar, 2007). Moreover, when applying clustering techniques new classes may be created which are ignored by the expert and are useful for diagnosis.

To assist to the operator in the task of supervision, fuzzy clustering methods have allowed finding classes by means of the historical data. In order to establish an automaton, these classes are associated to functional states and are interconnected. For these connections, Waissman, et al. (2005) proposed the *deterministic fuzzy automata*. This approach uses the classification information according to the number of samples belonging to a class, and the result is validated by the human operator. The graph of connections is constructed based on the sequences of changes among functional states by means of the classification with new data. Kempowsky, et al. (2006) defined transitions and frequencies matrices according to the changes among classes obtained by the classification of historical data. In these methods there is a fully dependence of the connections on the historical data. Kempowsky's approach only allows finding connections among classes with transitions in the historical data. Subsequently, there may appear connections that the expert considers necessary, but that the complex system does not detect due to the lack of examples of transition among functional states in the historical.

The *Fuzzy Automata* have been used to establish connections defined by a weight of the transitions by means of a set of initial, internal and final states (Klir and Yuan, 1995). Two types of automata exist, d*eterministic fuzzy automata* which allow obtaining a connection from an initial state to a final state only, and *non- deterministic fuzzy automata* that allow establishing connections from an initial state to all final states (Omlin, et al., 1998). The former was used to construct connections among states known by the expert (Waissman, et al., 2000). This approach used the historical data classification in to the functional states established by the expert. Consequently, this automaton does not help to create connections not considered in the historical data.

In this paper is proposed a method to construct a transition matrix that includes connections among functional states, and that does not exclusively depend on the transitions in the historical data. The method only depends on the membership degrees matrix; consequently, the method can be applicable to any fuzzy clustering technique. In our method, the transition matrix is constructed and updated by means of the Hebbian learning. The method is applied to find the automaton of a subsystem of boiler of a steam generator. This system was designed as a version to pilot scale of a real steam generator for a nuclear central (characteristics of the process are described in section 4). The same case was used to validate Kempowsky's approach (Kempowsky, 2004).

In section 2, general theory of the Fuzzy States Machines and the Hebbian Learning used in non-supervised learning in neural networks are presented. In section 3, the proposed method for obtaining and updating membership degrees transition matrix is presented. The implementation of the proposed method to a steam generator system, and a comparison of the results obtained with three Fuzzy Clustering methods are presented in section 4. Finally, in section 5 the conclusions and perspectives are presented.

2. FUZZY STATES MACHINE AND HEBBIAN LEARNING

The *Fuzzy States Machine* (FμMS) consists on a set of states and their respective transitions associated to membership degrees. The concept of FμMS was proposed for the first time by Wee and Fu (1969), as a fuzzy automaton that represents the graph of connections in a transitions matrix. For definition, a fuzzy automaton with finite states is the six-tuple $M = \{\Sigma, Q, R, Z, \delta, \omega\}$, where Σ is the set of alphabet, Q is the set of states, R is the initial state fuzzy automaton ($R \in Q$), Z is the finite output alphabet, δ: Σ x Q x [0,1] → Q is the fuzzy transitions map and ω: $Q \rightarrow Z$ is the output transitions map. If ω is ignored and a five-tuple $L(M) = \{\Sigma, Q, R, Z, \delta\}$ is considered, where L is the language that the automaton interprets, each transition between two states is going to be defined by a weight $\theta_{a,b,p} \rightarrow [0,1]$, where a and b are the indices of the states and p is the index of alphabet that correspond to the transition (Omlin, et al., 1999). Therefore, each transition of states, $q_a \rightarrow q_b$, with a input of symbol r_p and weight $\theta_{a,b,p}$ is going to be defined as $r_p/\theta_{a,b,p}$.

The methodology for automatic connections among functional states will use the subsets Σ, Q y δ because the graph of connections is static and constant in time if the classification is carried out with the historical data. The transition matrix, Δ, is going to represent all connections among states, as shown in (1):

$$\Delta = \begin{pmatrix} \theta_{1,1,1} & \theta_{1,2,1} & \cdots & \theta_{1,b,1} \\ \theta_{2,1,2} & \theta_{2,2,2} & \cdots & \theta_{2,b,2} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{a,1,p} & \theta_{a,2,p} & \cdots & \theta_{a,b,p} \end{pmatrix} \quad (1)$$

In section 3 a new representation of the matrix Δ will be shown. The difference in the new representation does not consider the input symbols p. The *Hebbian learning* is a technique that has allowed carrying out the updating of weight of an Artificial Neural Network. By definition, it is an updating function of a weight $w_{n,m}$ (Bishop, 2006), expressed in (2).

$$w_{n,m}(\eta+1) = w_{n,m}(\eta) + \varepsilon(x_n - y_m) \quad (2)$$

Where n is the number of the current iteration, x_n is the n-th input to the network, y_m is the m-th network output and ε is the learning parameter.

This proposal uses the theory of the FμMS to establish the initial frequencies of connections with all classes obtained by a fuzzy clustering method. In the case of automatic connections, if the connection has a likely transition among states, the Hebbian learning can use it to update the connections and thus remove those that are weaker. In addition, as the update of the matrix is non-supervised, *self-organizing maps algorithm* is used to propose an automaton. The expert analyses the new connections that are suggested and to decide the relationship they have with the behavior of the process (the algorithm is explained in section 3.2). In the next section, the implementation of fuzzy states machines and Hebbian learning are explained to construct and update transition matrix.

3. FROM FUZZY PARTION TO AUTOMATON

The method calculates a membership degrees transition matrix based on the membership degrees. The membership degrees are the result of the clustering of historical process data. When using the theory of Fuzzy States Machine, the method establishes all connections among functional states to update the matrix by using the Hebbian learning and self−organizing maps algorithm. Finally, the graph of connections among functional states is constructed by means of transition matrix; it represents the behaviour of the process. Fig. 1 presents the diagram of the method at each step which will be explained in detail in sections 3.1, 3.2 and 3.3.

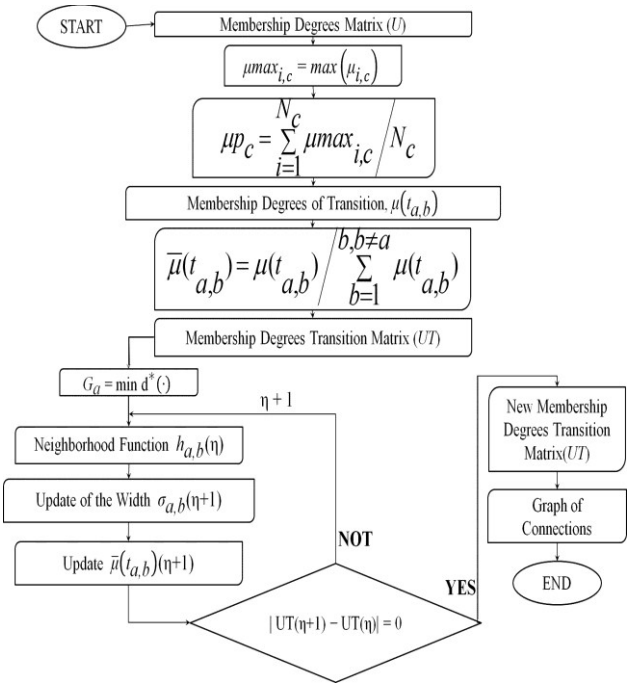


Fig 1. Scheme of the connection method.

3.1 Obtaining of the Membership Degrees Transition Matrix

Let U be a membership degrees matrix,

$$U = \begin{pmatrix} \mu_{1,1} & \mu_{1,2} & \cdots & \mu_{1,j} \\ \mu_{2,1} & \mu_{2,2} & \cdots & \mu_{2,j} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{i,1} & \mu_{i,2} & \cdots & \mu_{i,j} \end{pmatrix} \quad (3)$$

for i samples and j classes. The $\mu_{i,j}$ is the membership degree of the sample i to the class j. Using (3), the maximum membership degree is determined for each sample i associated to a class c, $\mu max_{i,c}$, where $c \rightarrow [1, \ldots, j]$, as in (4).

$$\mu max_{i,c} = max(\mu_{i,c}) \; for \; \forall c, 1 \le c \le j \quad (4)$$

μp_c is defined as the *average of the maximum membership degrees* of the samples i that belongs to a class c, as in (5).

$$\mu p_c = \sum_{i=1}^{N_c} \mu max_{i,c} \Big/ N_c \quad (5)$$

Where N_c is the number of samples with $\mu max_{i,c}$ that belongs to a class c. Using (5), a matrix UP is created. This matrix associates each membership degree μp_c with its correspondent class C_c, expressed in (6).

$$UP = \begin{pmatrix} \mu p_1 & C_1 \\ \mu p_2 & C_2 \\ \mu p_c & C_c \\ \vdots & \vdots \\ \mu p_j & C_j \end{pmatrix} \quad (6)$$

Using (6), the new transition matrix is constructed by applying the theory of Fuzzy States Machines (see section 2). In this case, a *Non-Deterministic Fuzzy States Machine* is proposed. There are connections from an initial *fuzzy state activity* S_a to multiple fuzzy states S_b, been $b \neq a$ (Reyneri, 1997). The interest of the method is focused on knowing the *membership degrees of transition* $\mu(t_{a,b})$ ranging from the states $a \rightarrow b$ and $b \rightarrow a$. The transitions matrix obtained assigns a frequency to each one of the possible connections among functional states. Beginning with all likely connections, the expert can avoid ignoring useful transitions which are not considered in the historical data. For a sample assigned to the state S_b with a membership degree $\mu'(S_b)$, it is possible defined $\mu'(S_b)$ as the contribution of $\mu(t_{a,b})$. $\mu(t_{a,b})$ defines the connection degree of the states $a \rightarrow b$. The relation between the final state S_b and the state S_a (with a membership degree $\mu(S_a)$), is defined in (7) (Alvim and Cruz, 2008).

$$\mu'(S_b) = \mu(t_{a,b})\mu(S_a) \quad (7)$$

Where for $\forall\, a$, $1 \leq a \leq j$ and $\forall\, b$, $1 \leq b \leq j$. The average of membership degrees of fuzzy state activities is known, which is associated with $\mu(S_a)$ and $\mu'(S_b)$; $\mu(t_{a,b})$ is estimated by using (7). When $\mu(t_{a,b}) > 1$, it is suggested to carry out a normalization, $\bar{\mu}(t_{a,b})$, as is proposed in (8).

$$\bar{\mu}(t_{a,b}) = \mu(t_{a,b}) \Big/ \sum_{1}^{b,b\neq a} \mu(t_{a,b}) \quad (8)$$

It is important consider that sub-index a as the initial state and the sub-index b as a final state. Transitions within the same active and final fuzzy state (that is $a = b$) have a membership degree of 1, but for reason of the construction of the graph of connections they are not taken into account. Hence, in the method connections of one state to another are estimated but not the possibility remaining in one same state. The sum of $\mu(t_{a,b})$ over all b, as in (8), satisfy the equation (9) (Demasi, 2003).

$$\sum_{1}^{b,b\neq a} \bar{\mu}(t_{a,b}) = 1 \quad (9)$$

Thus, the ratio between the sum of the membership degrees of the final states $\mu'(S_b)$, if $b \neq a$, and the sum of the $\bar{\mu}(t_{a,b})$ determine the membership degree of the state $\mu(S_a)$. Moreover, by using (9) is deduced that:

$$\mu(S_a) = \sum_{1}^{b} \mu'(S_b) \Big/ \sum_{1}^{a,a\neq b} \bar{\mu}(t_{a,b}) = \sum_{1}^{b} \mu'(S_b) \quad (10)$$

Equation (10) allows to permanently conserve all time the activation of the functional states, because adding the membership degrees of all the states $\mu'(S_b)$ will allow to conserve the membership degree of the initial state $\mu(S_a)$, if the sum of all transitions is equal to 1, as in (9). With this criterion, fuzzy states machine will preserve the belonging of the functional states all time (Demasi, 2003). Using (8), the *membership degrees transition matrix UT* is estimated as is shown in (11):

$$UT = \begin{pmatrix} 1 & \bar{\mu}(t_{1,2}) & \cdots & \bar{\mu}(t_{1,b}) & \cdots & \bar{\mu}(t_{1,j}) \\ \bar{\mu}(t_{2,1}) & 1 & \cdots & \bar{\mu}(t_{2,b}) & \cdots & \bar{\mu}(t_{2,j}) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \bar{\mu}(t_{a,1}) & \bar{\mu}(t_{a,2}) & \cdots & \bar{\mu}(t_{a,b}) & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{\mu}(t_{j,1}) & \bar{\mu}(t_{j,2}) & \cdots & \cdots & \cdots & 1 \end{pmatrix} \quad (11)$$

3.2 Updating of the Membership Degrees Transition Matrix

In an Automaton for monitoring, not all connections among functional states describe the real behavior of the process. For this reason, an update of the matrix UT (11) is required. By using the Hebbian learning theory and the self–organizing maps algorithm, the transition matrix is updated. The objective is to remove less representative connections. Let G_a be a function of minimum distance to a class (Palma and Marin, 2008):

$$G_a = min\, d^*(\cdot) \quad (12)$$

where $d^*(\cdot)$ is an arbitrary distance. In this case the criterion for distance between classes is necessary only based on fuzzy membership degrees, matrix U in (3). Isaza (2007) proposed a new distance between fuzzy classes a and b, $d^*(a,b)$. This measurement depends only on the membership degrees of the classes, so data is not required. Then, the distance between the classes a and b is estimated in (13).

$$d^*(a,b) = 1 - \left(\sum_{i=1}^{NU} \mu_{a,i} \cap \mu_{b,i} \Big/ \sum_{i=1}^{NU} \mu_{a,i} \cup \mu_{b,i} \right) \quad (13)$$

NU is the number of samples of the classes a and b, $\mu_{a,i} \cap \mu_{b,i} = min\left(\mu_{a,i}, \mu_{b,i}\right)$ and $\mu_{a,i} \cup \mu_{b,i} = max\left(\mu_{a,i}, \mu_{b,i}\right)$. Following the theory of the self-organizing maps, a Gaussian function $h_{a,b}$ is used (Kohonen, 1990). If the standard deviation or width, $\sigma_{a,b}(\eta)$ decreases, then $\bar{\mu}(t_{a,b})$ also decreases and by increasing $\sigma_{a,b}(\eta)$, $\bar{\mu}(t_{a,b})$ also increases. This function is shown in (14).

$$h_{a,b}(\eta) = exp\left(-G_a \Big/ 2\left(\sigma_{a,b}(\eta)\right)^2\right) \quad (14)$$

η is the actual iteration number of the algorithm. With the algorithm, $\sigma_{a,b}(\eta)$ is updated by using an initial random width. The updating $\sigma_{a,b}(\eta + 1)$ has the form of a Hebbian function, defined in (15) (Haykin, 2008).

$$\sigma_{a,b}(\eta+1) = \sigma_{a,b}(\eta) + \left[\varepsilon_o(\eta) h_{a,b}(\eta) / \sigma_{a,b}(\eta)\right] \cdot \left[\left(G_a^2 \Big/ m\left(\sigma_{a,b}(\eta)\right)^2\right) - 1\right] \quad (15)$$

Where m represents the size of the matrix UT (j x j dimensions or $m = j^2$) and $\varepsilon_o(\eta)$ is a learning factor for the updating of $\sigma_{a,b}(\eta + 1)$, expressed in (16).

$$\varepsilon_o(\eta) = \varepsilon_{in}\left(\varepsilon_f / \varepsilon_{in}\right)^{1/\eta} \quad (16)$$

Where, ε_{in} is an initial learning factor and ε_f is a final learning factor. To update the matrix *UT* (11), a *function of updating of the membership degrees of transition*, $\bar{\mu}(t_{a,b})$ (η + 1) is proposed as a Hebbian function, expressed in (17).

$$\bar{\mu}(t_{a,b})(\eta+1) = \bar{\mu}(t_{a,b})(\eta) + \left[\varepsilon_t(\eta) h_{a,b}(\eta)/\sigma_{a,b}(\eta)\right] \cdot \left[G_a^2\right] \quad (17)$$

Where $\varepsilon_t(\eta)$ is a learning factor for the updating of $\bar{\mu}(t_{a,b})$ (η+1), defined in (18).

$$\varepsilon_t(\eta) = \varepsilon_{in} + \left(\varepsilon_f - \varepsilon_{in}\right) \cdot \left(\eta/\eta f\right) \quad (18)$$

ηf is the maximum number of iterations. The algorithm updates the matrix *UT* (11) at each iteration and it is finalized when there exist a stability of the matrix, i.e. when $|UT(\eta + 1) - UT(\eta)| = 0$. As result the matrix *UT* (11) is automatically updated to conserve the properties shown in (9) and (10).

3.3 Construction of the Graph of Connections among Functional States

To establish the updating of the matrix *UT* (11), the graph of connections is constructed by observing output transitions that exist from an initial functional state with respect to the other. The automaton includes the transitions with a membership degree transition different to 0. The final graph represents the changes of the behavior of a process that are suggested to the expert.

4. CASE OF STUDY: STEAM GENERATOR PROCESS

4.1 Description of the Process

The proposed method is applied to a boiler subsystem of a steam generator. An estimation of fuzzy automaton is made with historical data and it is validated with 2 sets of new data and the concept of the expert in the process. The same process was used to validate the approach of matrices of connections presented Kempowsky, et al. (2006). In order to compare the results, the proposed method is evaluated in this paper with the same data. The steam generator test is designed as a version to pilot scale of a real steam generator for a nuclear central. Operation of the process is as follows: the feed water flow is generated by a pump that propels water to a boiler. To maintain constant water level in the boiler, a controller On−Off operates via the pump. Therefore, the heat power value of the boiler will depend on the steam accumulator pressure. When the accumulator pressure drops below of a minimum value, the heat resistance is activated which gives the maximum heat power, and when by achieving a maximum pressure the heat resistance is cut off to maintain the pressure in ±0.2 Bar in the set-point. The steam flow generated is measured by a flow sensor. The functional states analyzed by the expert correspond to: normal operation (C1), pressure regulation (C2) and (C3), level regulation (C4), and level and pressure regulation (C5). The historical data of the process contains 937 samples and 5 descriptors or process variables, corresponding to physical variables: the feed flow water (F3), heat power (Q4), boiler pressure (P7), boiler level (L8) and output steam flow (F10). The data are normalized with respect to the maximum and minimum value of each variable with the objective of homogenizing the influence of the innate dimensions of the variables.

4.2 Graph of Connections of Functional States

The first step is to obtain the matrix *U*, as in (3), as result of the application of a Fuzzy Clustering algorithm using the historical data. In order to analyze the independence with respect to the used Fuzzy Clustering method, three methods are used to find the initial transition matrix: *LAMDA* method (Piera, et al., 1989), *FCM* method (Bezdek, 1981) and *GK-Means* method (Gustafson and Kessel, 1978). In table 1, the parameters with which convergence and stability were achieved for the clustering algorithm are presented. The results (classes) of the classification correspond to those associated by the expert and those obtained in Kempowsky et al. (2006) by using *LAMDA* method. Fig. 2 illustrates the classification result, where each sample is classified in one of the 5 classes or functional states. To construct the graph of connections (see Fig. 3a), the membership degrees transition matrix is estimated in each case and it is also updated. The independence of the Clustering method is demonstrated (see Table 2).

Table 1. Parameters used for *LAMDA*, *FCM* and *GK-Means*

Algorithm	**Parameters**				
LAMDA	**Method**	**Exigency**	**Iterations**	**Probability**	**Connectivity**
	Self-Learning	0.8	2	Lamda3	Min-Max
FCM	**Classes**	**Factor of Fuzzification**	**Error**	**Iterations**	**Vol.**
	5	1.1	10^{-10}	1000	1
GK-Means	**Classes**	**Factor of Fuzzification**	**Error**	**Iterations**	**Vol.**
	5	2.1	10^{-10}	1000	1

Table 2. Membership degrees transition matrices.

LAMDA		**(C1)**	**(C2)**	**(C3)**	**(C4)**	**(C5)**
	(C1)	1	0.258	0.221	0.272	0.248
	(C2)	0.252	1	0.239	0.267	0.241
	(C3)	0.215	0.267	1	0.225	0.293
	(C4)	0	0	0	1	1
	(C5)	0	0	0	1	1
FCM		**(C1)**	**(C2)**	**(C3)**	**(C4)**	**(C5)**
	(C1)	1	0.251	0.251	0.249	0.248
	(C2)	0.25	1	0.249	0.250	0.249
	(C3)	0.222	0.274	1	0.235	0.269
	(C4)	0	0	0	1	1
	(C5)	0	0	0	1	1
GK-Means		**(C1)**	**(C2)**	**(C3)**	**(C4)**	**(C5)**
	(C1)	1	0.251	0.251	0.248	0.249
	(C2)	0.253	1	0.251	0.246	0.248
	(C3)	0.244	0.282	1	0.219	0.255
	(C4)	0	0	0	1	1
	(C5)	0	0	0	1	1

The automaton obtained by the membership degrees transition matrix is the same in the three cases. Six connections are removed and the stability of the matrices *UT* is achieved with η = 6 using ε_{in} = 0.01, ε_f = 0.1 and ηf = 50.

The values ε_{in} and ε_f are recommended by Haykin (2008). It is important to note that for the case of *LAMDA*, the matrix U is normalized before applying the proposed method, determining the maximum and minimum $\mu_{i,j}$ of all the elements of matrix U. By approximating ε_{in} to ε_f a low elimination of the connections among functional states is observed (see Fig. 3b), false transitions are presented that are not in agreement with the physical behavior of the process. Fig. 4 shows the graph of connections using the transitions matrices proposed in Kempowsky, et al. (2006). In these results, the method only allows to find transitions in the historical (see Fig. 4, continuous line) and after making validation with the expert and analysis of validation data, connection (C2)–(C5) is manually added (see Fig. 4, dotted line). By comparing the graph of connections of the Fig. 3a and Fig. 4, the proposed method in this paper is able to find and to validate the majority of connections among functional states proposed by Kempowsky (2004), except the connection (C5)–(C2). However, new connections are found that are not considered in the historical data, as:

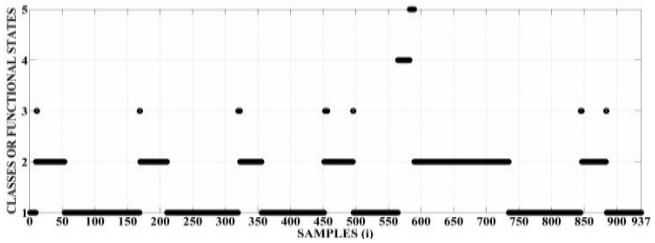


Fig 2. Classification of the historical data with *LAMDA*, *FCM* and *GK–Means*.

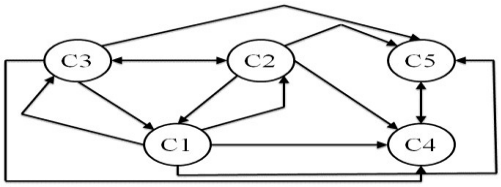


Fig 3a. Graph of connections obtained with *LAMDA*, *FCM* and *GK–Means*.

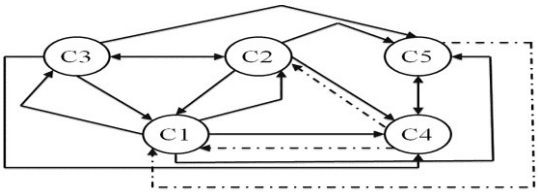


Fig 3b. Graph of connections with false transitions, if ε_{in} is approximated to ε_f.

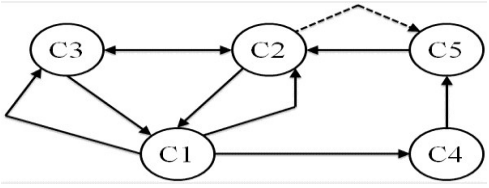


Fig 4. Graph of connections obtained with transition and frequencies matrices (Kempowsky, et al., 2006).

Connections (C2)-(C5) and (C3)-(C5): In the classification with historical data (see Fig. 2), connections (C2)–(C5) and (C3)–(C5) are not established. With the proposed method, connections (C2)–(C5) and (C3)–(C5) they are found, and when validating them with new data, it demonstrates that these transitions exist in the process (see Fig. 5, Fig. 6a and Fig. 6b). Physically, these connections occur because when detecting a low or high pressure in (P7), the pressure (C2) or (C3) are regulated by activating or cutting off the resistance (Q4) to increase/decrease the pressure. Therefore, the water level changes (more or less level) in (L8), (F3) is activated to regulate the level and (F10) is activated in a short instant of time to regulate the output pressure at boiler (it corresponds to (C5)).

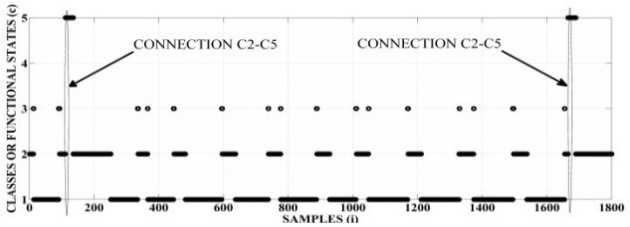


Fig 5. Connection (C2)–(C5) with validate data.

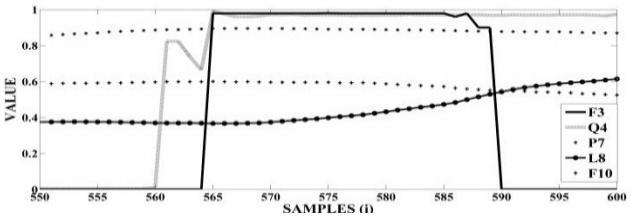


Fig 6a. Range of the validation data to analyze the connection (C3)–(C5).

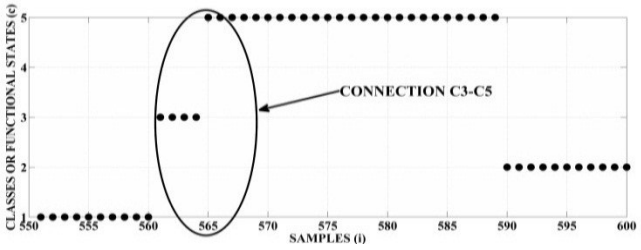


Fig 6b. Connection (C3)–(C5) in the validate data classification.

Connection (C1)-(C5): In normal state (C1), if for an instant of time a critical state of water level and pressure (C5) is detected; (F3) and (F10) are activated in a short time to stabilize the output pressure of the boiler and to regulate the water level.

Connection (C5)-(C4): After that connection (C1)–(C5) occurs, if in (C5) the water level does not achieve to stabilize totally in this short instant of time, when (L8) detects it, then the state (C4) allows to regulate the level by activating or cutting off (F3). Then, to avoid a change in the pressure because the water level changes, a transient occurs in the state regulate level (C4) which it must regulate pressure and level (C5) or connection (C4)–(C5) (See Fig. 3a and Fig. 4).

Connection (C3)-(C4): If there is a high pressure in (P7), (L3) is cut off to establish the pressure (C3), but critical state occurs because the water level may decrease in (L8), therefore (F3) must be activated to increase the water level to a stable level.

Connection (C2)-(C4): this case is a transition that cannot physically occur, because it is impossible that by detecting a low pressure in (P7) and the state (C2) begins to regulate pressure by activating (L3), the state (C4) regulates the water level when this state is activated if the increase of level is necessary. This last connection the expert suggested removing it.

It is important to take into account that the method suggests new connections to the expert, but he is the one who makes the final decision of the most important transitions for the process. Moreover, the expert cannot add connections.

5. CONCLUSIONS

A new methodology was defined in order to find an automaton that includes the connections among functional states. The connections are represented for a membership degrees transition matrix. By updating the transitions, the method determines the most "relevant" connections to predict the changes in the behavior of the process, plus other connections that were not evident in the historical data and were found when updating the matrix. The results suggest that the new connections may happen and should be considered by the expert for the decision-making of the process. The method is useful because clustering methods may detect classes not foreseen by the expert. It is clear that there exist dependence with respect to the learning parameters, but the recommended values of learning parameters allow finding the best graph of connections. As future work method to define the join of classes when two or more functional states represent similar behaviors is being considered, as well as the use of the membership degrees in order to eliminate false connections.

6. ACKNOWLEDGMENT

This paper is the result of project CODI MDC09-13 at the University of Antioquia. Thanks to Group DISCO, LAAS-CNRS for facilitate the data and Tatiana Kempowsky for her orientation in the analyses of the steam generator process.

REFERENCES

Aguilar–Martín, J. (2007). *Inteligencia Artificial Para la Supervisión de Procesos Industriales*. ULA, Mérida, Venezuela.

Alvim, L.G.M. and de Oliviera Cruz, A.J. (2008). A Fuzzy State Machine Applied to an Emotion Model for Electronic Game Characters. *IEEE World Congress on Computational Intelligence*, pp. 1956–1963.

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Edit Plenum Publishing Corporation, New York, USA.

Bishop, C.M. (2006). *Neural Networks For Pattern Recognition*. Second edition, Oxford University Press, New York, USA.

Demasi, P. (2003). *Estratégias adaptativas e evolutivas em tempo real para jogos eletrônicos*. UFRJ, Rio de Janeiro, Brasil.

Gustafson, D.E. and Kessel, W.C. (1978). Fuzzy Clustering with a Fuzzy Covariance Matrix. *Proceedings of IEEE CDC*, volume (17), pp. 761–766.

Haykin, S. (2008). *Neural Networks and Learning Machines*. Third edition, edit Prentice Hall, New Jersey, USA.l

Isaza, C. (2007). *Thèse Diagnostic Par Techniques D'Apprentissage Floues: Conception D'Une Méthodes de Validation et D'Optimisation dans Partitions*. LAAS–CNRS, Toulouse, France.

Isermann, R.(2006). *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Edit Springer-Verlag, Berlin, Germany.

Klir, G.J. and Yuan, Bo. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, Upper Saddle River, New Jersey, USA.

Lamrini, B., Le Lann, M.V., Benhammou A. and Lakhal, E.L. (2005). Détection des états fonctionnels par la méthode de classification «LAMDA». *Comptes Rendus Physique*, volume (6), issue (10), pp. 1161–1168.

Kempowsky, T. (2004). *Thèse Surveillance de Procèdes à Base de Méthodes de Classification: Conception d'un Outil d'Aide Pour la Détection et le Diagnostic des Défaillances*. LAAS – CNRS, Toulouse, France.

Kempowsky, T., Subias, A. and Aguilar – Martin, J. (2006). Process Situation Assessment: From a Fuzzy Partition to a Finite State Machine. *Engineering Applications of Artificial Intelligence*, volume (19), pp. 461–477.

Kohonen, T. (1990). The Self-Organizing Maps. *Proceedings of the IEEE*, volume (78), issue (9), pp. 1464–1480.

Omlin, C.W., Thornber, K.K. and Giles, L.C. (1998). Fuzzy Finite-State Automata Can be Deterministically Encoded into Recurrent Neural Networks. *IEEE Trans. on Fuzzy Systems*, volume (6), issue (1), pp. 86–79.

Omlin, C.W., Thornber, K.K. and Giles, L.C. (1999). Equivalence in Knowledge Representation: Automata, Recurrent Neural Network, and Dynamical Fuzzy Systems. *Proccedings of the IEEE*, volume (87), number (9), pp. 1623–1640.

Palma Méndez, J. and Marín Morales, R. (2008). *Inteligencia Artificial: Técnicas, Métodos y Aplicaciones*. Edit McGraw – Hill, Madrid, España.

Piera, N., Deroches, P. and Aguilar-Martin, J. (1989). *LAMDA: An Incremental Conceptual Clustering Method*. LAAS – CNRS, report (89420), Toulouse, France.

Reyneri, L.M. (1997). An Introduction to Fuzzy State Automata. *IWANN*, volume (1240), pp. 273-283.

Waissman, J., Sarrate, R., Escobet, T., Aguilar-Martin, J. and Dahhou, B. (2000). Wastewater Treatment Process Supervision By Means of a Fuzzy Automaton Model. *Proceedings of the ISIC*, pp. 163–168.

Waissman, J., Ben-Yousself, C. and Vazquez, G. (2005). Fuzzy Automata Identification Based on Knowledge Discovery in Datasets for Supervision of a WWT Process. *SETIT*.

Wee, W.G. and Fu, K.S. (1969). A formulation of Fuzzy Automata and its Applications as a Model of Learning Systems. *IEEE Trans on. Systems Science and Cybernetic*, volume (5), pp. 215–233.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

A GMDH TOOLBOX FOR NEURAL NETWORK-BASED MODELLING

Marcel Luzar * **Marcin Witczak** *

* *Institute of Control and Computation Engineering, University of Zielona Góra ul. Pogórna 50, 65-246 Zielona Góra (e-mail:{m.luzar,m.witczak}@issi.uz.zgora.pl)*

Abstract: In this paper, a MATLAB toolbox, which implements the idea of a group method of data handling is presented. First, a theoretical background regarding such a kind of neural networks is provided. Subsequently, design steps, tests and experimental results are presented. Moreover, the selection methods used in the considered approach are discussed. The final part of the paper exhibits practical applications of the developed toolbox.

Keywords: Neural network, parameter estimation, system identification.

1. THEORETICAL BACKGROUND

GMDH (ang. Group Method of Data Handling) (Farlow (1984)) is a specific type of neural networks. The method was developed by Ivakhenko (Ivakhenko and Mueller (1995)). The method makes it possible to identify multi-inputs systems. Because the accuracy of the neural model largely depends on a proper selection of the neural network structure, which is a complex task, the idea of GMDH is to replace a complex model with structure consisting of some sub-models (Patan et al. (2008)). The structure evolves during the synthesis of the network. Sub-models can be categorized as a linear, nonlinear or dynamic ones (Mrugalski (2003)). The toolbox uses dynamic sub-models. To protect GMDH network from expanding into infinity, the selection methods are needed. The implemented methods are as follows (Koza (1992)):

- constant population method,
- roulette method,
- ranking method,
- tournament method.

Moreover, a stop condition is needed that will prevent an excessive growth of a network (Duch et al. (2000)).In the toolbox, as a stop condition either the minimum estimation error or the maximum number of layers are used, respectively. Finally, the resulting network is shown in Fig. 1. For comprehensive description of the GMDH strategy the reader is referred to (Witczak (2007)), and hence it is omitted in this paper.

2. DESIGN CONDITIONS

Before implementation, the following conditions are taken into account:

- possibility to identify nonlinear and dynamic Multiple Input Multiple Output (MIMO) systems,
- identification for a series-parallel neural model, validation for a parallel neural model.,
- providing design parameters for modeling GMDH network as follows:

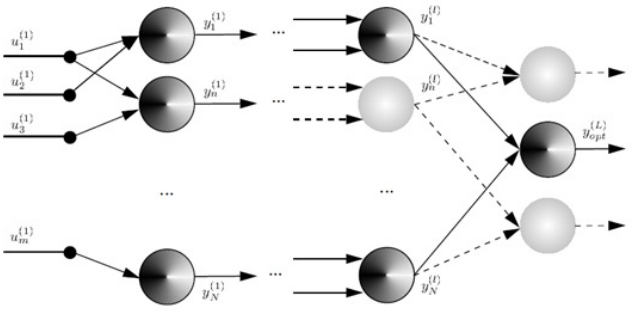


Fig. 1. Final structure of GMDH network

  (1) stop condition (minimum estimation error, maximum number of layers),
  (2) dynamic order,
  (3) selection method,
  (4) selection criterion,
  (5) initial parameter vector,
  (6) input and output lags,
  (7) state vector dimension for neural model.
- possibility to import testing and validation data from a file or a workspace,
- Graphical User Interface (GUI),
- error handling,
- help files.

The sub-model parameter estimation methods are as follows:

- bounded-error estimation (Jaulin and Walter (2001)),
- estimation with Adaptive Random Search (ARS) algorithm (Prudius (2007)),
- sub-space identification (Demuth et al. (2010)).

3. GMDH TOOLBOX

The entire implementation was done using the Matlab environment. A graphical user interface editor, called *guide* is used. The GMDH Toolbox consists of two windows: main window (GMDH Network Toolbox) and the data

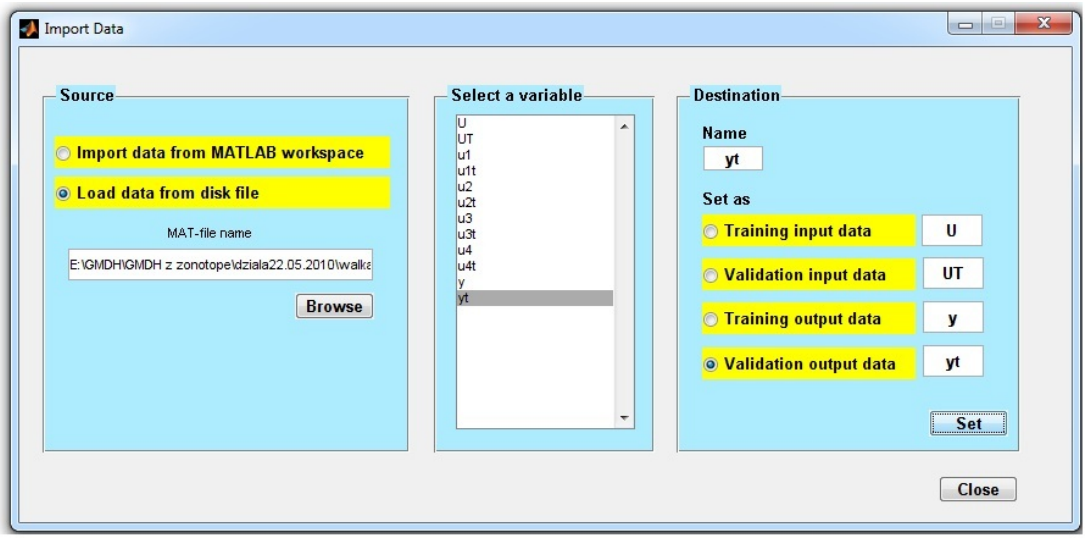


Fig. 2. Import data window

import window (Import Data). The ergonomic arrangement of the buttons and lists was achieved as a result of inspiration with the Neural Network Toolbox, which is available in Matlab. In the import data window (Fig. 2), there are three panels: the data source selection panel, the variable selection panel as well as the training input data, the validation input data, the training output data and, finally, the validation output data panel. The data can be imported either from an existing file with extension *.mat or from the Matlab workspace, respectively. After assigning the variables from the list as the relevant data (by the selection button *Set*), the import data window is closed. The main window (Fig. 3) is equipped with the tools for choosing model selection methods, the parameters of these methods, the set of the initial conditions and the network synthesis stop conditions. In the main window, menu bar is available, consisting of two items: *Menu* and *Help*. After pressing the corresponding button, the assigned submenu is opened. It is possible to open the import data window, to build the GMDH network structure or to close the application. The *Help* button submenu consists of *About toolbox* button (after pressing this button, the window with the information about the author and the version of the toolbox is opened), *Help* button (help files) and *Show GMDH Structure* (which shows an exemplary scheme of GMDH structure). The choice of the estimation method is performed by selecting an appropriate option from the *Estimation methods* panel. It is possible to choose the above mentioned estimation methods, namely: bounded-error estimation (Jaulin and Walter (2001)), ARS (Prudius (2007)) and sub-space identification method (Demuth et al. (2010)), respectively. Depending on the choice, the appropriate design parameters are available. In the case of bounded-error estimation, the user can specify the maximum number of layers, after which the GMDH network synthesis is terminated (the first stop condition). As the second stop condition, the minimum estimation error is defined. Besides the stop conditions, the user is required to choose the orders of the system dynamics. They are defined by setting a maximum lags in the inputs and outputs, respectively. If the non-dynamic models are needed, it is necessary to set zero value

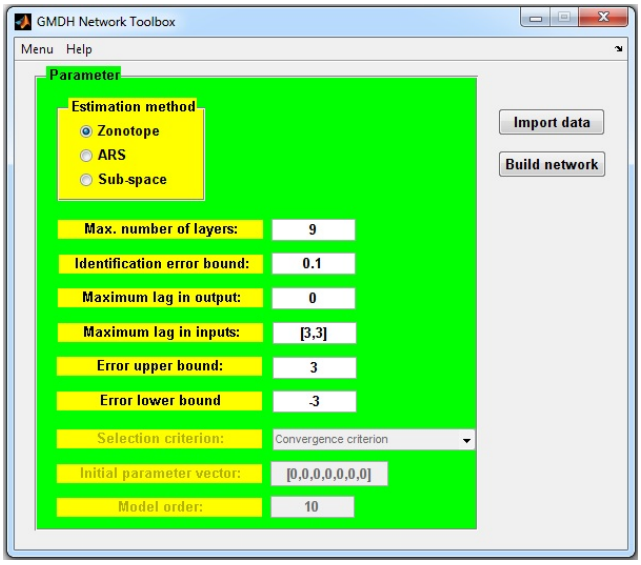


Fig. 3. Main window

in the delay text fields. The subsequent fields, which are available by selecting the bounded-error estimation, are the upper and lower bounds of the error, which may not be exceeded during the system identification phase. Another possible way to estimate the model parameters is the estimation method employing the ARS algorithm. After selecting the method, it is possible to set a GMDH network synthesis stop conditions. Furthermore, in a popup menu the selection criterions are available, namely (Soderstrom and Stoica (1989)):

- convergence criterion,
- absolute convergence criterion,
- mean square criterion,
- absolute mean criterion.

When the ARS method is selected, it is necessary to specify the initial vector of parameters. The last parameter estimation method is a sub-space one. In the case of this method, besides obvious stop conditions, it is necessary to choose the model order, i.e. the dimension of the sub-

model state vector. If the data are imported, the GMDH network synthesis can be initiated by pressing the *Build network* button.

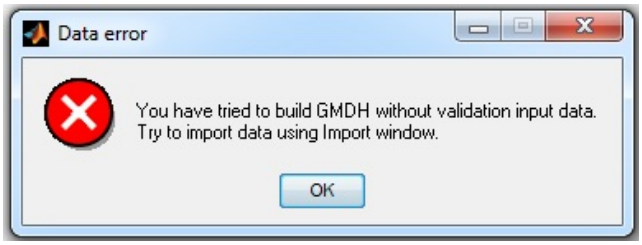


Fig. 4. Warning message window.

In the case, when data are not imported, or if are not correctly assigned as training/validation input/output data, the warning window is shown (Fig. 4). The results and plots made during building the GMDH network will be discussed in Section 4. To understand the toolbox more in depth, the help files are prepared, which are available after pressing the Help button (Fig. 5). In the help file, all necessary information, how to use the toolbox and the background on the GMDH idea are given.

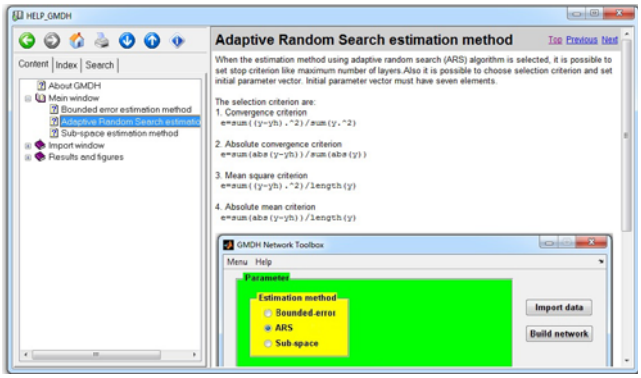


Fig. 5. Help file

Moreover, the Help files provide some hints on how the criterions are interpreted and how to interpret the resulting graphs.

4. EXPERIMENTAL RESULTS

After running the algorithm, which builds the GMDH network, by pressing the Build GMDH network, the synthesis of the GMDH network is started. After this process, the windows with graphs are opened, which show, how precisely the neural network fits the real system. The graphs are different, i.e. they depend on the selected estimation methods. The data used for the experiments were collected at the evaporation station of the Lublin Sugar Factory S.A. (Kościelny et al. (2002)). There was four inputs: juice pressure on control valve input, juice pressure on control valve output, juice flow after control valve and the juice temperature on control valve input. The output was rod displacement of servomotor. The first plot illustrates the experimental results obtained using the bounded-error estimation method. The first chart of the top shows the results of the GMDH neural modeling for the training data. The system output is marked by the solid line and the model output is marked with a dotted line. It is easy to see, that for over 800 samples both lines covers each other which means, that the neural network mimic the system with a good quality. The neural modeling results for the validation data are presents in Fig. 7. A comparison of outputs from the system and the model for the validation data suggests that neural modeling in is correct. Moreover, the identification error for the validation data is acceptable (Aubrun et al. (2008)). The third chart presents the bounded system output. It helps to check, if the interval between upper and lower bound is tight enough to identify with the best accuracy. Comparing the second and third chart, the biggest difference occurs while the bounds are exceeded. After neural modeling process, another figure is generated, which makes it possible to see, in which layer of a neural network the identification error was the smallest one. Figure 6 shows, that the smallest error for the validation data was in the second layer, in the next layers it begins to grow. This means that larger number of layers was not needed for identifying the system. Another experiment was made using the estimation ARS method. In the case of data from the Lublin Sugar Factory S.A., this method gives worse results than previous methods. In Fig. 8, it is easy to see, that the lines representing the output from the system and the model do not overlap as perfectly as in the case of the bounded-error estimation method. But it does not mean that the estimation method using ARS algorithm will be the worst one in every case. Another experiments were made for the data from the tunnel furnace available in the laboratory at University in Zielona Góra prove that the ARS method for this case is more appropriate. The results of this experiment are not included in this paper. The last experiment was concerned with the estimation with sub-space method. Similarly to the ARS method, also this gave worse identification results compared to the bounded-error estimation method (Fig. 9). Like for the ARS method, it does not mean that the sub-space method will give worst results for every data. The optimal number of layers was larger than in the first method. The synthesis was ended on third layer.

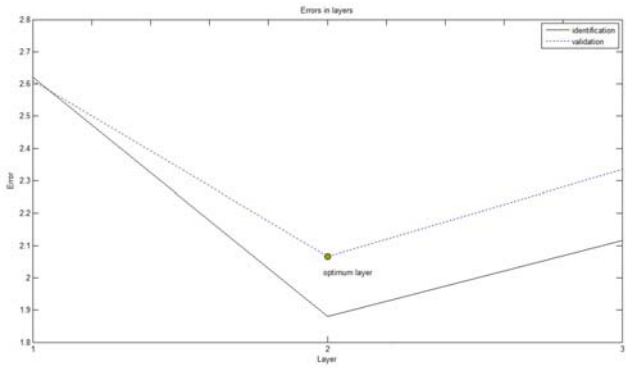


Fig. 6. Optimal number of layers in GMDH structure

5. CONCLUSIONS

In this paper, a GMDH toolbox for neural network-based modelling is presented. It was designed as an alternative approach to the system identification, different from the typical neural networks. The obtained results do not deviate from the widely accepted norms. The implemented three methods for estimating the parameters allow the

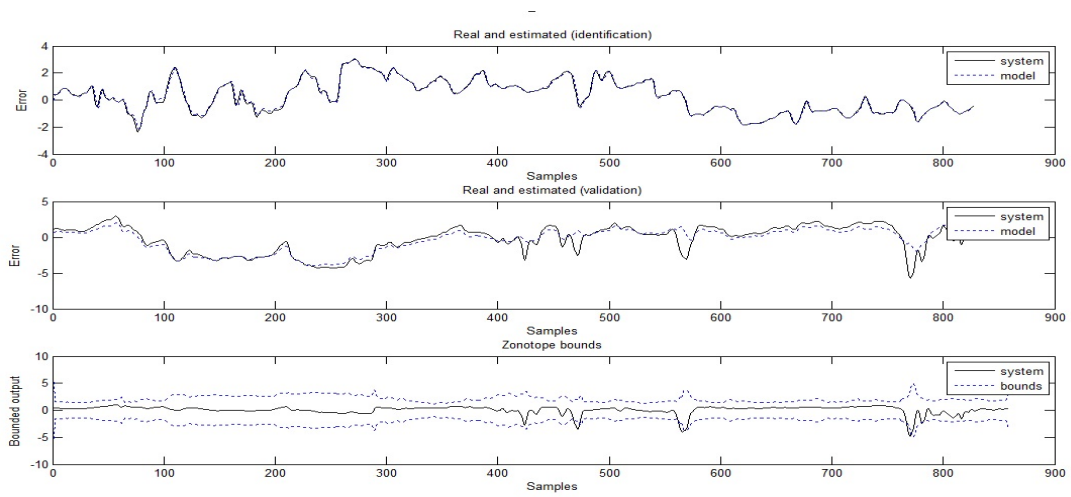


Fig. 7. System and model outputs for training and validation data from Lublin Sugar Factory S.A. obtained using bounded-error method.

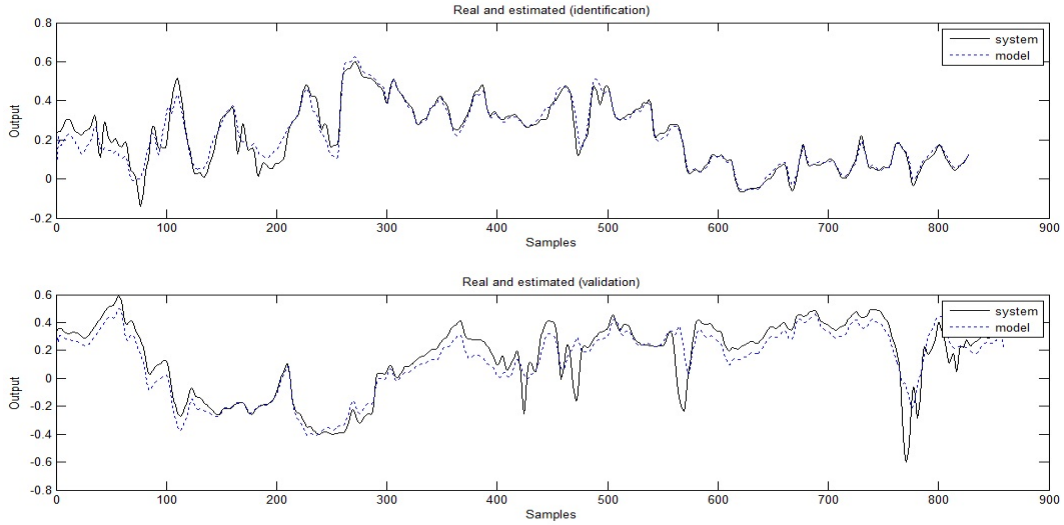


Fig. 8. System and model outputs for training and validation data from Lublin Sugar Factory S.A. obtained using estimation with adaptive random search (ARS) algorithm.

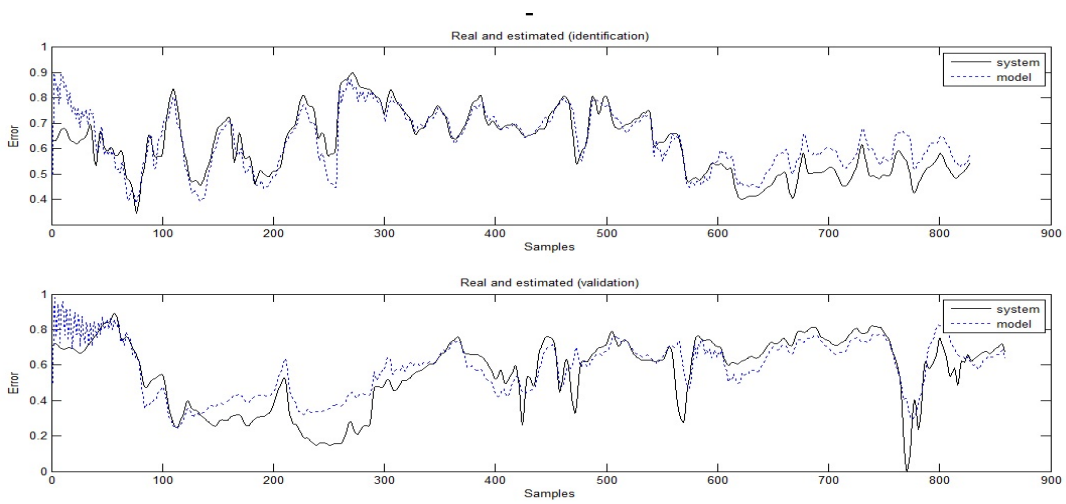


Fig. 9. System and model outputs for training and validation data from Lublin Sugar Factory S.A. obtained using sub-space method.

selection of an appropriate identification scenario. Further work will focus on drawing the GMDH network structure for a given case based on conection matrix. The Toolbox is used as an excellent way to understand the idea of the GMDH. The help file includes full error handling facilitates, which make the work with the toolbox more comfortable.

REFERENCES

Aubrun, C., Sauter, D., and Yame, J. (2008). Fault diagnosis of networked control systems. *International Journal of Applied Mathematics and Computer Science*, 18(4), 525–537.

Demuth, H., Beale, M., and Hagan, M. (2010). Neural network toolbox 6 users guide.

Duch, W., Korbicz, J., Rutkowski, L., and Tadeusiewicz, R. (2000). *Sieci neuronowe*. Biocybernetyka i inzynieria biomedyczna. Akademicka Oficyna Wydawnicza Exit, Warszawa.

Farlow, S.J. (1984). *Self-organizing Methods in Modeling - GMDH Type Algorithms*. Marcel Dekker, New York.

Ivakhenko, A.G. and Mueller, J.A. (1995). Self-organizing of nets of active neurons. *System analysis modeling simulation*, 20, 93–106.

Jaulin, L. and Walter, E. (2001). *Nonlinear bounded-error parameter estimation using interval computation*. Physica-Verlag, Heidelberg.

Kościelny, J.M., Bartyś, M., Syfert, M., and Pawlak, M. (2002). Industrial applications. In J. Korbicz, J. Kościelny, and Z. Kowalczuk (eds.), *Fault Diagnosis. Models, Artificial Intelligence, Applications*, 57–114. Springer, Berlin.

Koza, J.R. (1992). *Genetic Programming: On The Programming of Computers by Means of Natural Selection*. MIT Press, Massachusetts.

Mrugalski, M. (2003). *Neuronowe modelowanie systemów nieliniowych w ukadach detekcji uszkodzeń*. University of Zielona Góra, Zielona Góra.

Patan, K., Witczak, M., and Korbicz, J. (2008). Towards robustness in neural network based fault diagnosis. *International Journal of Applied Mathematics and Computer Science*, 18(4), 443–454.

Prudius, A.A. (2007). *Adaptive Random Search Methods for Simulation Optymalization*. Georgia Institute of Technology, Atlanta.

Soderstrom, T. and Stoica, P. (1989). *System Identification*. Prentice Hall, New York.

Witczak, M. (2007). *Modelling and estimation strategies for fault diagnosis of non-linear systems: from analytical to soft computing approaches*. Springer-Verlag, Berlin.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Decoupling model predictive control in a non-minimal state space representation

U. Hitzemann * **K. J. Burnham** *

* *Control Theory and Applications Centre, Coventry University, Coventry, United Kingdom (e-mail: hitzemau@uni.coventry.ac.uk)*

Abstract: This paper is concerned with a decoupling strategy using model predictive control (MPC) in a non-minimal state space (NMSS) form. For simulation experiments, a 2×2 multiple input, multiple output (MIMO) system is used, which exhibits cross-couplings in the input-output pathways. The decoupling approach follows from earlier research where it is applied to pole-placement controllers. Here, it is applied to NMSS-MPC where its formulation is modified in order to allow incorporation of individual control and prediction horizons of the inputs and outputs, respectively, in a straightforward way. Finally, constraints are imposed and a simulation example is presented.

Keywords: constraints, decoupling, model predictive control (MPC), non-minimal state space (NMSS)

1. INTRODUCTION

In the past decades, extensive research in the field of model predictive control by various researchers has been undertaken (Camacho and Bordons, 2007). The generalised predictive controller (GPC) (Clarke et al., 1987a) and its extensions (Clarke et al., 1987b) can be considered as a 'milestone' in transfer function based model predictive control. Since this earlier pioneering work, systems are also often modelled by making use of state space representations and consequently the GPC approach is modified for the use of state space models, see i.e. (Bitmead et al., 1990; Ikonen and Najim, 2002; Kwon and Han, 2005). Further reviews of the developments in MPC can be found in (Mayne et al., 2000). Here, one of the key features of MPC is mentioned, namely the relatively straightforward incorporation of constraints (Maciejowski, 2001). Since the cost function in MPC is of a quadratic form, in many cases, it turns out to be a quadratic programming (QP) convex optimization problem and can be solved by making use of QP methods, see i.e. (Boyd and Vandenberghe, 2004; Nocedal and Wright, 2006). Furthermore, robustness of MPC has been investigated, i.e. (Kothare et al., 1996) followed by developments targeting computational efficient algorithms by splitting the algorithm into offline and online computed parts (Kouvaritakis et al., 2000).

In this paper, the MPC configured within a non-minimal state space (NMSS-MPC) form is used. The advantage of the NMSS structure over the minimal state space structure is that the necessity of a state observer is eliminated. Since the states are, in many cases, not available as measurements, an estimate is required which is then used in the state feedback control law. In linear state feedback control, the dynamics additionally introduced by the observer are designed to be sufficiently faster than the system dynamics hence, by the separation theorem, may be considered negligible. In the case when constraints become active, non-linearities are introduced and the effect of the observer dynamics can become a crucial issue and can result in constraint violation or even instability (Wang and Young, 2006). The NMSS structure overcomes the necessity of using an observer since the state vector contains the current output and its past values as well as the past inputs. These data are measurable, hence directly available and an estimate of the states is not required (Young et al., 1987; Wang and Young, 1988). In these developments, the NMSS structure incorporating an '*integral-of-error-state*' is used in order to ensure set point tracking. In (Wang and Young, 2006) a NMSS-MPC in difference form is proposed thereby eliminating the '*integral-of-error-state*' and yet able to ensure set-point tracking.

The issue of decoupling has been addressed by (Plummer and Vaughan, 1997; Lee et al., 1995) and it is upon this development that this paper is based. Here, the decoupling is achieved by diagonalizing the system matrices, so that the cross-couplings are nullified. Other approaches are also possible. One such approach makes use of an appropriate choice of the controller weighting matrices, see (Exadaktylos and Taylor, 2010).

The paper is organized as follows: In Section 2, the non-minimal state space structure is introduced as well as the model predictive control structure. Section 3 is concerned with the decoupling MPC and imposing constraints. Finally, a simulation example is presented in Section 4 and conclusion given in Section 5.

2. THE MODEL PREDICTIVE CONTROLLER IN A NON-MINIMAL STATE SPACE FORM

2.1 Non-minimal state space model structure

Consider a mathematical plant model with q inputs and p outputs expressed in the following difference equation form

$$\begin{aligned} \mathbf{y}(k) + F_1\mathbf{y}(k-1) + \ldots + F_{n_a}\mathbf{y}(k-n_a) \\ = H_1\mathbf{u}(k-1) + \ldots + H_{n_b}\mathbf{u}(k-n_b) \end{aligned} \quad (1)$$

where $\mathbf{u} = [u_1 \ldots u_q]^T$ denotes the vector of inputs and $\mathbf{y} = [y_1 \ldots y_p]^T$ denotes the vector of outputs. F_i and H_i denote constant coefficient matrices of appropriate dimension. For simplicity, suppose that the number of inputs is equal to the number of outputs $q = p = n$. The above system representation is now represented in an equivalent, general state space form where subscript g denotes 'general'

$$\mathbf{x}_g(k) = A_g\mathbf{x}_g(k-1) + B_g\mathbf{u}(k-1) \tag{2a}$$
$$\mathbf{y}(k) = C_g\mathbf{x}_g(k) \tag{2b}$$

where $A_g \in \mathbb{R}^{n(n_a+n_b)\times n(n_a+n_b)}$, $B_g \in \mathbb{R}^{n(n_a+n_b)\times n}$ and $C_g \in \mathbb{R}^{n\times n(n_a+n_b)}$ denote the state transition matrix, input transition matrix and output transition matrix, respectively, and $\mathbf{x}_g \in \mathbb{R}^{n(n_a+n_b)\times 1}$ denotes the state vector.

$$A_g = \begin{bmatrix} -F_1 & -F_2 & \cdots & -F_{n_a} & H_2 & H_3 & \cdots & H_{n_b-1} & H_{n_b} \\ I & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & I & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & I & 0 \end{bmatrix} \tag{3}$$

$$B_g = \begin{bmatrix} H_1^T & 0 \cdots I & 0 & 0 \cdots 0 & 0 \end{bmatrix}^T \tag{4}$$

$$C_g = \begin{bmatrix} I & 0 \cdots 0 & 0 & 0 \cdots 0 & 0 \end{bmatrix}^T \tag{5}$$

Delays can be taken into account by setting the corresponding H_i matrices zero. As mentioned above, the state vector in the NMSS representation contains current and previous output values as well as previous input values, hence it is of the following form

$$\begin{aligned}\mathbf{x}_g(k) = [&\mathbf{y}(k)\,\mathbf{y}(k-1)\ldots\mathbf{y}(k-n_a+1)\\ &\mathbf{u}(k-1)\ldots\mathbf{u}(k-n_b+1)]^T\end{aligned} \tag{6}$$

In order to ensure set point tracking in the MPC design, integral action is introduced by multiplying (2) with the difference operator $\Delta = 1 - z^{-1}$ where z^{-1} denotes the backward time shift operator and by augmenting the state vector by $\mathbf{y}(k)$, this is the structure proposed by (Wang and Young, 2006). Performing these operations, (2) becomes

$$\mathbf{x}(k) = G\,\mathbf{x}(k-1) + D\,\Delta\mathbf{u}(k-1) \tag{7a}$$
$$\mathbf{y}(k) = C\,\mathbf{x}(k) \tag{7b}$$

where

$$\mathbf{x} = \begin{bmatrix}\Delta\mathbf{x}_g \\ \mathbf{y}\end{bmatrix} \tag{8}$$

$$G = \begin{bmatrix} A_g & 0 \\ C_gA_g & I\end{bmatrix} \tag{9}$$

$$D = \begin{bmatrix} B_g \\ C_gB_g\end{bmatrix} \tag{10}$$

$$C = [0\ I] \tag{11}$$

Here, I and 0 denote the identity and null matrices of appropriate dimensions, respectively.

2.2 NMSS-MPC

The MPC formulation in the SISO case is well described in the literature, see for example (Ikonen and Najim, 2002; Kwon and Han, 2005; Camacho and Bordons, 2007; Wang, 2009) but the extension to MIMO models is often brief since it only requires a slight modification. In this paper, a further modification to the MPC formulation in the MIMO case is proposed which allows the implementation of individual control and prediction horizons of the n inputs and outputs to be realised in a straightforward manner.

At first, define the vectors of predicted outputs $\mathbf{Y}$, future input differences $\Delta\mathbf{U}$ and future reference trajectory $\mathbf{R}$ as follows

$$\begin{aligned}\mathbf{Y}_i &= [y_i(k+1|k)\ y_i(k+2|k)\ \cdots\ y_i(k+Np_i|k)]^T\\ \mathbf{Y} &= [\mathbf{Y}_1^T\ \mathbf{Y}_2^T\ \cdots\ \mathbf{Y}_n^T]^T\end{aligned} \tag{12}$$

$$\begin{aligned}\Delta\mathbf{U}_i &= [\Delta u_i(k|k)\ \Delta u_i(k+1|k)\ \cdots\ \Delta u_i(k+Nc_i-1|k)]^T\\ \Delta\mathbf{U} &= [\Delta\mathbf{U}_1^T\ \Delta\mathbf{U}_2^T\ \cdots\ \Delta\mathbf{U}_n^T]^T\end{aligned} \tag{13}$$

$$\begin{aligned}\mathbf{R}_i &= [r_i(k+1|k)\ r_i(k+2|k)\ \cdots\ r_i(k+Np_i|k)]^T\\ \mathbf{R} &= [\mathbf{R}_1^T\ \mathbf{R}_2^T\ \cdots\ \mathbf{R}_n^T]^T\end{aligned} \tag{14}$$

$\forall\ i = 1,\ldots,n$ and $y(k+j|k)$ denotes the jth prediction based on the current time instant k. The cost function which is required to be minimized can then be stated as

$$\mathbf{J} = (\mathbf{R}-\mathbf{Y})^T\mathbf{Q}(\mathbf{R}-\mathbf{Y}) + \Delta\mathbf{U}^T\mathbf{\Lambda}\Delta\mathbf{U} \tag{15}$$

where $\mathbf{Q} = diag(Q_1\ Q_2\ \cdots\ Q_n)$ and $\mathbf{\Lambda} = diag(\Lambda_1\ \Lambda_2\ \cdots\ \Lambda_n)$ are positive definite and positive semi-definite block diagonal weighting matrices, respectively, with $Q_i > 0 \in \mathbb{R}^{Np_i\times Np_i}$ and $\Lambda_i \geq 0 \in \mathbb{R}^{Nc_i\times Nc_i}$, $\forall\ i = 1,\ldots,n$ being diagonal matrices. Minimizing (15) with respect to the decision variables $\Delta\mathbf{U}$ requires that the output predictions are expressed in terms of $\Delta\mathbf{U}$. Considering the ith output and taking (7) into account

$$\begin{aligned} y_i(k+1|k) &= C_i\mathbf{x}(k+1|k) = C_i(G\mathbf{x}(k|k) + D\Delta\mathbf{u}(k|k))\\ y_i(k+2|k) &= C_i\mathbf{x}(k+2|k)\\ &= C_i(G\mathbf{x}(k+1|k) + D\Delta\mathbf{u}(k+1|k))\\ &= C_iG^2\mathbf{x}(k|k) + C_iGD\Delta\mathbf{u}(k|k)\\ &\qquad + C_iD\Delta\mathbf{u}(k+1|k)\end{aligned}$$

where C_i denotes the ith row of C. Proceeding in this manner and assuming that $\Delta u_i(k+j|k) = 0,\ \forall\ j \geq Nc_i$, the generalisation can be summarized by

$$\begin{aligned}\mathbf{Y} = \begin{bmatrix}\mathbf{Y}_1\\ \mathbf{Y}_2\\ \vdots\\ \mathbf{Y}_n\end{bmatrix} &= \begin{bmatrix}F_1\\ F_2\\ \vdots\\ F_n\end{bmatrix}\mathbf{x}(k) + \begin{bmatrix}\Phi_{11} & \Phi_{12} & \cdots & \Phi_{1n}\\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ \Phi_{n1} & \Phi_{n2} & \cdots & \Phi_{nn}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{U}_1\\ \Delta\mathbf{U}_2\\ \vdots\\ \Delta\mathbf{U}_n\end{bmatrix}\\ &= \mathbf{F}\,\mathbf{x}(k) + \mathbf{\Phi}\Delta\mathbf{U}\end{aligned} \tag{16}$$

where

$$F_i = \begin{bmatrix} C_iG\\ C_iG^2\\ \vdots\\ C_iG^{Np_i}\end{bmatrix} \tag{17}$$

$$\Phi_{jl} = \begin{bmatrix} C_jD_l & 0 & 0 & \cdots\\ C_jGD_l & C_jD_l & 0 & \cdots\\ C_jG^2D_l & C_jGD_l & C_jD_l & \cdots\\ \vdots & & & \\ C_jG^{Np_j-1}D_l & C_jG^{Np_j-2}D_l & C_jG^{Np_j-3}D_l & \cdots\\ & & & 0\\ & & & 0\\ & & & 0\\ & & & \\ & & & C_jG^{Np_j-Nc_l}D_l\end{bmatrix} \tag{18}$$

and D_l denotes the lth column of D.

Also, (16) indicates that cross-coupling effects are eliminated if the off-diagonal matrices in $\mathbf{\Phi}$ are null. This is the case if the input coefficient matrices in (1) are diagonal matrices. In this case, the ith output is only dependent on the ith input and each input-output pair can be considered as individual SISO systems. Certainly, the state vector (8) still contains values of all n inputs

and n outputs but these are previously measured values, hence it can be considered as a known constant vector at each current time instant k. Furthermore, it should be mentioned that the smallest prediction horizon is required to be greater or equal the largest control horizon in order to avoid negative powers of G in (18).

Substituting (16) into (15) yields the cost function depending on $\Delta\mathbf{U}$ only, which is required to be minimized.

$$\underset{\Delta\mathbf{U}}{min.}\ \Delta\mathbf{U}^T(\mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi}+\mathbf{\Lambda})\Delta\mathbf{U}+2\Delta\mathbf{U}^T\mathbf{\Phi}^T\mathbf{Q}(\mathbf{F}\mathbf{x}(k)-\mathbf{R}) \tag{19}$$

The solution of this optimization problem is

$$\Delta\mathbf{U}=-(\mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi}+\mathbf{\Lambda})^{-1}\mathbf{\Phi}^T\mathbf{Q}(\mathbf{F}\mathbf{x}(k)-\mathbf{R}) \tag{20}$$

which is in the state variable feedback form

$$\Delta\mathbf{U}=-\mathbf{K}_x\mathbf{x}(k)+\mathbf{K}_s \tag{21a}$$

where

$$\mathbf{K}_x=(\mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi}+\mathbf{\Lambda})^{-1}\mathbf{\Phi}^T\mathbf{Q}\mathbf{F} \tag{21b}$$

$$\mathbf{K}_s=(\mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi}+\mathbf{\Lambda})^{-1}\mathbf{\Phi}^T\mathbf{Q}\mathbf{R} \tag{21c}$$

and since the current control input is applied to the plant

$$\Delta\mathbf{u}(k)=\bar{\mathbf{C}}\,\Delta\mathbf{U} \tag{22}$$

where $\bar{\mathbf{C}}\in\mathbb{R}^{n\times(Nc_1+Nc_2+\ldots+Nc_n)}$ and the ith row is

$$\bar{\mathbf{C}}_i=[\ \overbrace{0\ \cdots\ 0}^{\sum_{j=1}^{i-1}Nc_j}\ \ 1\,0\ \cdots\ 0]$$

whereby the first row begins with unity and is filled with zeros.

Formulating the MPC in the above proposed representation allows a straightforward implementation in the case of dealing with MIMO systems, in particular when individual control and prediction horizons are desired, since these can be used as effective tuning parameters. Also, in combination with the NMSS model representation the potential applicability to practical applications is increased.

3. DECOUPLING

The decoupling of the ith output from the jth input $\forall\ i\neq j$ is achieved by diagonalizing the plant model transfer function matrix, following the approach by (Plummer and Vaughan, 1997), where a pole-placement control structure was used to control the system. In this paper, the above elaborated MPC controller is used and an approach to handle imposed constraints is presented.

3.1 System diagonalization

Consider the system (1) reformulated as follows

$$A(z^{-1})\mathbf{y}(k)=B(z^{-1})\mathbf{u}(k) \tag{23}$$

with

$$A(z^{-1})=I+A_1z^{-1}+A_2z^{-2}+\ldots+A_{n_a}z^{-n_a} \tag{24}$$

and

$$B(z^{-1})=B_1z^{-1}+B_2z^{-2}+\ldots+B_{n_b}z^{-n_b} \tag{25}$$

so that the plant transfer function matrix is given by

$$\mathbf{y}(k)=A^{-1}(z^{-1})B(z^{-1})\mathbf{u}(k) \tag{26}$$

such that when this is represented in a left matrix fraction description (MFD), and $A(z^{-1})$ is in a diagonal form so that no cross-couplings are induced but the off-diagonals in the $B(z^{-1})$ matrix are non zero thus giving rise to cross-couplings. In order to avoid this cross-coupling effect, a transformation of the input is performed

$$\mathbf{u}(k)=E(z^{-1})\mathbf{v}(k) \tag{27}$$

where

$$E(z^{-1})=\text{adj}[B(z^{-1})]z^d \tag{28}$$

which yields, that

$$B_d(z^{-1})=B(z^{-1})E(z^{-1})=|B(z^{-1})|\ z^d\ I \tag{29}$$

where $|\ \cdot\ |$ denotes the determinant. Substituting (29) into both (27) and (26) gives a diagonalized system representation from the input $\mathbf{v}(k)$ to the output $\mathbf{y}(k)$, i.e.

$$\begin{aligned}\mathbf{y}(k)&=A^{-1}(z^{-1})B(z^{-1})E(z^{-1})\mathbf{u}(k)\\&=A^{-1}(z^{-1})B_d(z^{-1})\mathbf{v}(k)\end{aligned} \tag{30}$$

Here, an additional polynomial matrix $E(z^{-1})$ is required to be incorporated into the MPC structure. Additionally, z^d is required to be chosen. The choice in a predictive control structure is not restricted to choosing d such that $E(z^{-1})$ is non-causal, as in the case of non predictive control, since future control actions are available provided that the corresponding control horizons are sufficiently large.

3.2 Decoupling MPC

From (30) it seems natural to consider the system in this representation and to use the input coefficient matrix in polynomial form $B_d(z^{-1})$ in the control structure so that effectively, the input coefficient matrices in (1) are replaced by the coefficients of $B_d(z^{-1})$ and the input is now $\mathbf{v}(k)$ instead of $\mathbf{u}(k)$. Now, these coefficient matrices are of a diagonal form and when deriving the MPC as described in Section 2.2, the off-diagonal block matrices of $\mathbf{\Phi}$ in (16) are zero and consequently the cost function in the form (19) is the sum of the individual SISO cost functions of the n input-output $(\mathbf{Y}_i,\ \Delta\mathbf{V}_i)$ pairs. As a result, the polynomial matrix $E(z^{-1})$ is incorporated into the model, hence in the MPC structure. However, the input to this system and subsequently the controller output is now $\mathbf{v}(k)$. This means that $\mathbf{u}(k)$ is required to be recovered, since $\mathbf{u}(k)$ has to be applied to the real plant. This can be achieved by making use of (27).

Furthermore, when the polynomials in $E(z^{-1})$ are non-causal, $\mathbf{u}(k)$ is dependent on current and previous values of $\mathbf{v}(k)$. It is noted, however, as indicated earlier, that causal polynomials can be used since predictive control is employed. Certainly, making use of predicted values of the input might cause inaccuracy and a loss of performance. This issue is also discussed in the simulation example, see Section 4.

Assuming that z^d is chosen such that $E(z^{-1})$ is non-causal

$$E(z^{-1})=\begin{bmatrix}\mathbf{e_{11}}(z^{-1}) & \mathbf{e_{12}}(z^{-1}) & \cdots & \mathbf{e_{1n}}(z^{-1})\\ \mathbf{e_{21}}(z^{-1}) & \mathbf{e_{22}}(z^{-1}) & \cdots & \mathbf{e_{2n}}(z^{-1})\\ \vdots & \vdots & \cdots & \vdots\\ \mathbf{e_{n1}}(z^{-1}) & \mathbf{e_{n2}}(z^{-1}) & \cdots & \mathbf{e_{nn}}(z^{-1})\end{bmatrix} \tag{31}$$

with

$$\mathbf{e_{jl}}(z^{-1})=e_{jl_1}+e_{jl_2}z^{-1}+e_{jl_3}z^{-2}+\ldots \tag{32}$$

and the entire predicted input vector $\mathbf{U}$ can be recovered by

$$\mathbf{U}=\begin{bmatrix}\tilde{E}_{11} & \tilde{E}_{12} & \cdots & \tilde{E}_{1n}\\ \tilde{E}_{21} & \tilde{E}_{22} & \cdots & \tilde{E}_{2n}\\ \vdots & \vdots & \cdots & \vdots\\ \tilde{E}_{n1} & \tilde{E}_{n2} & \cdots & \tilde{E}_{nn}\end{bmatrix}\mathbf{V}+\begin{bmatrix}\tilde{b}_1\\ \tilde{b}_2\\ \vdots\\ \tilde{b}_n\end{bmatrix}\mathbf{V}^-=\tilde{\mathbf{E}}\,\mathbf{V}+\tilde{\mathbf{b}}\,\mathbf{V}^- \tag{33}$$

with

$$\tilde{E}_{jl} = \begin{bmatrix} e_{jl_1} & 0 & 0 & \cdots \\ e_{jl_2} & e_{jl_1} & 0 & \cdots \\ e_{jl_3} & e_{jl_2} & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \in \mathbb{R}^{Nc_j \times Nc_l} \tag{34}$$

and

$$\tilde{b}_j = \begin{bmatrix} e_{j1_2} & e_{j1_3} & \cdots & \cdots & \cdots & e_{j2_2} & e_{j2_3} & \cdots & \cdots & \cdots \\ e_{j1_3} & e_{j1_4} & \cdots & \cdots & 0 & e_{j2_3} & e_{j2_4} & \cdots & \cdots & 0 \\ e_{j1_4} & e_{j1_5} & \cdots & 0 & 0 & e_{j2_4} & e_{j2_5} & \cdots & 0 & 0 \\ \cdots & & & & & & & & & \end{bmatrix}$$

$$\begin{bmatrix} \cdots & e_{jn_2} & e_{jn_3} & \cdots & \cdots & \cdots \\ \cdots & e_{jn_3} & e_{jn_4} & \cdots & \cdots & 0 \\ \cdots & e_{jn_4} & e_{jn_5} & \cdots & 0 & 0 \\ \cdots & & & & & \end{bmatrix} \tag{35}$$

where the number of rows is Nc_j and the number of columns depend on the order of $\mathbf{e_{jl}}(z^{-1})$, hence on the choice of d and the order of the polynomials in $B(z^{-1})$.

Equation (33) is partitioned into two terms. The first term depends on the current value and predictions $\mathbf{V}$ and the second term on past values only

$$\mathbf{V}^- = [\mathbf{v}^T(k-1) \quad \mathbf{v}^T(k-2) \quad \mathbf{v}^T(k-3) \ \cdots]^T \tag{36}$$

Note here that the matrices (34) and (35) are in a Toeplitz form, hence (33) basically describes the convolution of (31) with the sequence $\mathbf{V}$. Also note that the MPC controller makes use of $\Delta\mathbf{V}$ instead of $\mathbf{V}$. However, this does not require any changes to the formulation above except that multiplying (33) by Δ, i.e. there are no changes to the matrices involved in (34) and (35).

3.3 Constraints

When constraints are required to be imposed on the plant input, input difference or output, again the difficulty appears that plant input is not available directly. For this reason, use is made of (33) whereby the term depending on previous values only can be considered as a constant term in the optimization problem at each sampling instance

$$\mathbf{Z} = \tilde{\mathbf{b}}\,\mathbf{V}^- \tag{37}$$

In this paper, use is made of the closed-loop paradigm (CLP) (Rossiter, 2004) in order to handle constraints, where a perturbation term is added to the optimal, unconstrained control law (21), which is in the decoupling case

$$\Delta\mathbf{V} = -\mathbf{K}_x\mathbf{x}(k) + \mathbf{K}_s + \beta = \mathbf{K} + \beta \tag{38}$$

where $\beta \in \mathbb{R}^{(Nc_1+\ldots+Nc_n)\times 1}$ denotes the disturbance vector whose elements are zero when constraints are not active. Thus when $\beta = 0$ the control action is optimal and subsequently, $\beta^T\beta$ is required to be as small as possible in the case of active constraints.

At first, consider constraints on $\Delta\mathbf{U}$. Multiplying (38) by $\tilde{\mathbf{E}}$ and taking (33) as well as (37) into account, gives

$$\Delta\mathbf{U} = \tilde{\mathbf{E}}(\mathbf{K} + \beta) + \Delta\mathbf{Z} \tag{39}$$

so that

$$\Delta\mathbf{U}^{min} \le \tilde{\mathbf{E}}(\mathbf{K} + \beta) + \Delta\mathbf{Z} \le \Delta\mathbf{U}^{max} \tag{40}$$

and the constraints can be formulated as

$$\tilde{\mathbf{E}}\beta \le \mathbf{c_1} \tag{41}$$

with

$$\mathbf{c_1} = \Delta\mathbf{U}^{max} - \Delta\mathbf{Z} - \tilde{\mathbf{E}}\mathbf{K} \tag{42}$$

which implies that when $\mathbf{c_1}_j < 0$, then the optimal control law violates constraints and action is required.
Similarly

$$-\tilde{\mathbf{E}}\beta \le \mathbf{c_2} \tag{43}$$

with

$$\mathbf{c_2} = -(\Delta\mathbf{U}^{min} - \Delta\mathbf{Z} - \tilde{\mathbf{E}}\mathbf{K}) \tag{44}$$

where constraints are violated by the optimal control law when $\mathbf{c_2}_j < 0$. In the case where action is required, the following optimization problem is required to be solved

$$\begin{aligned} &\min_{\beta} \quad \beta^T\beta \\ &subject\ to \quad \begin{bmatrix} \tilde{\mathbf{E}} \\ -\tilde{\mathbf{E}} \end{bmatrix} \beta \le \begin{bmatrix} \mathbf{c_1} \\ \mathbf{c_2} \end{bmatrix} \end{aligned} \tag{45}$$

In this manner, also constraints on $\mathbf{U}$ and $\mathbf{Y}$, when additionally using (16), can be handled.

4. SIMULATION EXAMPLE

A simulation example is considered by making use of the following 2 input, 2 output MIMO system

$$G(s) = \begin{bmatrix} \dfrac{4e^{-27s}}{45s+1} & \dfrac{2e^{-21s}}{60s+1} \\ \dfrac{5.5e^{-17s}}{15s+1} & \dfrac{6e^{-12s}}{50s+1} \end{bmatrix} \tag{46}$$

which is discretised by making use of a sampling interval $T_s = 7$ s. The simulation results corresponding to a decoupling and a non-decoupling MPC controller are shown in Figure 1. In the decoupling case, $d = 2$ is chosen such that $E(z^{-1})$ is non-causal. Furthermore, the weighting matrices are $Q_1 = 0.05I, Q_2 = 0.55I$ in the decoupling MPC and $\mathbf{Q} = 0.1I$ in the non-decoupling MPC. The weighting matrices are chosen such that similar rise and settling times are achieved. $\mathbf{\Lambda} = I$ in both cases and the prediction and control horizons are chosen to be $Np_1 = 100, Np_2 = 80$ and $Nc_1 = Nc_2 = 50$. The reference signal is a step of amplitude 3 units where this reference change of the first output is applied at $t = 150$ s and to the second output at $t = 300$ s. As shown in Figure 1, the non-decoupling

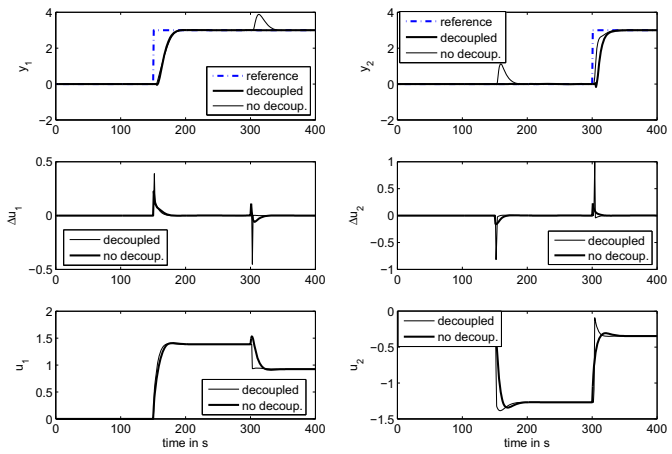


Fig. 1. Inputs, input differences and outputs of the decoupling and non-decoupling MPC

MPC controller allows significant response in the output where the reference signal does not change compared to decoupling MPC where almost no cross-coupling effect is observable. The control actions of the decoupled and non-decoupling MPC are

similar, but the amplitude of the decoupling MPC exceeds the amplitude of the non-decoupling MPC. This is perhaps not surprising.

Now, constraints are imposed on $\Delta\mathbf{U}$, where $-0.05 \leq \Delta\mathbf{U}_1 \leq 0.05$ and $-0.1 \leq \Delta\mathbf{U}_2 \leq 0.1$ and the results are shown in Figure 2. The response is compared to the unconstrained non-decoupling MPC and it is shown that the cross-coupling effects are still not as intensive as in the unconstrained non-decoupling case. Finally, $d = 4$ is chosen so that $E(z^{-1})$ is causal, hence

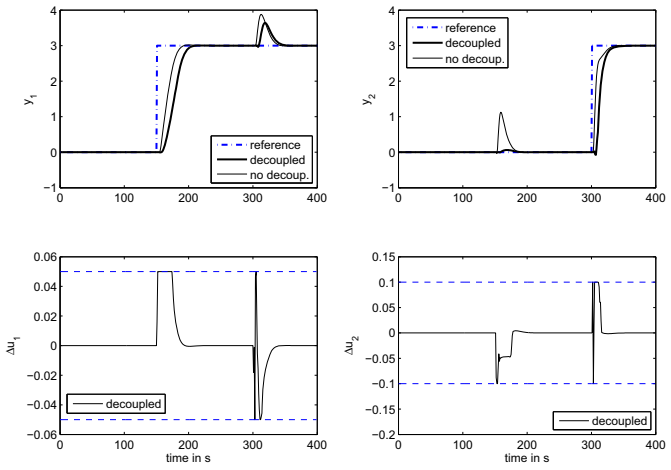


Fig. 2. Constraints imposed on $\Delta\mathbf{U}$

it is required to make use of the predicted control inputs in order to obtain the current value. Figure 3 shows the simulation results where the weighting matrices of the decoupling MPC are changed to $Q_1 = 0.0005I$ and $Q_2 = 0.0055I$ in order to obtain satisfactory results. Here cross-coupling effects can

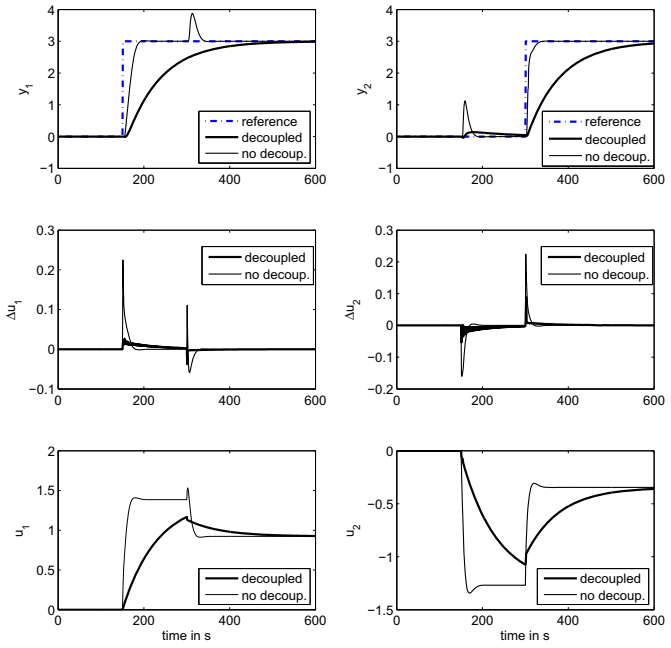


Fig. 3. Case of causal $E(z^{-1})$

be observed and the settling time is increased. This is believed to be caused by the inaccuracy of the predictions used in the calculation of the control.

In practice the system parameters are usually not perfectly known. This can be caused by non-linearities, which are approximated to be linear, or imperfectly estimated system parameters, hence, the performance of the decoupling MPC is of interest when uncertainties in the model parameters are present. This issue is addressed in the following.

The matrix $E(z^{-1})$ is chosen to be non-causal and the controller parameters are the same as used in Figure 1.

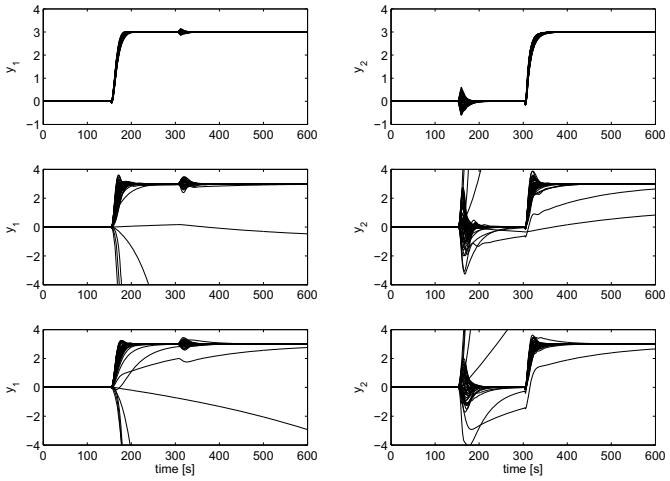


Fig. 4. Upper: uncertainties in the numerator parameters; Middle: uncertainties in the denominator parameters; Lower: uncertainties in the numerator and denominator parameters

In Figure 4 in the upper plot, the model parameters of the numerator are varied in the range of $\pm 10\,\%$ of the nominal value. This means, that a uniformly distributed random number in this interval is added to the nominal value and 50 Monte-Carlo runs are performed. In this manner, the denominator parameters are varied in the plot in the middle. But here, the interval is chosen such that the time constants vary in the range of $\pm 50\,\%$ of the nominal time constants. The lower plot shows the performance when the numerator and denominator parameters are varied in the same intervals as in the plots above.

As indicated in Figure 4, the numerator parameter uncertainties do not impact the performance significantly. Compared to Figure 1, the cross-coupling effects are still less in amplitude than in the non-decoupling case. But, it seems that the performance is more sensitive to uncertainties in the denominator parameters. Here, it can lead to unstable results. This can be overcome by increasing the weighting matrix $\mathbf{\Lambda}$. As shown in Figure 5, when the weighting matrix is $\mathbf{\Lambda} = 700I$, stable results are obtained. Since the poles of the model are still inside the unit circle in the entire interval, the unstable results are caused by ill-conditioned matrix $(\mathbf{\Phi}^T\mathbf{Q}\mathbf{\Phi} + \mathbf{\Lambda})$ in (20). One way to overcome this issue is to increase $\mathbf{\Lambda}$. The lower plot in Figure 4 shows the performance when uncertainties are applied to the numerator and denominator parameters and this performance is similar to that when denominator uncertainties are applied only. This is not unexpected since the numerator uncertainties affect the performance marginally compared to denominator uncertainties.

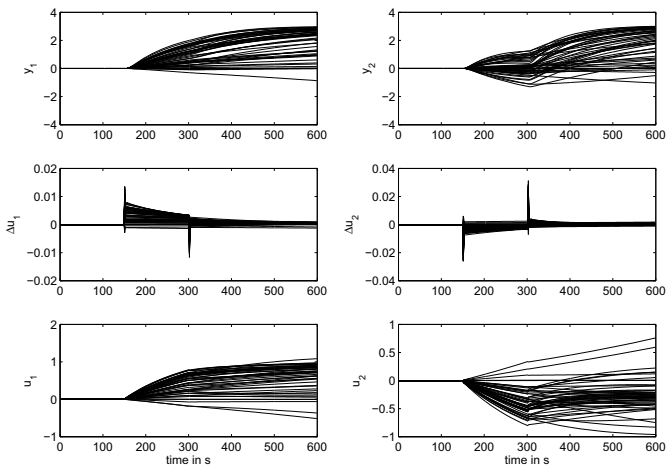


Fig. 5. Uncertain denominator parameters and increased $\mathbf{\Lambda}$

5. CONCLUSION

In this paper, an approach for reducing the cross-coupling effects by making use of a model predictive control structure is presented. The model used is in a non-minimal state space form and the MPC controller allows the incorporation of individual control and prediction horizons in a straightforward manner. By diagonalising the system, an additional matrix in polynomial form is introduced. This matrix requires the adjoint of the input coefficient matrix in polynomial form to be computed as well as the determinant. Furthermore, the choice of the forward shift z^d affects the performance significantly since the control input predictions may not be sufficiently accurate in order to achieve a desired performance criteria.

REFERENCES

Bitmead, R., Gevers, M., and Wertz, V. (1990). *Adaptive Optimal Control: The thinking Man's GPC*. Prentice Hall Int., Australia.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, New York.

Camacho, E.F. and Bordons, C. (2007). *Model Predictive Control*. Springer, London, 2nd edition.

Clarke, D., Mohtadi, C., and Tuffs, P. (1987a). Generalized predictive control - part I. the basic algorithm. *Automatica*, 23(2), 137–148.

Clarke, D., Mohtadi, C., and Tuffs, P. (1987b). Generalized predictive control - part II. extensions and interpretations. *Automatica*, 23(2), 149–160.

Exadaktylos, V. and Taylor, C.J. (2010). Multi-objective performance optimisation for model predictive control by goal attainment. *International Journal of Control*, 83(7), 1374–1386.

Ikonen, E. and Najim, K. (2002). *Advanced Process Identification and Control*. Marcel Dekker Inc., New York.

Kothare, M.V., Balakrishnan, V., and Morari, M. (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10), 1361–1379.

Kouvaritakis, B., Rossiter, J.A., and Schuurmans, J. (2000). Efficient Robust Predictive control. *IEEE Transaction on Automatic Control*, 45(10), 1545–1549.

Kwon, W.H. and Han, S. (2005). *Receding Horizon Control: Model Predictive Control for State Models*. Springer, Leipzig.

Lee, M.J., Young, P.C., Chotai, A., and Tych, W. (1995). A non-minimal state variable feedback approach to multivariable control of glasshouse climate. *Trans. Inst. Measurement and Control*, 17(4), 200–211.

Maciejowski, J. (2001). *Predictive Control with Constraints*. Prentice Hall.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., and Scokaert, P. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789–814.

Nocedal, J. and Wright, S.J. (2006). *Numerical Optimization*. Springer, New York, 2nd edition.

Plummer, A.R. and Vaughan, N.D. (1997). Decoupling pole-placement control, with application to a multi-channel electro-hydraulic servosystem. *Control Eng. Practice*, 5(3), 313–323.

Rossiter, J.A. (2004). *Model-based Predictive Control: A Practical Approach*. CRC Press, Boca Raton.

Wang, C.L. and Young, P.C. (1988). Direct digital and adaptive control by input-output state variable feedback pole assignment. *International Journal of Control*, 47(1), 97–109.

Wang, L. and Young, P.C. (2006). An improved structure for model predictive control using non-minimal state space realisation. *Journal of Process Control*, 16(4), 355–371.

Wang, L. (2009). *Model Predictive Control System Design and Implementation using MATLAB*. Springer, London.

Young, P.C., Behzadi, M.A., Wang, C.L., and Chotai, A. (1987). Direct digital and adaptive control by input-output state variable feedback pole assignment. *International Journal of Control*, 46(6), 1867–1881.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Design of Unknown Input Reconstruction Algorithm in Presence of Measurement Noise

Małgorzata Sumisławska * **Tomasz M. Larkowski** *
Keith J. Burnham *

* *Control Theory and Applications Centre, Coventry University, Priory Street, Coventry, CV1 5FB, UK*
(e-mail: sumislam@uni.coventry.ac.uk).

Abstract: In this paper a novel approach to unknown input estimation based on a parity equations concept is developed. Unlike the unknown input observers based on the Kalman filtering approach, the observer proposed here is independent of the system state vector. Therefore, due to the reduction of the number of estimated signals, a higher accuracy of the input estimation is achieved. This makes the scheme advantageous in cases when the accuracy of the input estimate is crucial and the knowledge about the system states is not required. By increasing the order of the parity space, which is a tuning parameter of the algorithm proposed, the new approach allows the influence of the effects of measurement noise to be reduced. A Lagrange multiplier method is used to obtain an analytical solution for the filter parameters.

Keywords: Filtering, observers, parity equations, unknown input reconstruction

1. INTRODUCTION

The history of observers dates back to the 1960s with the Luenberger system state observers (see Luenberger (1964)). Subsequently, state observers have been extended to the class of systems with both, known and unknown system inputs (see for example Darouach and Zasadzinski (1997)). Over the last decade the simultaneous estimation of both, state vector and unknown inputs, based on a Kalman filtering approach has gained an interest (cf. Floquet and Barbot (2006); Hsieh (2000)). Gillijns and De Moor (2007a) combined the state observer proposed by Darouach and Zasadzinski (1997) and the unknown input estimator of Hsieh (2000) creating a state and unknown input observer, which is optimal in the minimum variance sense. This approach has subsequently been extended to the case of a linear system with a direct feedthrough (see Gillijns and De Moor (2007b)).

In this paper a new approach to the problem of unknown input estimation based on parity equations (PE) is proposed. A detailed explanation of PE can be found in Ding (2008); Gertler (1991); Li and Shah (2002). A very general relationship between the PE and the left inverse of the minimum-phase deterministic system has been presented by Edelmayer (2005). On the contrary, in this paper PE are used to obtain an approximation (estimate) of the unknown input of a stochastic system, whose measurements are affected by noise. The method is suitable for both minimun-phase and nonminimum-phase systems. The contribution of this paper is to utilise the Lagrange multiplier method to provide an analytical solution for the filter parameters, which minimise effects of measurement noise. Furthermore, unlike the unknown input observers (UIOs) based on the Kalman filtering approach (see for example Gillijns and De Moor (2007a,b)), the developed observer is orthogonal to the system state vector.

In the framework of this paper, firstly, the PE theory is explained. Subsequently, the derivation of the novel filter is provided. Use is made of the Lagrange multiplier method (see for example Bertsekas (1982)) to obtain an analytical solution to the unknown input estimator parameters. Then, based on a numerical example, the influence of the parity space order (which is a tuning parameter of the filter) on the efficacy of the algorithm is analysed. Finally, the accuracy of the novel method is compared to that of the Kalman filter-based minimum variance unbiased unknown input estimator proposed by Gillijns and De Moor (2007b).

2. DESCRIPTION OF APPROACH

In this section the new algorithm is derived. Firstly, for completeness, PE are described in Subsection 2.1, see e.g. Li and Shah (2002). Then, in Subsection 2.2, using existing concepts, a new unknown input observer based on PE (further referred to as PE-UIO) is developed.

2.1 Parity Equations

Assume that a linear dynamic discrete time two-input single-output system is represented by an n^{th} order state space equation of the following form:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu_0(t) + Gv(t) \\ y_0(t) &= Cx(t) + Du_0(t) + Hv(t) \\ u(t) &= u_0(t) + \tilde{u}(t) \\ y(t) &= y_0(t) + \tilde{y}(t) \end{aligned} \tag{1}$$

where $A \in \mathcal{R}^{n\times n}$, $B \in \mathcal{R}^{n\times 1}$, $C \in \mathcal{R}^{1\times n}$, $D \in \mathcal{R}^{1\times 1}$, $G \in \mathcal{R}^{n\times 1}$ and $H \in \mathcal{R}^{1\times 1}$. The terms $u_0(t)$, $v(t)$ and $y_0(t)$ refer to, respectively, known and unknown input to the system and the system output. An errors-in-variables (EIV) case is considered (see, for example, Söderström (2007)), i.e. all measured variables, which are input $u(t)$ and $y(t)$, are affected by a zero mean, white Gaussian mutually uncorrelated measurement noise sequences denoted by $\tilde{u}(t)$ and $\tilde{y}(t)$, respectively. Hence, the noise free but unmeasured system input and output are denoted as $u_0(t)$ and $y_0(t)$, respectively.

The following stacked vector of the unknown input, $v(t)$, is created (see, for example, Li and Shah (2002)):

$$V = [\, v(t-s)\ v(t-s+1)\ \cdots\ v(t)\,]^T \tag{2}$$

where the term s denotes the order of the parity space. Analogously, one can build stacked vectors of $y(t)$, $y_0(t)$, $\tilde{y}(t)$, $u(t)$, $u_0(t)$ and $\tilde{u}(t)$ which are denoted, respectively, as Y, Y_0, $\tilde{Y}$, U, U_0 and $\tilde{U}$. By making use of this notation the system (1) can be expressed in the form of:

$$Y_0 = \Gamma x(t-s) + QU_0 + TV \tag{3}$$

where Γ is an extended observability matrix:

$$\Gamma = \left[\, C^T\ A^TC^T\ \cdots\ (A^s)^TC^T\,\right]^T \in \mathcal{R}^{(s+1)\times n} \tag{4}$$

and Q is the following block Toeplitz matrix:

$$Q = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{s-1}B & CA^{s-2}B & \cdots & D \end{bmatrix} \in \mathcal{R}^{(s+1)\times(s+1)} \tag{5}$$

Analogously, the matrix T is built by replacing D with H and B with G in the matrix Q. In order to eliminate the unknown state vector from (3), a row vector W is defined, which belongs to the left nullspace of Γ, i.e.

$$W\Gamma = 0 \tag{6}$$

Hence (3) can be reformulated as:

$$WTV = WY_0 - WQU_0 = W(Y - \tilde{Y}) - WQ(U - \tilde{U}) \tag{7}$$

By rearranging the measured (known) variables on the right-hand side of (7) and the unknowns on the left-hand side, the following PE is obtained (cf. Li and Shah (2002)):

$$WTV + W\tilde{Y} - WQ\tilde{U} = WY - WQU \tag{8}$$

In the next section use is made of the PE in order to derive a novel algorithm for the unknown input estimation.

2.2 Input reconstruction with measurement noise filtering

Denote the matrix spanning the left nullspace of Γ as $\Gamma^\perp$. Consequently, the row vector W is a linear combination of rows of $\Gamma^\perp$. It is assumed here that the system (1) is observable, hence the extended observability matrix Γ is of full rank. Therefore, the dimension of $\Gamma^\perp$ is $(s - n + 1) \times (s + 1)$, and since T is square, it is true that $\Gamma^\perp T \in \mathcal{R}^{(s-n+1)\times(s+1)}$. Thus in the case of noise-free input and output measurements, i.e. when $U{=}U_0$ and $Y{=}Y_0$, the following equation holds (cf. (7)):

$$\Gamma^\perp TV = b \tag{9}$$

where b is a column vector of $(s - n + 1)$ elements, and:

$$b = \Gamma^\perp Y - \Gamma^\perp QU \tag{10}$$

Note, that the matrix T consists of the Markov parameters of the relation between the unknown input and the output, which are given by (see Kirtikar et al. (2009)):

$$T_i = \begin{cases} H & ,i = 0 \\ CA^{i-1}G & ,i > 0 \end{cases} \tag{11}$$

The relative degree of the system $G_v(z) = C(zI-A)^{-1}G+H$, denoted as r, is the smallest number for which $T_r \neq 0$ (cf. Edelmayer (2005)). Hence, one can note that (10) is a homogenous set of equations (i.e. the sequence of unknown input values can be determined explicitly from (10)) only, if the system $G_v(z)$ has no zeros, i.e. the relative degree is equal to the order of $G_v(z)$. (Which means that the last r columns of the matrix $\Gamma^\perp T$ are equal zero.) Nevertheless, the unique solution to the set of equations (10) can be seriously affected by the measurement noise $\tilde{u}(t)$ and $\tilde{y}(t)$. The algorithm proposed here minimises the effects of the unwanted measurement noise. Furthermore, the technique can be utilised to yield an approximation of $v(t)$ in the case when $G_v(z)$ has zeros. The proposed method is suitable for both minimum-phase and nonminimum-phase systems.

It is proposed to calculate the value of the unknown input as:

$$\hat{v}(t) = WY - WQU \tag{12}$$

which, in the case of a noise-free input and output measurements, is:

$$\hat{v}(t) = WTV \tag{13}$$

Thus, based on the assumption that the unknown input is slowly varying, its estimate can be calculated as a linear combination of the sequence $v(t-s)$, $v(t-s+1), \cdots, v(t)$.

$$\hat{v}(t) = \alpha_0 v(t) + \alpha_1 v(t-1) + \cdots + \alpha_s v(t-s) \tag{14}$$

where the α parameters are dependent on the choice of the vector W, such that:

$$WT = [\, \alpha_s\ \alpha_{s-1}\ \cdots\ \alpha_0\,]^T \tag{15}$$

One can note that (14) is an equation of a moving average finite impulse response filter with the gain given by the sum of the α parameters, i.e. the sum of elements of the vector WT. Thus, it is suggested here that the vector W should be selected in such a way, that the sum of elements of the vector WT is equal unity.

It is anticipated that the choice of the order of the parity space s, as well as the vector W may influence a lag in the estimate of the unknown input (due to the moving average filtering property of the unknown input estimator).

In the next subsection an algorithm for the selection of an optimal vector W is derived based on the Lagrange multiplier method.

2.3 Selection of optimal W

In the case of noisy input and output measurements, equation (12) becomes:

$$\hat{v}(t) = WTV + W\tilde{Y} - WQ\tilde{U} \tag{16}$$

Hence, the estimate of the unknown input is affected by a coloured noise. However, by a careful choice of W, the degrading effect of noise can be minimised. Due to the fact that $\tilde{y}(t)$ and $\tilde{u}(t)$ are uncorrelated, white and zero mean (i.e. the expected values $E\{\tilde{y}(t)\} = E\{\tilde{u}(t)\} = 0$), it is true that:

$$E\{W\tilde{Y} - WQ\tilde{U}\} = 0 \tag{17}$$

Hence asymptotically, a presence of the measurement noise does not cause a bias in the unknown input estimate. Furthermore, an influence of the measurement noise on the unknown input estimate can be minimised by reducing the variance of the term $W\tilde{Y} - WQ\tilde{U}$, i.e.:

$$E\{(W\tilde{Y} - WQ\tilde{U})(W\tilde{Y} - WQ\tilde{U})^T\} = W\Sigma_{\tilde{y}}W^T + \\ + WQ\Sigma_{\tilde{u}}Q^TW^T - W\Sigma_{\tilde{u}\tilde{y}}^TQ^TW^T - WQ\Sigma_{\tilde{u}\tilde{y}}W^T \quad (18)$$

where $\Sigma_{\tilde{u}} = E\{\tilde{U}\tilde{U}^T\}$, $\Sigma_{\tilde{y}} = E\{\tilde{Y}\tilde{Y}^T\}$, $\Sigma_{\tilde{u}\tilde{y}} = E\{\tilde{U}\tilde{Y}^T\}$. Due to the fact that the input and output measurement sequences are considered to be white, zero mean and mutually uncorrelated:

$$\Sigma_{\tilde{u}} = \sigma_{\tilde{u}}^2 I, \quad \Sigma_{\tilde{y}} = \sigma_{\tilde{y}}^2 I, \quad \Sigma_{\tilde{u}\tilde{y}} = 0 \quad (19)$$

where the terms $\sigma_{\tilde{u}}^2$ and $\sigma_{\tilde{y}}^2$ refer to the variance of the measurement error of the system input and output, respectively, whilst I is an identity matrix of appropriate dimension.

Subsequently, the vector W should be selected to minimise the cost function $f(W)$:

$$f(W) = W\Sigma_{\tilde{y}}W^T + WQ\Sigma_{\tilde{u}}Q^TW^T \quad (20)$$

subject to the following constraints:

(1) Sum of elements of WT is equal to 1.
(2) $W\Gamma = 0$.

The cost function (20) can be minimised by making use of the Lagrange multipliers method (see, for example, Bertsekas (1982)). Denote the rows of $\Gamma^\perp$ by γ_1, γ_2, ..., $\gamma_{(s-n+1)}$:

$$\Gamma^\perp = \left[\gamma_1^T \ \gamma_2^T \ \cdots \ \gamma_{(s-n+1)}^T \right]^T \quad (21)$$

The vector W is a linear combination of rows of $\Gamma^\perp$, i.e.

$$W = \sum_{i=1}^{s-n+1} p_i\gamma_i \quad (22)$$

Hence the cost function (20) can be reformulated as a function of the parameter vector $P = [\, p_1 \ p_2 \ \cdots \ p_{s-n+1} \,]^T$:

$$f(P) = \left(\sum_{i=1}^{k} p_i\gamma_i \right) \Sigma \left(\sum_{j=1}^{k} p_j\gamma_j^T \right) = \sum_{i=1}^{k}\sum_{j=1}^{k} p_ip_j\gamma_i\Sigma\gamma_j^T \quad (23)$$

where $k = s - n + 1$ and:

$$\Sigma = \Sigma_{\tilde{y}} + Q\Sigma_{\tilde{u}}Q^T \quad (24)$$

The cost function $f(P)$ is required to be minimised subject to the constraint:

$$g(P) = \mathrm{sum}_{row}(WT) - 1 = 0 \quad (25)$$

where the operator $\mathrm{sum}_{row}(A)$ denotes a column vector whose elements are sums of the appropriate rows of the matrix A.

The solution to the Lagrange minimisation problem is given by (see Bertsekas (1982)):

$$\nabla f(P) = \lambda \nabla g(P) \quad (26)$$

The partial derivative of $f(P)$ with respect to the i^{th} element of the vector P (denoted as p_i) is:

$$\frac{\partial f(P)}{\partial p_i} = \sum_{j=1}^{k} p_j\gamma_i\Sigma\gamma_j^T + \sum_{j=1}^{k} p_j\gamma_j\Sigma\gamma_i^T \quad (27)$$

After some manipulations the gradient of $f(P)$ is reformulated as:

$$(\nabla f(P))^T = \left(\Gamma^\perp\Sigma(\Gamma^\perp)^T + (\Gamma^\perp\Sigma\left(\Gamma^\perp\right)^T)^T \right) P \quad (28)$$

The partial derivative of the constraint function $g(P)$ with respect to p_i is calculated via:

$$\frac{\partial g(P)}{\partial p_i} = \mathrm{sum}_{row}(\gamma_i T) \quad (29)$$

Thus, the gradient of $g(P)$ can be reformulated as:

$$(\nabla g(P))^T = \mathrm{sum}_{row}(\Gamma^\perp T) \quad (30)$$

By making use of the notation:

$$S = \left(\Gamma^\perp\Sigma(\Gamma^\perp)^T + (\Gamma^\perp\Sigma\left(\Gamma^\perp\right)^T)^T \right) \quad (31)$$

and

$$\psi = \mathrm{sum}_{row}(\Gamma^\perp T) \quad (32)$$

the solution to the Lagrange optimisation problem (26) can be rewritten as:

$$SP = \lambda\psi \quad (33)$$

Hence, the optimal parameter vector P is given by:

$$P = \lambda S^{-1}\psi \quad (34)$$

The constraint function $g(P) = 0$ can be rewritten as:

$$P^T\psi - 1 = 0 \quad (35)$$

Incorporating (34) into (35):

$$\lambda \left(S^{-1}\psi\right)^T \psi - 1 = 0 \quad (36)$$

Hence, the Lagrange multiplier is given by:

$$\lambda = \left(\left(S^{-1}\psi\right)^T \psi \right)^{-1} \quad (37)$$

The algorithm for calculating the optimal vector W is summarised below:

(1) Select the order of the parity space $s \geq n$ and build matrices Γ, Q and T.
(2) Obtain $\Gamma^\perp$ (the left nullspace of Γ).
(3) Compute Σ using (24).
(4) Calculate the column vector ψ and the matrix S making use of (32) and (31), respectively.
(5) Obtain the Lagrange multiplier λ using (37).
(6) Calculate the parameter vector P by (34).
(7) Compute the vector W as:

$$W = P^T\Gamma^\perp \quad (38)$$

3. NUMERICAL EXAMPLE

Consider an exemplary system, described by (1), whose state space matrices are:

$$A = \begin{bmatrix} 0 & 0.765 \\ 1 & -0.050 \end{bmatrix} B = \begin{bmatrix} 0.005 \\ 0.5 \end{bmatrix} \quad G = \begin{bmatrix} 1.383 \\ 0.975 \end{bmatrix} \quad (39)$$
$$C = [\, 0 \ 2 \,] \qquad D = [0] \qquad H = [1]$$

The efficacy of the PE-UIO filter, designed for the system (39), for different cases of s and different cases of the input and output measurement noise variances is evaluated. Two efficiency indices are considered in order to assess the efficacy of the algorithms examined, namely the minimal

mean square error (MSE_{min}) and the estimation lag (EL). Consider a classical mean square error (MSE) index, where the true input $v(t)$ is delayed by i samples with respect to the estimated input $\hat{v}(t)$:

$$\text{MSE}(i) = \frac{\sum_t \left(\hat{v}(t) - v(t-i)\right)^2}{\sum_t v^2(t-i)} \tag{40}$$

The word *minimum* here refers to the the fact that the minimal value of MSE, as a function of the delay i, is considered (i.e. MSE_{min}). The delay, for which the function $\text{MSE}(i)$ achieves its minimum, is denoted EL (EL is the argument of $\text{MSE}(i)$, i.e. i =EL):

$$\begin{aligned} \text{EL} &= \arg\min_i \text{MSE}(i) \\ \text{MSE}_{min} &= \text{MSE}(\text{EL}) \end{aligned} \tag{41}$$

Hence, the EL indicates the number of samples by which the unknown input estimate is delayed with respect to the input. The MSE_{min} provides the accuracy measure of the unknown input estimate (delayed by the EL).

A Monte-Carlo simulation comprising of 100 runs has been carried out. Mean values of the MSE_{min} and the EL for each simulation setup are presented in Table 1. As expected, an increase in the parity space order results in a corresponding increase in the EL. For the first two cases of the measurement noise ($\sigma_{\tilde{u}}^2, \sigma_{\tilde{y}}^2$), i.e. ((0.1,2) and (1,1)) the MSE_{min} reduces as s increases from 3 to 6, however, a further increase of s degrades the efficacy of the PE-UIO (in terms of MSE_{min}).

The last row of Table 1 shows the efficacy of the Kalman filter-based minimum variance unbiased (MVU) state and input estimator proposed by Gillijns and De Moor (2007b). Due to the moving average filtering properties of the PE-UIO and that only a single signal (the unknown input) is estimated, the PE-UIO appears to be advantageous in comparison to the MVU approach in the case examined. Fig. 1 presents an exemplary visual illustration of the unknown input estimation using MVU and PE-UIO with $s = 6$.

Table 1. Simulation results ($\sigma_{\tilde{u}}^2, \sigma_{\tilde{y}}^2$)

$(\sigma_{\tilde{u}}^2, \sigma_{\tilde{y}}^2)$	(0.1,2)		(1,1)		(10,1)	
s	EL	MSE_{min}	EL	MSE_{min}	EL	MSE_{min}
3	0	2.454	0	2.764	0	13.639
4	0	2.268	0	2.210	0	6.469
6	1	1.572	1	1.422	1	3.012
8	2	2.496	2	2.086	2	2.668
MVU	0	19.4971	0	18.841	0	28.7843

4. CONCLUSIONS AND FURTHER WORK

A new approach for reconstructing the unknown input has been developed. The main advantage of the new scheme is that it does not rely on state estimation. Therefore, since the number of estimated signals is reduced, the estimation accuracy (in terms of the discrepancy between the true and the estimated input) is increased. Consequently, the proposed approach offers advantages in cases, when knowledge about the system states is not required. Due to the moving average filtering properties of the proposed scheme, the effect of measurement noise on both signals, i.e input and output, is minimised.

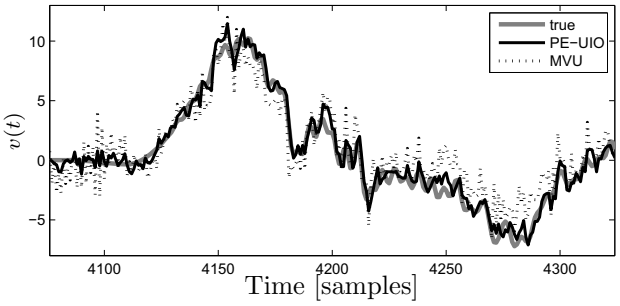


Fig. 1. Comparison of efficacy of PE-UIO and MVU ($\sigma_{\tilde{u}}^2 = 1$, $\sigma_{\tilde{y}}^2 = 1$)

Further work aims towards the development of the algorithm for multiple-input multiple-output systems. Consideration is also to be given to develop the algorithm for the more practical case of coloured measurement noise. Furthermore, the problem of choice of the order of the parity space depending on the noise variance remains open.

REFERENCES

Bertsekas, D.P. (1982). *Constrained Optimisation and Lagrange Multiplier Methods*. Academic press, Inc., London.

Darouach, M. and Zasadzinski, M. (1997). Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica*, 33(4), 717–719.

Ding, S.X. (2008). *Model-based Fault Diagnosis Techniques: Design Schemes, Algorithms and Tools*. Springer Verlag, Berlin.

Edelmayer, A. (2005). *Fault Detection in Dynamic Systems: From State Estimation to Direct Input Reconstruction Methods*. D.Sc. Thesis, Hungarian Academy of Sciences, Budapest.

Floquet, T. and Barbot, J. (2006). State and unknown input estimation for linear discrete-time systems. *Automatica*, 42(11), 1883–1889.

Gertler, J. (1991). Analitical redundancy methods in fault detection ond isolations. In *Procedings of the IFAC Symposium SAFEPROCESS*, 9–21. Baden-Baden.

Gillijns, S. and De Moor, B. (2007a). Unbiased minimum variance input and state estimation for linear discrete-time systems. *Automatica*, 43, 111–116.

Gillijns, S. and De Moor, B. (2007b). Unbiased minimum variance input and state estimation for linear discrete-time systems with direct feedtrough. *Automatica*, 43, 934–937.

Hsieh, C. (2000). Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Transactions on Automatic Control*, 45(12), 2374–2378.

Kirtikar, S., Palanthandalam-Madapusi, H., Zattoni, E., and Bernstein, D.S. (2009). l-delay input recontruction for discrete-time linear systems. In *Procedings of the Conferenceon Decision and Control*, 1848–1853. Shanghai, China.

Li, W. and Shah, S. (2002). Structured residual vector-based approach to sensor fault detection and isolation. *Journal of Process Control*, 12, 429–443.

Luenberger, D.G. (1964). Observing the state of linear systems. *IEEE Trans. Mil. Electr.*, MIL-8, 70–80.

Söderström (2007). Errors-in-variables methods in system identification. *Automatica*, 43(6), 939–958.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fault Tolerant Control Schemes for Nonlinear Models of Aircraft and Spacecraft: Preliminary Results

P. Castaldi * **N. Mimmo** * **S. Simani** **

* *Dipartimento di Elettronica, Informatica e Sistemistica, Università di Bologna, Facoltà di Ingegneria Aerospaziale, Via Fontanelle 40, 47100 Forlì(FC), ITALY (Ph/fax: +390543786943; e-mail: {paolo.castaldi,nicola.mimmo2}@unibo.it).*
** *Dipartimento di Ingegneria, Università di Ferrara, Via Saragat 1/E, 44100 Ferrara (FE), ITALY (Ph/fax: +390532974844; e-mail: silvio.simani@unife.it).*

Abstract: This paper explains the design method of an innovative active fault tolerant control scheme and the achieved results regarding its application to aerospace nonlinear models. The proposed method keeps the already in–place control and guidance laws and adds a feedback loop that accommodates the fault. The kernel of this active fault tolerant control consists in the fault detection and diagnosis module projected by using the non–linear geometric approach. Thanks to this approach fault estimate are analytically decoupled from both other faults and disturbances. The novel active fault tolerant control has been tested by using high fidelity simulators of aircraft and spacecraft systems and the performance show the method's robustness with respect to disturbance effects and measurement errors. The results obtained demonstrate how the proposed design methodology could be a successful approach for the reliable design of fault tolerant control schemes in real aircraft and spacecraft applications.

Keywords: Fault tolerant control, nonlinear geometric approach, fault diagnosis, disturbance de-coupling, aircraft and spacecraft systems.

1. INTRODUCTION

A conventional feedback control design for a complex system may result in an unsatisfactory performance, or even instability, in the event of malfunctions in actuators, sensors or other system components. This is particularly important for safety–critical systems, such as aircraft and spacecraft applications. To overcome these drawbacks, new approaches to control system design have been developed in order to tolerate component malfunctions, while maintaining desirable stability, and performance properties. These types of control systems are often known as Fault Tolerant Control (FTC) systems, which possess the ability to recover component faults automatically.

In general, methods for fault tolerant control systems are classified into two types, i.e. Passive Fault Tolerant Control Scheme (PFTCS), and Active Fault Tolerant Control Scheme (AFTCS). In a PFTCS, controllers are fixed, and designed to be robust against a class of presumed faults. In contrast to a PFTCS, an AFTCS reacts to the system component faults actively by reconfiguring control actions. An AFTCS relies heavily on realtime Fault Detection and Diagnosis (FDD) schemes, which are exploited for providing the most uptodate information about the true status of the system. Usually, the information coming from FDD schemes can be used from reconfiguration of logicbased switching controller or from a feedback of the fault estimate.

This paper illustrates a comprehensive novel methodology for Active Fault Tolerant Control Systems. The AFTCS is obtained by keeping the already in-place guidance and control (GC) laws and by adding a loop for feedback of the fault estimate. One of the main advantages of this strategy is that a structure of logicbased switching controllers is not required also avoiding the relative time delays required for the appropriate controller choice. It's worth noting that the design of the proposed FDD scheme and the design of the guidance and control (GC) scheme can be done independently. These features could significantly improve the applicability scope of these approaches since the modification of the validated and certified in-place nominal control law could be a major concern and especially for aerospace systems. Concerning the FDD procedure, this paper shows a novel nonlinear method based on the work of (Castaldi et al. (2010a)) that belong to the framework of Non Linear Geometric Approach (NLGA). The presented FDD scheme uses structurally robust adaptive filters (AF) which are analytically decoupled, thanks to NLGA, from disturbances as both non–controllable signals and other faults, resulting in this way, in an unbiased fault estimate (filters that aren't decoupled from disturbances shows an analytical bias in the fault estimate). The consequence is that the reliability of the overall AFTC system increases.

Generally a system can be affected by several categories of fault and by multiple faults per category. This work will treat the case of *single* faults, *i. e.* multiple faults that

occur one per time, modeled by additive step functions that can well represent a loss of efficiency (Falcoz et al. (2010); Marcos et al. (2010)).

The organization of the remainder of this paper is as follows. Section 2 provides a description of the NLGA–AF scheme, by highlighting the generic algorithm to obtain a new subsystem affected by the fault to be estimated and decoupled from the other faults and the disturbs. Also the adaptive filters structure and the resulting AFTCS strategy are shown in detail. Section 3 illustrates the implementation of the FDD module and the simulation results for two aerospace examples: unmanned aerial vehicle and spacecraft. Going through these examples are shown several mathematical technique suitable to solve applicative constraints: *weak* decoupling and banks of *combined* filters. In Section 4 the stability verification for the overall AFTC scheme is furnished by mean of MonteCarlo simulation. Concluding remarks are summarized in Section 5.

2. NLGA–AF METHODOLOGY AND AFTCS STRATEGY

This section describes the implementation of the FDD scheme and the structure of the AFTCS strategy.

Regarding the presented FDD scheme, it belongs to the NLGA framework, where a coordinate transformation, highlighting a sub–system affected by the fault and decoupled by the disturbances, is the starting point to design a set of adaptive filters. They are able to both detect additive fault acting on a single actuator and estimate the magnitude of the fault. It is worth observing that, by means of this NLGA approach, the fault estimate is decoupled from disturbance d. The proposed approach can be properly applied to the nonlinear affine model of the system in the form:

$$\begin{cases} \dot{x} = n(x) + g(x)\, c + \ell(x)\, f + p_d(x)\, d \\ y = h(x) \end{cases} \tag{1}$$

where the state vector $x \in \mathcal{X}$ (an open subset of $\mathbb{R}^{\ell_n}$), $c(t) \in \mathbb{R}^{\ell_c}$ is the control input vector, $f(t) \in \mathbb{R}^{\ell_f}$ is the fault vector (faults to be estimated), $d(t) \in \mathbb{R}^{\ell_d}$ the disturbance vector (embedding also the faults which have to be decoupled, in order to perform the fault isolation) and $y \in \mathbb{R}^{\ell_m}$ the output vector, whilst $n(x)$, $\ell(x)$, the columns of $g(x)$, and $p(x)$ are smooth vector fields, with $h(x)$ is a smooth map.

The design of the strategy for the diagnosis of the fault f with disturbance decoupling, by means of the considered NLGA, is shown in (De Persis and Isidori (2001)) and organized as follows:

- computation of Σ_*^P, *i.e.* the minimal conditioned invariant distribution containing P (where P is the distribution spanned by the columns of $p_d(x)$);
- computation of Ω^*, *i.e.* the maximal observability codistribution contained in $(\Sigma_*^P)^\perp$;
- if $\ell(x) \notin (\Omega^*)^\perp$, fault detectability condition, the fault is detectable and a suitable change of coordinate can be determined.

As mentioned above, the considered NLGA to the fault diagnosis problem, is based on a coordinate change in the state space and in the output space, $\Phi(x)$ and $\Psi(y)$, respectively. They consist in a surjection Ψ_1 and a function Φ_1 such that $\Omega^* \cap \operatorname{span}\{dh\} = \operatorname{span}\{d(\Psi_1 \circ h)\}$ and $\Omega^* = \operatorname{span}\{d\Phi_1\}$, where:

$$\begin{cases} \Phi(x) = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} \Phi_1(x) \\ H_2 h(x) \\ \Phi_3(x) \end{pmatrix} \\ \Psi(y) = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} = \begin{pmatrix} \Psi_1(y) \\ H_2 y \end{pmatrix} \end{cases} \tag{2}$$

are (local) diffeomorphisms, whilst H_2 is a selection matrix, *i.e.* its rows are a subset of the rows of the identity matrix. This transformation can be applied to the system (1) if and only if a the fault detectability condition is satisfied. The system (1) in the new reference frame can be decomposed into 3 subsystems, namely $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$, where the first one is always de–coupled from the disturbance vector and affected by the fault as follows:

$$\begin{cases} \dot{\bar{x}}_1 = n_1(\bar{x}_1, \bar{y}_2) + g_1(\bar{x}_1, \bar{y}_2)\, c + \ell_1(\bar{x}_1, \bar{y}_2, \bar{x}_3)\, f \\ \bar{y}_1 = h(\bar{x}_1) \end{cases} \tag{3}$$

where the variable $\bar{y}_2$ in (3) is assumed to be measured and considered as independent input.

With reference to (3), the NLGA–AF can be designed if the condition in (De Persis and Isidori (2001)) and the following new constraints are satisfied:

- the $\bar{x}_1$–subsystem is independent from the $\bar{x}_3$ state components;
- the single fault is a step function of the time; hence an element of vector f is a constant to be estimated;
- there exists a proper scalar component $\bar{x}_{1s}$ of the state vector $\bar{x}_1$ such that the corresponding scalar component of the output vector is $\bar{y}_{1s} = \bar{x}_{1s}$ and the following relation holds (Castaldi et al. (2007)):

$$\dot{\bar{y}}_{1s}(t) = M_1(t) \cdot f_s + M_2(t) \tag{4}$$

where $M_1(t) \neq 0, \forall t \geqslant 0$. Depending on te application (see the following examples), the term f_s can be read as a single scalar fault or a combination of single scalar faults weighted by nonlinear state functions. Moreover $M_1(t)$ and $M_2(t)$ can be computed for each time instant, since they are functions just of input and output measurements. The relation (4) describes the general form of the system under diagnosis. Under these conditions, the design of the adaptive filter is achieved, with reference to the system model (4), in order to provide a fault estimation $\hat{f}_s(t)$, which asymptotically converges to the magnitude of the fault f_s. Assume that the subsystem (4) is determined with the proposed NLGA procedure. Than f_s can be estimated by means of the following adaptive filter based on the least–squares algorithm with forgetting factor (Castaldi et al. (2010a)). The adaptation law is given by:

$$\begin{cases} \dot{P} = \beta\, P - \dfrac{1}{N^2} P^2 \breve{M}_1^2, & P(0) = P_0 > 0 \\ \dot{\hat{f}}_s = P\, \epsilon\, \breve{M}_1, & \hat{f}_s(0) = 0 \end{cases} \tag{5}$$

with the following equations representing the output estimation, and the corresponding normalised estimation error:

$$\begin{cases} \hat{\breve{y}}_{1s} &= \breve{M}_1 \hat{f}_s + \breve{M}_2 + \lambda \breve{\bar{y}}_{1s} \\ \epsilon &= \dfrac{1}{N^2}\left(\bar{y}_{1s} - \hat{\bar{y}}_{1s}\right) \end{cases} \tag{6}$$

where all the involved variables of the adaptive filter are scalar. In particular, $\lambda > 0$ is a parameter related to the bandwidth of the filter, $\beta \geqslant 0$ is the forgetting factor and $N^2 = 1 + \breve{M}_1^2$ is the normalisation factor of the least–squares algorithm. Moreover, the proposed adaptive filter adopts the signals $\breve{M}_1$, $\breve{M}_2$, $\breve{\bar{y}}_{1s}$ which are obtained by means of a low–pass filtering of the signals M_1, M_2, $\bar{y}_{1s}$ as follows:

$$\begin{cases} \dot{\breve{M}}_1 = -\lambda \breve{M}_1 + M_1, & \breve{M}_1(0) = 0 \\ \dot{\breve{M}}_2 = -\lambda \breve{M}_2 + M_2, & \breve{M}_2(0) = 0 \\ \dot{\breve{\bar{y}}}_{1s} = -\lambda \breve{\bar{y}}_{1s} + \bar{y}_{1s}, & \breve{\bar{y}}_{1s}(0) = 0 \end{cases} \tag{7}$$

Thus, the considered adaptive filter is described by the systems (5), (6), and (7). It can be proved that the asymptotic relation between the normalised output estimation error $\epsilon(t)$ and the fault estimation error $f_s - \hat{f}_s(t)$ is the following:

$$\lim_{t\to\infty} \epsilon(t) = \lim_{t\to\infty} \frac{\breve{M}_1(t)}{N^2(t)} \left(f_s - \hat{f}_s(t)\right) \tag{8}$$

Moreover, it can be proved that the adaptive filter described by the relations (5), (6), and (7) provides an estimation $\hat{f}_s(t)$ that asymptotically converges to the magnitude of the step fault f_s. The proofs are similar to those of (Castaldi et al. (2010a)) and have been omitted here.

With reference to Figure 1, u_r is the reference input, u is the actuated input, u_c is the controlled input, u_{GC} represents the output signal from the GC system, y is the measured output, f the actuator fault, whilst $\hat{f}$ is the estimated actuator fault. Therefore, Figure 1 shows that the AFTCS strategy is obtained by integrating the FDD module with the existing GC system. The FDD module consisting of the generalised bank of NLGA–AF provides the correct estimation $\hat{f}$ of the f actuator fault. This estimated signal is injected into the control loop, in order to compensate the effect of the actuator fault.

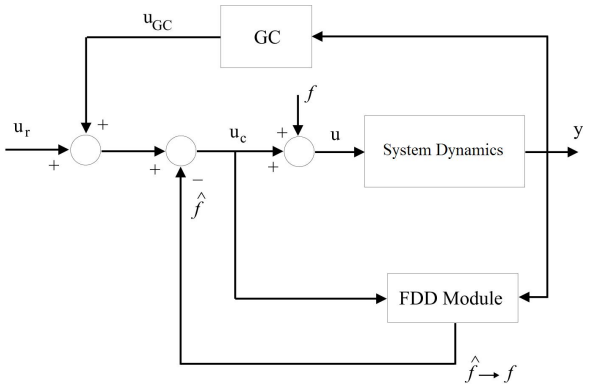


Fig. 1. Logic diagram of the integrated AFTCS strategy.

3. NLGA–AF METHODOLOGY APPLICATIONS

The NLGA methodology showed in Section 2 is directly applicable only to systems that are affine *w. r. t.* both inputs and disturbances. If an observable sub–system can be found one fault can be isolated from disturbances and other faults. Unfortunately a lot of nonlinear models of interest in the aerospace field don't match these requirements and, in particular, there are two main cases:

(1) The system is not affine *w. r. t.* inputs (*ex.* UAV);
(2) The system is affine *w. r. t.* both inputs and disturbances but it's impossible to isolate only one fault (*ex.* spacecraft);

In this section the NLGA–AF methodology is applied to three examples belonging to the enumerated categories and, for each of these, are showed the models and their peculiarity, the eventual hypothesis used in the models approximation, the NLGA application strategy, the adaptive filters obtained and the simulation results for the resulting global AFTCS.

3.1 Example 1: Unmanned Aerial Vehicle, UAV

The dynamic model of the fixed wing UAV assumed for this application is often used for control system design purposes (Boskovic et al. (2004)). The control inputs are the thrust, T, the load factor, n, and the bank angle, μ. There are two faults respectively on the thrust and the load factor. In this work there are not external or environment disturbances and the filters have to be decoupled only from the other faults. In this case the mathematical matter is that the load factor appears in the model with in a non–affine way while the thrust is an affine input. On the other hand it can be possible to obtain an estimation of fault acting on n independently from that on T. The strategy used to solve this problem (Castaldi et al. (2010b)) consists in a fall of two adaptive filters: the first one estimates the fault on n and the second on T. The adaptive filter on n communicates the fault estimate to the second filter which can considers known the quantities present in its equation and so it can provides the estimation of fault on T. Figure 2 depicts the logic of the modified FDD module. As shown in the following simulation results,

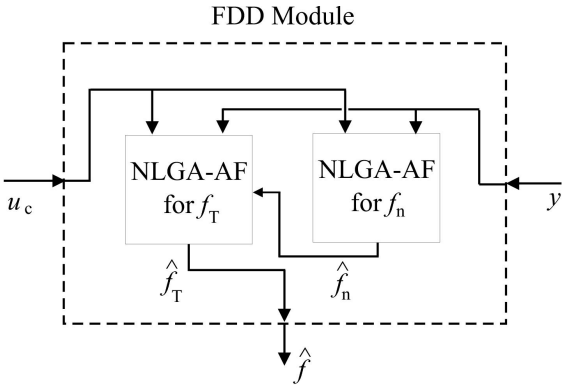


Fig. 2. FDD *weak* decoupling logic.

the fault decoupling is perfect in steady–state conditions, and negligible in transient conditions, if the NLGA–AF is designed to provide a prompt fault reconstruction.

With reference to Section 2, the design of the NLGA–AF described by Eq. (4) for the signal $f_s = f_n$, is based on the dynamic system in the form:

$$\begin{cases} \dot{\bar{y}}_{1s,n} = M_{1,n}\, f_n + M_{2,n} \\ M_{1,n} = \dfrac{g\,\cos\mu}{V} \\ M_{2,n} = \dfrac{g}{V}\,(n\,\cos\mu - \cos\gamma) \end{cases} \tag{9}$$

where V is the airspeed, g is the gravity acceleration and γ is the ramp angle. It is worth observing that $M_{1,n}(t) \neq 0$, $\forall\, t \geqslant 0$, since the bank angle, μ, has been kept far from the value of 90^o during the set of simulations performed in this work.

On the other hand, the NLGA–AF for the signal f_T has the form:

$$\begin{cases} \dot{\bar{y}}_{1s,T} = M_{1,T}\, f_T + M_{2,T} \\ M_{1,T} = \dfrac{g}{W} \\ M_{2,T} = g\left(\dfrac{T-D}{W} - \sin\gamma\right) \end{cases} \tag{10}$$

with the drag expression given by:

$$D = \frac{1}{2}\,\rho\, V^2\, S\, C_{D0} + 2\,k\,\frac{\left(n + \hat{f}_n\right)^2 W^2}{\rho\, V^2\, S} \tag{11}$$

where W is the UAV's weight, ρ is the air density, S is the wing reference area, C_{D0} it the parasite drag coefficient and k is the induced drag coefficent.

Figure 3 shows the estimate of the signal $\hat{f}_n$ (dashed line), when compared with the actual simulated fault f_n (continuous line). It's worth noting that the $\hat{f}_T$ asymptotically has zero mean value thanks to the *weak* decoupling: after the estimation transient of $\hat{f}_n$ the estimate on T take into account the f_n effect resulting in a unbiased estimate. Without the *weak* decoupling the detectability range on $\hat{f}_T$ will be greater than that shown and, more precisely, it will depend on f_n value resulting in an hardly usable filter.

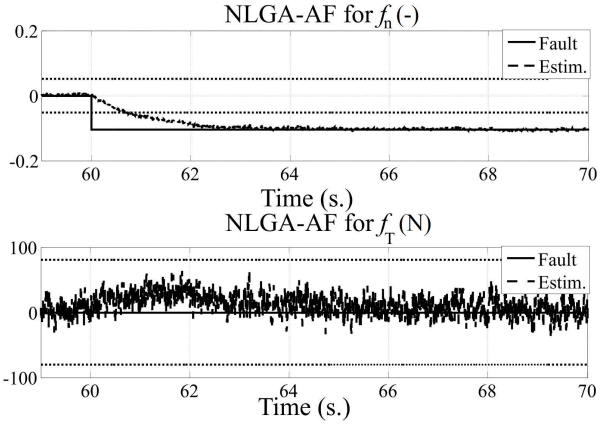


Fig. 3. Estimate $\hat{f}_n$ of f_n fault.

Figure 4 shows the estimate of the signal $\hat{f}_T$ (dashed line) and compares it to the actual simulated fault f_T (continuous line).

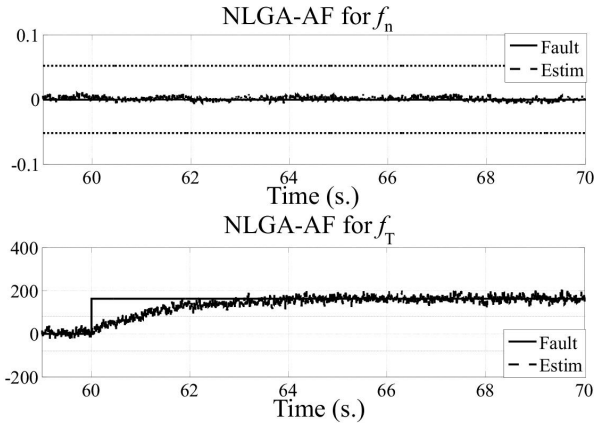


Fig. 4. Estimate $\hat{f}_T$ of f_T fault.

In particular, Figure 5 shows the airspeed signal V when the AFTCS recovers the fault (dashed line) and without fault accommodation (continuous line).

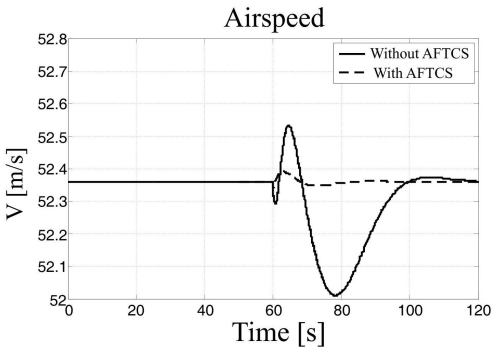


Fig. 5. Airspeed V with and without fault recovery.

Figure 6 shows the tracking error, one of the most meaningful aircraft variables that should assess the performances of the and AFTCS strategy, with (dashed line) and without (continuous line) fault recovery.

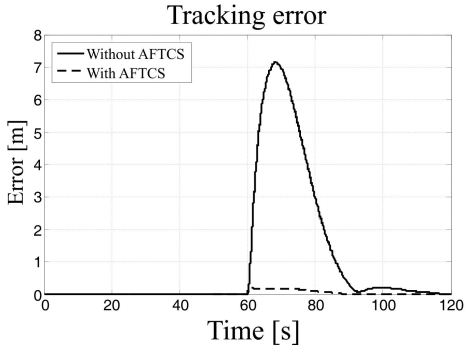


Fig. 6. Euclidean position error with and without recovery.

3.2 Example 2: Spacecraft

The spacecraft is considered to be a rigid body, whose attitude, described in Euler's angles, can be changed by using three actuators that generate the control momentums aligned with the three principal axis of inertia. In this case the faults are three momentums acting along the same direction of control torques. The disturbance from which the filters have to be decoupled is the gravitational

one. In this example the model is affine but it is impossible to obtain a new sub–system sensible only to one fault and decoupled from both disturbance and other faults. Then, the strategy consists in a bank of the three adaptive filters each of these sensible to all faults except one and decoupled from the gravitational disturbance. With a simple decision logic the fault detection, isolation and estimation can be accomplished. In order to design the NLGA–AF scheme estimating a single fault f_i decoupled from gravitational disturbance, it is possible to design three AF organized as a generalised scheme by means of the following procedure.

(1) Let $g_i(x)$, $i = 1, \ldots, 3$ (*i.e.* the columns of $g(x)$ in the model of Eq. (1)). Let $p_d(x)$ be the vector field related to the gravitational disturbances;
(2) Find a $\bar{x}_1$–subsystem sensitive, for example, to a combination of the faults f_1 and f_2, and decoupled from gravitational disturbances by defining: $p(x) = [\, g_3(x) \;\; p_d(x) \,]$ and then applying the NLGA procedure; it results

$$f_s = f_{12} \triangleq f_1 A_{12}(x) + f_2 B_{12}(x) \tag{12}$$

(3) Repeat the step 2 to determine other two $\bar{x}_1$–subsystems sensitive, respectively, to the fault couples $\{f_1, f_3\}$ and $\{f_2, f_3\}$;
(4) Organise these 3 filters as a generalised scheme : if, for example, the single fault f_1 is present, only the filter sensitive to the combination of $\{f_2, f_3\}$ has zero output, thus allowing fault isolation. Moreover, by exploiting the estimate $\hat{f}_{12}$, the fault f_1 estimate can be obtained by means of relation $\hat{f}_1 = \frac{\hat{f}_{12}}{A_{12}(x)}$.

The coordinate change necessary to select the $\bar{x}_1$–subsystem decoupled from gravitational disturbances, determined by means of the above described NLGA procedure for the case with $f = [f_1 \quad f_2]^T$ has the form:

$$\bar{x}_{1s} = \bar{x}_{11} = P \sin\theta + Q \frac{I_Z - I_Y}{I_X - I_Z} \frac{I_Y}{I_X} \cos\theta \sin\phi \tag{13}$$

The $\bar{x}_1$–subsystem sensitive to $f_{s_{12}}$ is given by (Baldi et al. (2010)):

$$\dot{\bar{y}}_{1s} = M_1 f_{12} + M_2 \tag{14}$$

where $M_1 = 1$ and

$$\begin{aligned} M_2 = {} & \frac{(I_y - I_z)}{I_x} R\,[Q \sin\theta + P \cos\theta \sin\phi] + \\ & [Q \cos\phi - R \sin\phi]\, P \cos\theta + \\ & + \frac{1}{I_x} \left[M_{x_C} \sin\theta + M_{y_C} \frac{I_z - I_y}{I_x - I_z} \cos\theta \sin\phi \right] + \\ & + Q\,[P + \tan\theta\,(Q \sin\phi + R \cos\phi)] \times \\ & \times \frac{I_Z - I_Y}{I_X - I_Z} \frac{I_Y}{I_X} \cos\theta \cos\phi \end{aligned}$$

Moreover, the model of Eq. (14) is completed by:

$$\begin{aligned} A_{12}(x) &= \sin\theta \\ B_{12}(x) &= \frac{I_z - I_y}{I_x - I_z} \cos\theta \sin\phi \end{aligned} \tag{15}$$

where $M_{\sharp C}$, $f_\sharp$, $I_\sharp$, ($\sharp = x, y, z$) are the control torque, fault torque and inertia on $\sharp$-axis, respectively. The rates of turn are respectivelly P, Q and R while the ϕ, θ, ψ are the attitude angles.

The global scheme with a generic nonlinear attitude control system is tested in noisy simulation. In the case of presence of fault f_2, figure 7 shows the estimates $\hat{f}_{12}$, $\hat{f}_{13}$ and $\hat{f}_{23}$ (now considered as *residuals*). Only the residuals decoupled from gravitational and fault on the second axis, *i.e.* f_{13}, have zero mean value after the fault occurrence, thus allowing the correct fault isolation. Hence, as described in this section, $\hat{f}_2(t)$ can be determined. Figure

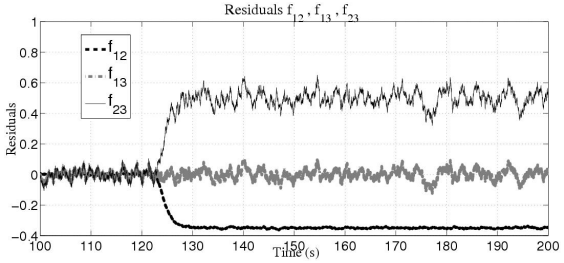


Fig. 7. Residuals.

8 shows the estimates of f_2 obtained with both the proposed gravitational decoupled filter (bold black line) and the the not decoupled filters (gray line), when compared with the actual simulated actuator fault (fine black line). With reference to the variation $\Delta\theta$ of the attitude angle

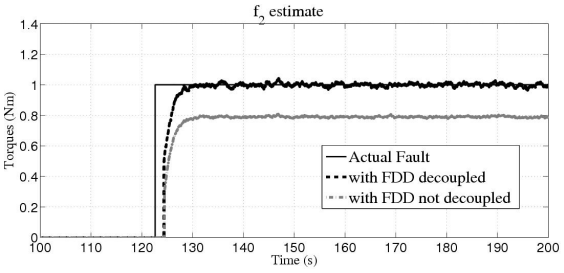


Fig. 8. Fault estimate $\hat{f}_2$.

θ with respect to the reference value, the fault recovery performances are obtained by using three different control strategies: without FTC, with FTC using not decoupled filters and, finally, with FTC using the proposed decoupled filters. As shown in Figures 9, the comparison highlights the better performances of the AFTCS relying on a FDD module decoupled from the gravitational field.

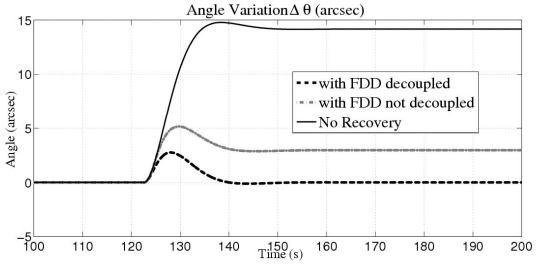


Fig. 9. Angle variation: $\Delta\theta$.

Figure 10 shows the fault estimates provided by an FDD module not decoupled from gravity disturbances, while

the attitude increases during stellar pointing. Due to increasing fault estimation errors, the corresponding AFTCS scheme is unable to recover the spacecraft's state to the correct value.

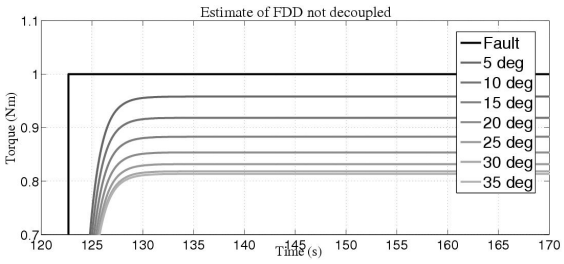


Fig. 10. Not decoupled filter estimates deteriorate as attitude increases.

4. STABILITY ANALYSIS

The stability properties of the overall AFTCS are checked by means of a MonteCarlo campaign based on high fidelity nonlinear simulators of the proposed aerospace systems. Initial state conditions are changed randomly and a fault affecting the system occurs during the transient related to the stability analysis. All simulations have been performed by considering noise signals modeled as a band limited white processes. As an example, Figure 11 shows that the UAV's system state variables $[V\,\gamma\,\chi]^T$ return to the equilibrium values, proving the overall system stability, and even the fault occurrence does not affect the stability properties.

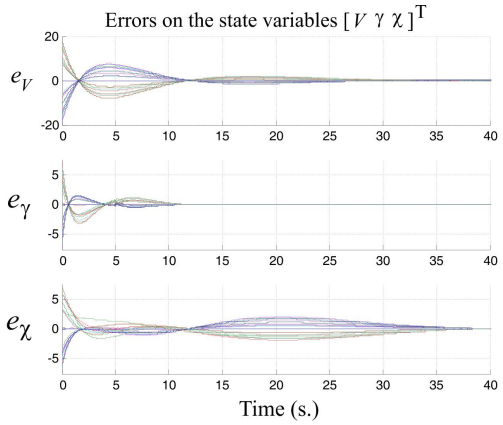


Fig. 11. Monitored state variables used for the stability analysis.

5. CONCLUSION

This paper proposes a new NLGA–AF methodology for Fault Tolerant Controls and shows its applications in the aerospace field. Starting from an existing system constituted by an in–place Guidance and Control System and its relative plant the illustrated methodology adds a feedback loop. In this loop there's a bank of adaptive filters which constituent equations are derived using the NonLinear Geometric Approach. The bank of filters uses the inputs and the outputs of the controlled plant and provides the estimation of the fault. The loop is closed by injecting the fault estimate in the input lines. The powerful of this approach stays in the analytic disturbances decoupling that generate a structurally robust FDD scheme and, consequently, a very reliable AFTC system. The overall scheme results in a Fault Tolerant Control System. The performance of the original GCS in presence of fault is substantially improved by the added loop both in the transient phases and asymptotically. Two different aerospace applications show the potentiality of this innovative nonlinear scheme. In the case of satellite, differently from other schemes already present in literature, for the first time, the fault estimates were decoupled from gravitational disturbances. Beyond a new solution was proposed for the UAV non–affine system by using a *weak* decoupling. The realistic adopted models of plants and the pragmatic mathematical hypothesis at the base of these applications were tested by an intensive campaign of lifelike noisy simulations. The results show that the FTC schemes produced are reliable and ready to physical implementations. Future works will take into consideration also other fault types depending on the considered application.

REFERENCES

Baldi, P., Castaldi, P., Mimmo, N., and Simani (2010). Fault diagnosis and control reconguration in earth satellite model engines. *UKACC International Conference on Control*.

Boskovic, J.D., Chen, L., and Mehra, R.K. (2004). Adaptive Control Design for Nonaffine Models Arising in Flight Control. *Journal of Guidance, Control & Dynamics*, 27(2), 209–217.

Castaldi, P., Geri, W., Bonfè, M., Simani, S., and Benini, M. (2010a). Design of residual generators and adaptive filters for the FDI of aircraft model sensors. *Control Engineering Practice*, 18(5), 449–459.

Castaldi, P., Geri, W., Simani, S., and Bonfè, M. (2007). Nonlinear Actuator Fault Detection and Isolation for a General Aviation Aircraft. *Space Technology – Space Engineering, Telecommunication, Systems Engineering and Control*, 27(2–3), 107–113. Special Issue on Automatic Control in Aerospace.

Castaldi, P., Mimmo, N., Simani, S., and Bonfè, M. (2010b). Active Fault Tolerant Control Scheme for a Nonlinear Model of a Fixed Wing Unmanned Aerial Vehicle. *International Journal of Applied Mathematics and Computer Science*. Submitted.

De Persis, C. and Isidori, A. (2001). A Geometric Approach to Nonlinear Fault Detection and Isolation. *IEEE Trans. on Automatic Control*, AC–46, 853–865.

Falcoz, A., Boquet, F., Dinh, M., Polle, B., Flandin, G., and Bornschlegl, E. (2010). Robust fault diagnosis strategies for spacecraft application to lisa pathfinder experiment. *ACA'10 – 18th IFAC Symposium on Automatic Control in Aerospace*.

Marcos, A., Kerr, M., and Peñín, L.F. (2010). Application of a fault accommodation approach to a re–entry vehicle. *ACA'10 – 18th IFAC Symposium on Automatic Control in Aerospace*.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Robust Model Matching for Geometric Fault Detection Filters: A Commercial Aircraft Example ⋆

József Bokor * **Peter Seiler** ** **Bálint Vanek** * **Gary J. Balas** **

* *Systems and Control Laboratory, Computer and Automation Research Institute, Hungarian Academy of Sciences (bokor@sztaki.hu).*
** *Aerospace and Engineering Mechanics Department, University of Minnesota (seiler@aem.umn.edu).*

Abstract: Geometric fault detection and isolation filters are known for having excellent fault isolation, fault reconstruction and sensitivity properties under small modeling uncertainty and noise. However they are assumed to be sensitive to model uncertainty and noise. This paper proposes a method to incorporate model uncertainty into the design. First, a geometric filter is designed on the nominal plant. Next a robust model matching problem is solved to design a filter that robustly matches the performance of the geometric filter over the set of uncertain plants. Several existing methods for robust filter synthesis are described to solve the robust model matching problem. It is then shown that the robust model matching problem has an interesting self-optimality property for multiplicative input uncertainty sets. Finally, an aircraft dynamics example is presented to detect and isolate aileron actuator faults to asses the performance of the geometric filter.

1. INTRODUCTION

Modern fly-by-wire aircraft flight control systems are becoming more complex with many actuators controlling several aerodynamic surfaces. While performance goals, like aerodynamic drag minimization and structural load suppression are becoming more and more important flight must be kept at the same highest safety level. In parallel, there is a clear trend towards the All-Electric Aircraft. Recently, Airbus introduced on the A380 a new hydraulics layout [Van den Bossche, 2006], where the three Hydraulics circuitry is replaced by a two Hydraulics plus two Electric layout, which saves one ton mass for the aircraft. Each primary surface has a single hydraulically powered actuator and electrically powered back-up with the exception of the outer aileron, which uses the two hydraulic systems together. Consequently, the trends of complexity and more-electric architectures, like Electromechanical Actuators (EMA) with more fault sources, raise the importance of availability, reliability and operating safety. For safety critical systems, like aircraft, the consequence of faults in the control system hardware and software can be extremely serious in terms of human mortality and economical impact. This is the reason why all aircraft manufacturers are compliant with stringent safety regulations of FAA, EASA and other aviation authorities. However, there is a growing need for on-line supervision and fault diagnosis to satisfy the newer societal imperatives towards an environmentally-friendlier aircraft with still the highest level of safety and reliability. The traditional approach to fault diagnosis in the wider application context is based on hardware redundancy methods which use multiple sensors, actuators computers and software to measure and control a particular variable [Goupil, 2009a]. Based on the mathematical model of the plant, analytical relation between different sensor outputs can be used to generate residual signals. There is a growing interest in methods which do not require additional hardware redundancy, and only rely on the ever increasing level of computational power onboard the aircraft. In analytical redundancy schemes, the resulting difference generated from the consistency checking of different variables is called as a residual signal.

The residual should be zero when the system is normal, and should diverge from zero when a fault occurs in the system. This zero and non-zero property of the residual is used to determine whether or not faults have occurred. Analytical redundancy makes use of a mathematical model and the goal is the determination of faults of a system from the comparison of available system measurements with a priori information represented by the mathematical model, through generation of residual quantities and their analysis. Various approaches have been applied to the residual generation problem, the parity space approach [Chow and Willsky, 1984], the multiple model method [Chang and Athans, 1978], detection filter design using geometric approach [Massoumnia, 1986], frequency domain concepts [Frank, 1990], unknown input observer concept

⋆ This material is based upon work supported by the National Science Foundation under Grant No. 0931931 entitled "CPS: Embedded Fault Detection for Low-Cost, Safety-Critical Systems". Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work is supported by the ADDSAFE (Advanced Fault Diagnosis for Safer Flight Guidance and Control) EU FP7 project, Grant Agreement: 233815, Coordinator: Dr. Andrés Marcos. This work is also supported by the Control Engineering Research Group of HAS at Budapest University of Technology and Economics. The authors are also thankful for Zoltán Szabó, for providing insight on geometric FDI methods.

[Chen and Patton, 1999], dynamic inversion based detection [Edelmayer et al., 2003], and using rational nullspace bases [Varga, 2003]. Most of these design approaches refer to linear time-invariant (LTI) systems. The geometric concept is further generalized to linear parameter-varying (LPV) systems by Balas et al. [2003], while input affine nonlinear systems are considered by De Persis et al. [2001].

The geometric design approach, for example, is known for its excellent fault isolation, fault reconstruction and sensitivity properties under small modeling uncertainty and noise. This paper proposes a method incorporate model uncertainty into the design. First, a geometric filter is designed on the nominal plant. Next a robust model matching problem is solved to design a filter that robustly matches the performance of the geometric filter over the set of uncertain plants. It is then shown that the robust model matching problem has an interesting self-optimality property for multiplicative input uncertainty sets. Specifically, the filter designed on the nominal plant is the optimal filter in the robust model matching problem. Finally, an aircraft aileron FDI example is detailed in the present article.

The importance of this paper is on the application (simulation) of the geometric approach based LTI FDI technique to a nonlinear high-fidelity aircraft, where issues of model uncertainty, realistic disturbances and robustness have to be accounted for in the design stage. The remainder of the paper is structured as follows. Section 2 formulates the robust fault detection filter design problem and describes the proposed solution method. The application example of a civil aircraft is described in Section 3. The method is applied to the high fidelity aircraft example, which demonstrates the proposed approach, given in Section 4. Finally, the paper is concluded in Section 5.

2. ROBUST MODEL MATCHING

This section considers a robust model matching problem for geometric filter design on uncertain plants. Then several existing methods for robust filter synthesis are described. The final subsection shows that the robust model matching problem has an interesting self-optimality property for multiplicative input uncertainty sets.

2.1 Problem Formulation

Let G_u denote an uncertain plant for which the filter will be designed. The standard linear fractional transformation (LFT) framework [Packard and Doyle, 1993, Zhou et al., 1996] can be used to model the uncertainties. Let $G \in \mathbb{RH}_\infty^{(n+k)\times(n+m)}$ and $\mathbf{\Delta} \subseteq \mathbb{RH}_\infty^{n\times n}$ be given.[1] Define the set of models

$$\mathcal{M} := \{G_u = F_u(G,\Delta) : \Delta \in \mathbf{\Delta},\ \|\Delta\|_\infty \leq 1\} \quad (1)$$

It is assumed that $F_u(G,\Delta)$ is well defined for all $\Delta \in \mathbf{\Delta}$ with $\|\Delta\|_\infty \leq 1$. $\mathbf{\Delta}$ is typically a set describing a block structured system that can include (repeated) real parametric and LTI dynamic system uncertainties. Nonlinear and/or time-varying uncertainties can also be modeled using integral quadratic constraints [Megretski and Rantzer, 1997]. The restriction that Δ be a square system is only for notational simplicity.

Each $G_u \in \mathcal{M}$ is a system that relates the faults and plant inputs to the signals available to the fault detection filter:

$$\begin{bmatrix} y \\ u \end{bmatrix} = G_u \begin{bmatrix} f \\ u \end{bmatrix} \quad (2)$$

The objective is to design a filter F with inputs $\left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ and output residuals r such that the residuals have "good" fault decoupling properties for all models $G_u \in \mathcal{M}$.

A robust model matching problem is now described to meet this objective. The nominal plant in the set $\mathcal{M}$ is given by $\Delta = 0$, i.e. $G_0 := F_u(G,0)$ is the nominal plant. First, design a geometric filter F_0 to solve the fundamental problem of residual generation on the nominal plant G_0. The model matching method attempts to design a filter F such that the performance on the uncertain plant G_u robustly matches the designed behavior of F_0G_0. Mathematically, the proposed design problem is:

Problem 1. Let $G \in \mathbb{RH}_\infty^{(n+k)\times(n+m)}$, $\mathbf{\Delta} \subseteq \mathbb{RH}_\infty^{n\times n}$ and $F_0 \in \mathbb{RH}_\infty^{l\times k}$ be given. The *robust model matching problem* is:

$$\min_{F\in\mathbb{RH}_\infty^{l\times k}} \max_{G_u\in\mathcal{M}} \|F_0G_0 - FG_u\|_\infty \quad (3)$$

The interconnection for this robust model matching problem is shown in shown in Figure 1. The reference model is given by F_0G_0. The nominal residual response r_0 will have the desired decoupling properties given by the fundamental problem of residual generation. The optimization in Equation 3 designs a filter F that most closely matches, in a worst-case sense, the desired residual generation behavior F_0G_0. In this paper the focus is on fault detection filters designed using the geometric approach but the model matching problem can, in principle, be used to robustly match the behavior of any filter F_0 designed on the nominal system G_0.

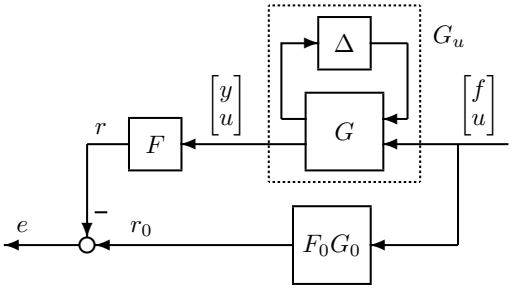


Fig. 1. Robust model matching.

2.2 Filter Synthesis

There are several approaches to solve the robust model matching problem. Sun and Packard observed that robust filter design (Equation 3) is an infinite-dimensional convex optimization in the filter [Sun and Packard, 2003]. They developed an algorithm to compute the globally optimal robust filter for the special case where $\mathbf{\Delta}$ only models repeated real uncertainties. It does not seem possible to

[1] G and F were used in the previous section to denote gain matrices in the geometric filter. In this section G and F will denote systems in the model matching design.

extend this algorithm to sets $\mathbf{\Delta}$ that include dynamic uncertainties, nonlinearities and/or time-varying operators.

The standard approach to handle more complicated uncertainty sets is to replace $\max_{G_u \in \mathcal{M}} \|F_0 G_0 - F G_u\|_\infty$ with an upper-bound. For example, when $\mathbf{\Delta}$ contains only LTI uncertainty the maximization over $\mathcal{M}$ can be replaced with the μ upper bound which involves a minimization over D scales [Dullerud and Paganini, 2000]. The design problem can then be recast as a μ-synthesis problem involving a search for the filter and the D scales. μ-synthesis is, in general, a nonconvex problem and the coordinate-wise D-K iteration has been applied to solve for the filter and uncertainty multipliers [Appleby et al., 1991]. The D-K iteration yields sub-optimal solutions but is a standard method to handle the nonconvexity that arises in robust control synthesis.

In robust filter design problem, the filter enters the design interconnection in an open loop (rather than a feedback) configuration and this structure can be exploited. There are two different approaches to convert the μ-synthesis problem into an infinite dimensional convex optimization problem ([Scherer and Köse, 2008] and [Seiler et al., 2011]). Both approaches use the more general IQC framework to model the uncertainty and obtain an upper bound on the worst-case performance. In [Scherer and Köse, 2008], the filter synthesis problem is converted into an infinite-dimensional (convex) semi-definite program (SDP) [Boyd et al., 1994]. The set of allowable IQC multipliers is infinite dimensional and a finite dimensional optimization is obtained by restricting the multipliers to be a combination of chosen basis functions. In [Seiler et al., 2011], the robust filter design problem is turned into a frequency-dependent, infinite dimensional linear matrix inequality (LMI) in the filter and multipliers. Next, a finite dimensional optimization is obtained by enforcing the frequency-dependent LMI on a dense frequency grid and restricting the filter to be a linear combination of chosen basis functions. The frequency-dependent IQC multipliers are allowed to be arbitrary functions on the frequency grid. To summarize, the two approaches use roughly dual methods to convert the robust filter design problem to a finite dimensional convex optimization: In [Seiler et al., 2011], basis functions are used for the filter but the multipliers (scalings) are allowed to be arbitrary functions on the frequency grid. In [Scherer and Köse, 2008] basis functions are chosen for the multipliers but the filter is allowed to be an arbitrary, linear system.

The various methods to solve the robust filter design problem have benefits and drawbacks in terms of computational complexity and ease of formulating the problem (e.g. picking basis functions for the filter or for the uncertainty scalings). The next section shows that the robust model matching problem has an interesting self-optimality property for multiplicative input uncertainty sets. Specifically, F_0 itself is the optimal filter for this uncertainty structure.

2.3 Multiplicative Input Uncertainty

This section considers the robust model matching problem for input multiplicative uncertainty. The uncertain system is given by $G_u := G_0(I + w\Delta)$ where $w \in \mathbb{RH}_\infty$ is a weight that specifies the level of uncertainty at each frequency by $|w(j\omega)|$. $|w(j\omega)| = 1$ corresponds to 100% input uncertainty at frequency ω and hence weights typically satisfy $\|w\|_\infty \leq 1$. Input multiplicative uncertainty is a commonly used uncertainty model because the effect of uncertainty can be quickly assessed by choosing simple weights w. For example, a reasonable uncertainty model is obtained by choosing w to be a first order system with small magnitude at low frequencies and magnitude close to one at high frequencies. Alternatively, the Matlab function `ucover` [Balas et al., 2010] can be used to compute a w so that the uncertainty set $\mathcal{M}$ contains a given, finite set of LTI systems. The weight can generally be chosen as a full matrix but the result in this section is restricted to weights of the form $w(s)I$.

The design interconnection for the robust model matching problem with input multiplicative uncertainty is shown in Figure 2. G_0 again denotes the nominal system and F_0 is a filter that has been designed to achieve some desired performance on the nominal plant. For this uncertainty structure the robust model matching problem can be equivalently stated as:

Problem 2. Let $F_0 \in \mathbb{RH}_\infty^{m\times n}$, $G \in \mathbb{RH}_\infty^{n\times k}$ and $w \in \mathbb{RH}_\infty$ be given. The *robust model matching problem* is:

$$\min_{F\in\mathbb{RH}_\infty^{m\times n}} \max_{\Delta\in\mathbb{RH}_\infty^{k\times k},\|\Delta\|_\infty\leq 1} \|F_0 G - FG(I + w\Delta)\|_\infty \quad (4)$$

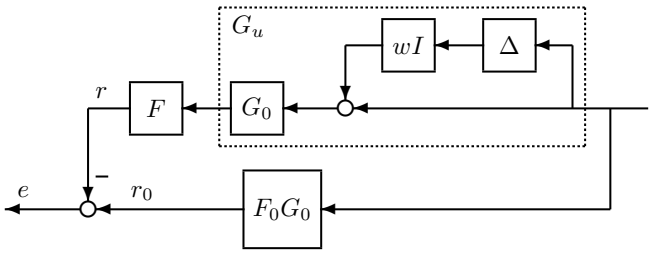


Fig. 2. Robust model matching with multiplicative input uncertainty.

The next theorem presents the main result of this section.

Theorem 3. If $\|w\|_\infty \leq 1$ then F_0 is the optimal filter for the robust model matching problem.

Proof 1. The robust model matching problem can be equivalently written as:

$$\min_{F\in\mathbb{RH}_\infty^{m\times n}} \max_{\omega} \max_{\substack{\Delta\in\mathbb{RH}_\infty^{k\times k}\\ |\Delta(j\omega)|\leq|w(j\omega)|}} \|\left(F_0 G - FG(I+\Delta)\right)(j\omega)\|$$

The min-max is always greater than the max-min and hence a lower bound on the model matching problem is obtained by:

$$\max_{\omega} \min_{F\in\mathbb{RH}_\infty^{m\times n}} \max_{\substack{\Delta\in\mathbb{RH}_\infty^{k\times k}\\ |\Delta(j\omega)|\leq|w(j\omega)|}} \|\left(F_0 G - FG(I+\Delta)\right)(j\omega)\| \quad (5)$$

Next, the constraints that F and Δ be stable are dropped:

$$\max_{\omega} \left[\min_{F\in\mathbb{C}^{m\times n}} \max_{\substack{\Delta\in\mathbb{C}^{k\times k}\\ |\Delta|\leq|w(j\omega)|}} \|(F_0 G)(j\omega) - FG(j\omega)(I+\Delta)\| \right] \quad (6)$$

The max over Δ is unchanged by dropping the stability constraint but the min over F is potentially lower once we drop the stability constraint. Thus the result of Equation 6 is no greater than the optimal value for Equation 5.

Next, apply Lemma from [Seiler et al., 2011] with $A := F_0(j\omega)$, $B := G(j\omega)$, and $\alpha := |w(j\omega)|$. By this lemma and the assumption $\|w\|_\infty \leq 1$, the optimization in the brackets of Equation 6 has an optimal cost equal to $|w(j\omega)|\|(F_0G)(j\omega)\|$ at each ω and the optimal value is achieved by $F = F_0(j\omega)$.

Thus the optimal cost for the robust model matching problem is lower bounded by $\|wF_0G\|_\infty$. This cost is achieved by the choice $F = F_0$ and hence F_0 is the optimal filter.

Roughly, this result implies that the robust model matching filter design is self optimal for this input multiplicative uncertainty set. The uncertainty degrades the performance but it does so in a way that apparently cannot be exploited by any other filter. Note that this result is not specific to nominal filters F_0 designed with the geometric method. The result only depends on the formulation of the robust model matching problem and the specific structure of the input multiplicative uncertainty.

3. AIRCRAFT MODEL

3.1 General Aircraft Characteristics

The aircraft model used in this paper is an aircraft from Airbus. The aircraft has two engines and a nominal weight of 200 tons. Some of its performance at cruise flight condition are speed of 240 knots, altitude of 30000 ft. The aircraft has 19 control inputs, and measurement of 6-DOF motion with load factor (n_x, n_y, n_z), body rate (p, q, r), velocity (V_T), aerodynamic angles (α, β), position (X, Y, Z) and attitude (ϕ, θ, ψ) outputs. The inputs are: $pi1$ left and $pi2$ right engine; AF (airbrake), which is disabled at cruise flight condition, $\delta_{a,IL}$ Aileron internal Left; $\delta_{a,IR}$ Aileron internal Right; $\delta_{a,EL}$ Ail external Left; $\delta_{a,ER}$ Ail external Right; $\delta_{sp,1L}$ Spoiler 1 Left; $\delta_{sp,1R}$ Spoiler 1R; Spoiler 23L; Spoiler 23R; Spoiler 45L; Spoiler 45R; $\delta_{sp,6L}$ Spoiler 6L; $\delta_{sp,6R}$ Spoiler 6R; $\delta_{e,L}$ Elevator Left; $\delta_{e,R}$ Elevator Right; δ_r Rudder; and δ_{ih} Trimmable Horizontal Stabilizer which is used for trimming purposes.

The aerodynamic database, propriety of Airbus Operations S.A.S, is of high-fidelity. The rigid body aircraft equations of motion are augmented with actuator [Goupil, 2009b] and sensor characteristics. The nonlinear body-axes rigid body dynamics includes 13 states using quaternion formalism: p, q, r body rates, u, v, w velocities all in body axes, q_0, q_1, q_2, q_3 quaternions, representing the rotation between the body and inertial axes, and X, Y, Z positions in the North-East-Down coordinate frame, assuming Flat Earth for simplicity.

3.2 Linearized Aircraft Model

In the present article one design point, cruise flight condition, is considered. The LTI model of the aircraft is obtained at level flight, with $p = q = r = 0$ rad/s, $v_x = 194.36$ m/s, $v_y = 0m/s$, $v_z = 15.13$ m/s, at 9144 m altitude, see Vanek et al. [2011] for details. The airbrake, which is disabled at high Mach numbers, is removed from the control inputs since it has no effect on the aircraft. Since the original aircraft model uses quaternions, which impose additional constraints on the state equations, the model used for trim and linearization is rewritten using conventional Euler angles [Stengel, 2004]. The model used for trim is an open-loop model without the control loop and, since the actuators and sensors are assumed to have unit steady state gain and low-pass characteristics, their dynamics are omitted. Trim is obtained with zero aileron, rudder and elevator deflection, left and right engines are providing the same amount of thrust to balance the yawing motion. Pitch axis trim is obtained with the Trimmable Horizontal Stabilizer, while the aircraft has 2.66 degrees Angle-of-attack. The resulting 12 state linear model is unstable.

The open loop aircraft model is slightly unstable around the yaw angle (ψ), and has two modes (X, Y) which are integrators. Since the FDI problem is invariant of X, Y positions and yaw angle these states are removed from the dynamics. The resulting model with nine states, as described in [Vanek et al., 2011], almost perfectly matches the original 12 states model in the behavior of the remaining states, and outputs. The resulting system with nine states is stable which is necessary for linear estimator based FDI techniques.

The resulting LTI model can be augmented with first order sensor and actuator dynamics derived from the high-fidelity simulation, to account for their effect on the aircraft behavior. Since the filters obtained by geometrical FDI methods require intense computation onboard the aircraft, only the pure rigid body dynamics model is used for filter synthesis here.

4. FDI FILTER DESIGN FOR THE AIRCRAFT

A geometric LTI FDI filter is designed for the aileron fault detection problem of the aircraft. First, the filter design steps are detailed and supported by linear analysis plots to show the optimality of the geometric filter. Detailed simulations on the high-fidelity aircraft model with injected aileron faults follows.

4.1 Filter Design Steps

The main idea behind the filter design formulation is that aileron faults appear on the filter residual output, while elevator and rudder faults are embedded in the unobservability subspace of the filter. For that reason the LTI model derived in Section 3.2 is augmented with left inner aileron, left elevator, and rudder faults, by using the successive input directions from the B and D matrices as fault directions in the linear model. Load factor, n_x, n_y, and n_z, measurement is omitted from the model, since the D matrix associated with these acceleration outputs is nonzero, which makes the geometric FDI synthesis more complicated. The resulting filter, using the methods developed in [Massoumnia, 1986], has 1 residual output, 27 inputs, and 7 states. Since perfect decoupling is possible, the transfer functions between elevator to residual and rudder

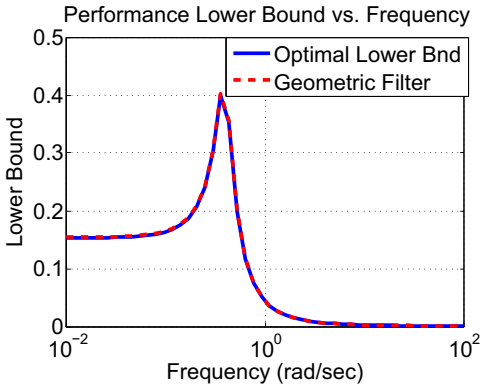


Fig. 3. Theoretical lower bound and achieved lower bounds of the FDI problem formulation with input multiplicative uncertainty.

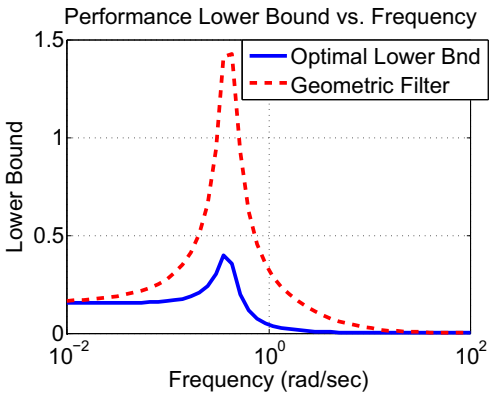


Fig. 4. Theoretical, and achieved lower bounds of the FDI problem formulation with input multiplicative uncertainty, augmented actuator and sensor models.

to residual are zero, while the residual have $0.394 rad/s$ time constant tracking response for aileron faults.

A lower bound on the optimal performance is computed using frequency-gridding method described in [Seiler et al., 2011], when the system is exposed to uncertainty. In the nominal case, with no uncertainty, the geometric filter is optimal for the decoupling, and according to Theorem 3 the filter is also optimal when input multiplicative uncertainty is considered. The effect of structured, input multiplicative uncertainty with the weights of $w_1 = \frac{2s+2}{s+60}$ on engines,$w_2 = \frac{2s+8}{1160}$ on spoilers, $w_3 = \frac{1.5s+3}{1120}$ on ailerons, elevators, and rudders , and $w_4 = \frac{14}{1160}$ on trimmable horizontal stabilizers are considered, with time constants comparable with the different actuator bandwidths. These weights corresponds to more than 100% uncertainty at high frequencies and 5% uncertainty at low frequencies on the input channels, and the block structure of the uncertainty Δ_a is grouped according to the actuator functional groups: $\Delta_a = diag < \Delta^{2\times2}_{engine}, \Delta^{4\times4}_{aileron}, \Delta^{8\times8}_{spoiler}, \Delta^{3\times3}_{longitudinal}, \Delta^{1\times1}_{rudder} >$.

The frequency grid consisted of 50 logarithmical spaced points between 0.01 and $100 rad/sec$. Figure 3 shows the lower bounds versus frequency. The dashed curve in Figure 3 shows the worst-case performance of F_0. The performance of F_0 degrades by approximately 41% over the uncertainty set, from perfect decoupling corresponding to 0 lower bound of the nominal case. The solid curve in Figure 3 shows the lower bound on the best achievable filter performance with uncertainty set included. The two curves are equal as expected based on Theorem 3. Thus F_0 is the optimal filter for robustly matching its own performance on the nominal plant. To further investigate the performance of the obtained filter, the uncertain LTI aircraft model is augmented with first order sensor and actuator models, on all input and output channels. Since the corresponding mathematical models are Airbus propriety, they are not discusses here. A lower bound calculation indicates in Figure 4 that the achievable performance is not significantly higher, compared with the actuator- and sensorless case, but the performance of the nominal filter is significantly lower than the achievable minimum. Due to these results, it is desirable to have actuator and sensor dynamics included in the filter design, which is not the

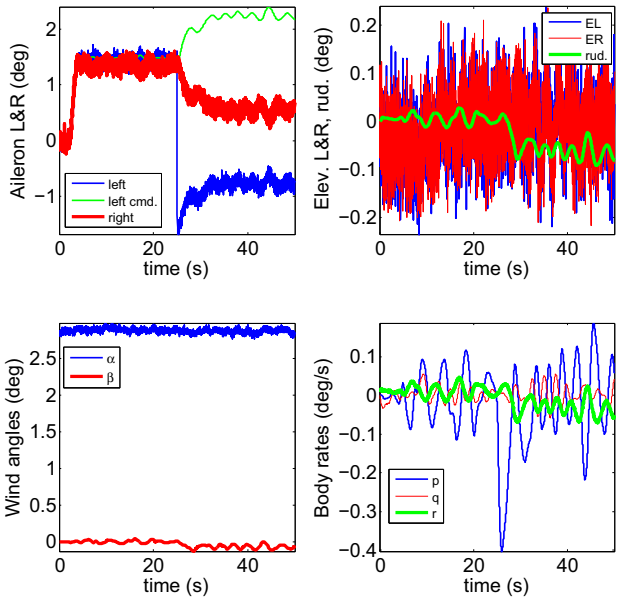


Fig. 5. Left aileron liquid jamming scenario, fault occurs at $25s$.

case here since computational complexity of those filters are significantly higher.

The filters are applied to the nonlinear aircraft model after taking the trim values into consideration, on both control input and sensor output signals. Since the simulation is implemented under SIMULINK with $0.01 sec$ fixed step size, the corresponding filters are also discretized with the same sampling time using bilinear transformation. It is also worth mentioning, that the simulation is in closed-loop with the flight control system set to altitude and heading hold mode and moderate atmospheric windgust disturbances are perturbing the aircraft flight.

The first fault scenario is left inboard aileron liquid jamming as seen on Figure 5, this means that a bias occurs on the rod sensor and the actuator shifts from it nominal $1.5 deg$ deflection to $-0.75 deg$ deflection and it remains $-2.25 deg$ apart from it commanded position. Figure 5 also shows the abrupt change in roll rate at $25 sec$ when the fault occurs, otherwise slight deflection can be seen on the

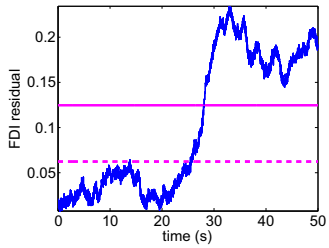


Fig. 6. Aileron liquid jamming, geometric FDI filter residual.

rudder but elevator and THS is unaffected, mainly the right aileron compensates the effect of the failure.

After investigation of fault free flight profiles, a detection threshold of 0.125 is selected. This corresponds to 100% margin over the largest observed residual signal with no fault. It is worth to note, that significantly lower detection threshold is achievable when the atmospheric windgust disturbances have lower level. Using the above mentioned threshold a detection time of 3.12 seconds is achieved, as shown on Figure 6, which is satisfactory since the level of fault only affects optimal aircraft configuration and is not critical to be detected instantaneously.

5. CONCLUSIONS

This paper considers the design of geometric fault detection filters and their application to a high fidelity aircraft model, and shows the advantages of advanced model-based methods, those are candidates for future industrial implementation. First, a geometric filter is designed on the nominal plant. Next a robust model matching problem is solved to design a filter that robustly matches the performance of the geometric filter over the set of uncertain plants. It is then shown that the robust model matching problem has an interesting self-optimality property for multiplicative input uncertainty sets. The proposed LTI filter is then applied to a high-fidelity aircraft model, where different aileron faults are successfully detected and when designed properly isolated from elevator and rudder faults in reasonable time. Further research should extend the validity of the present approach and based on the present findings provide a fault detection approach for a larger flight envelope.

REFERENCES

B. Appleby, J. Dowdle, and W. VanderVelde. Robust estimator design using μ synthesis. In *Proc. of the IEEE Conference on Decision and Control*, pages 640–645, 1991.

G. Balas, J. Bokor, and Z. Szabo. Invariant subspaces for LPV systems and their applications. *IEEE Transactions on Automatic Control*, 48(11):2065–2069, 2003.

G. Balas, R. Chiang, A. Packard, and M. Safonov. *Robust Control Toolbox*. MathWorks, 2010.

S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, 1994.

C. B. Chang and M. Athans. State estimation for discrete systems with switching parameters. *IEEE Transactions on Aerospace and Electronic Systems*, 14:418–425, 1978.

J. Chen and R. J. Patton. *Robust Model-based Fault Diagnosis for Dynamic Systems*. Kluwer Academic. Boston., 1999.

E.Y. Chow and A.S. Willsky. Analytical redundancy and the design of robust failure detection systems. *IEEE Trans. on Automatic Control*, 29(7):603–614, 1984.

C. De Persis, R. De Santis, and A. Isidori. Nonlinear actuator fault detection and isolation for a VTOL aircraft. In *Proceedings of the 2001 American Control Conference, Vols 1-6*, pages 4449–4454, 2001.

G. Dullerud and F. Paganini. *A course in robust control theory: A convex approach*. Springer Verlag, 2000.

A. Edelmayer, J. Bokor, and Z. Szabo. A geometric view on inversion-based detection filter design in nonlinear systems. In *Proceedings of the 5th IFAC symposium on fault detection, supervision and safety of technical processes. SAFEPROCESS 2003, Washington*, pages 783–788, Washington, 2003.

P.M Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - a survey and some new results. *Automatica*, 26:459–474, 1990.

P. Goupil. Airbus state of the art and practices on fdi and ftc. In *Proceedings of Safeprocess'09*, 2009a.

Philippe Goupil. Oscillatory failure case detection in the a380 electrical flight control system by analytical redundancy. *Control Engineering Practice*, 18:1110–1119, 2009b.

M.A. Massoumnia. A geometric approach to the synthesys of failure detection filters. *IEEE Transactions on Automatic Control*, 31:839–846, 1986.

A. Megretski and A. Rantzer. System analysis via integral quadratic constraints. *IEEE Trans. on Automatic Control*, 42(6):819–830, 1997.

A. Packard and J. Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

C.W. Scherer and I.E. Köse. Robustness with dynamic IQCs: An exact state-space characterization of nominal stability with applications to robust estimation. *Automatica*, 44:1666–1675, 2008.

P. Seiler, B. Vanek, J. Bokor, and G.J. Balas. Robust H_∞ filter design using frequency gridding. In *Submitted to the 2011 American Control Conference*, 2011.

Robert F. Stengel. *Flight Dynamics*. Princeton University Press, 2004.

K. Sun and A. Packard. Optimal, worst-case filter design via convex optimization. In *Proc. of the IEEE Conference on Decision and Control*, pages 1380–1385, 2003.

D. Van den Bossche. The a380 flight control electrohydrostatic actuators, achievements and lessons learnt. In *Proc. 25th Congress of the International Council of the Aeronautical Sciences*, 2006.

B. Vanek, P. Seiler, J. Bokor, and G.J. Balas. Robust fault detection filter design for commercial aircraft. In *Euro GNC 2011 1st CEAS Specialist Conference on Guidance, Navigation & Control, Münich*, 2011.

A. Varga. On computing least order fault detectors using rational nullspace bases. In *In Proceedings of the IFAC Symp. SAFEPROCESS'2003, Washington D.C.*, 2003.

Kemin Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, 1996.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

System Programmable Logic Controller Computer Aided Development Procedure

S. Chiesa*. S. Corpino. G. Medici***.**

** Politecnico di Torino Dipartimento di Ingegneria Aeronautica e Spaziale, Torino, 10129, Italy (Tel: 0115646818; e-mail: sergio.chiesa@polito.it) / S.P.A.I.C. srl, Torino, 10129,*
*** Politecnico di Torino Dipartimento di Ingegneria Aeronautica e Spaziale, Torino, 10129, Italy (Tel: 0110906867; e-mail: sabrina.corpino@polito.it) / S.P.A.I.C. srl, Torino, 10129,*
**** Politecnico di Torino Dipartimento di Ingegneria Aeronautica e Spaziale, Torino, 10129, Italy (Tel: 0115646807; e-mail: giovanni.medici@polito.it)*

Abstract: This paper describes the modeling procedure of complex, multisystem Programmable Logic Controller (PLC), using a general purpose, object-oriented, simulation tool. The procedure is a ladder-type, three step development method. It helps the specialists throughout the whole system requirement definition, allowing them to produce a more reliable and better defined Functional Requirement Document (FRD). The structure of the logic controller is defined during the first step. The system specialists use a fast, customizable, flow chart-based, visual software. Once the main control law core has been defined, the development team switches to the second step. During this phase each control law branch is developed and refined, using ANSI C programming language. In the last step the structured and refined code is implemented inside the so called status model. The status model is a Matlab Simulink-based model which uses an embedded Matlab function to model the Flight Control Computer (FCC). This core is linked (but can not interact) with sensors and general inputs for the system. Through the status model the user can simulate the system reacting in real time to its inputs. Faulted system behavior can also be simulated by introducing inconsistent inputs or by neglecting them. The aim of this tool is to test and validate the control laws dynamically, helping the specialists to debug the FRD. The development team can verify their control laws and test them in various failure scenarios. The FRD debug enables the development team to produce a more reliable and effective FRD, while enhancing their awareness of the entire system they are designing. By means of an adequate FRD the software design cycle is more reliable and communication between specialists and software developers is more effective.

Keywords: Simulation; UAV - Unmanned Aerial Vehicles; Computer Aided Control System Design; Flowcharts; Systems Architecture; Programmable Logic Controller.

1. INTRODUCTION

In the last twenty years the use of Programmable Logic Controller (PLC) in industry has seen a dramatic growth. In almost any application, when designing a system or component, the logic controller must be able to satisfy the requirements with an adequate degree of reliability. When a complex system is designed, each sub-systems' logic controller must be created. In this process there are two main actors: the specialist team (a team that knows the system and defines the requirements it must satisfy), and the software development team (a team that develops the software, using the information provided by the specialists).

There are a number of published software products (such as IBM Telelogic Rhapsody) that help both actors throughout the system development. They are based on a scripting language called SYS-UML that can lead to an automatic requirement matrix definition starting from the system design. These products contain some interesting tools such as a system input output simulator, black boxes and requirement list automatic import / export functions. These platforms usually represent the best choice for the development of complex systems, but are usually quite expensive and need adequate training to be used effectively. As a result these tools are usually more suitable for use in larger companies, making them simply out of the range of small businesses.

Specialists need a flexible, modular environment that allows them to rapidly perform any necessary changes when defining the requirements.

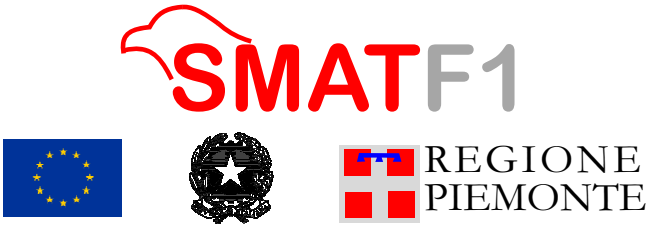


Fig. 1. Main Funders and Project Logo.

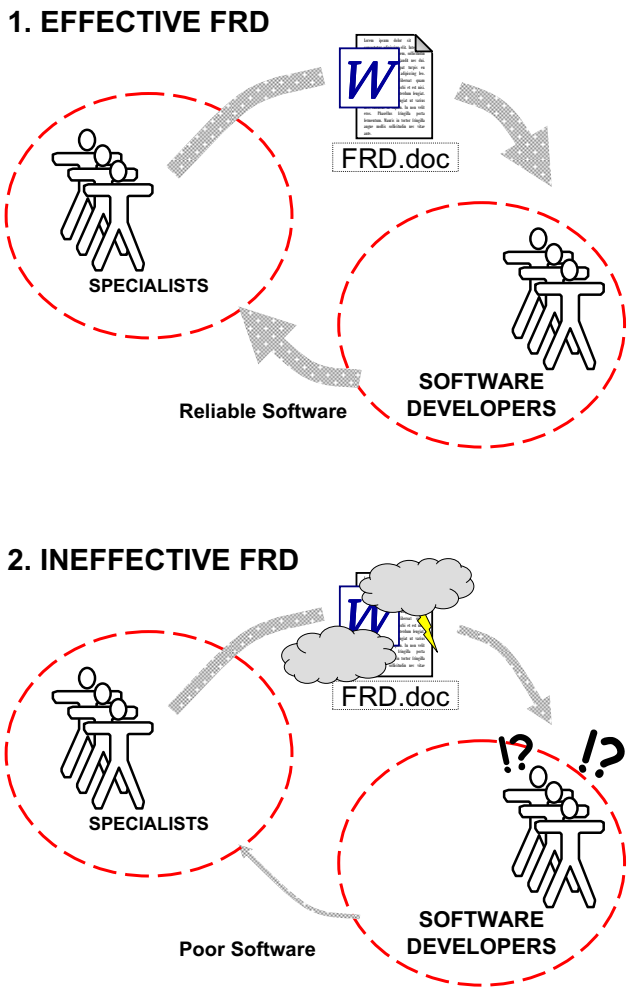


Fig. 2. Functional Requirement Document: Effectiveness and Reliability.

Once the system specialists evaluate the logic controller laws to be completed, a document is produced. Such documents (Functional Requirement Document or FRD) contain, in a brief, simple format, all the functional requirements the system should satisfy. The FRD is usually then sent to the software development team (or enterprise).

Writing an effective FRD is not an easy task. The specialists may forget crucial information or "obvious" requirements that the software development team (or contractor enterprise) may ignore.

Poor communication or a lack of a specific requirement may lead to an unnecessary or malfunctioning logic controller behaviour (Fig. 2). This would cause delays in the development time schedule and the need to adjust the software via multiple issue updates, which can lead to an increase in costs.

The authors developed the method described in this paper while cooperating with the SMAT-F1 project.

SMAT-F1 ("Sistema di Monitoraggio Avanzato del Territorio", the Italian translation of Advanced Environment Monitoring System) project is funded by "Regione Piemonte", managed by Finpiemonte and promoted through the Promoter Board of Piedmont's Aerospace District ("Comitato Promotore Distretto Aerospaziale Piemonte"). It is also co-funded by the European Regional Development Fund (ERDF), within the Regional Operative Program 2007/2013. This project is currently under development and has now reached the first phase (named SMAT-F1). Fig. 1 provides an overview of the main funders and project logos.

SMAT-F1 is currently under development under the direction of ALENIA AERONAUTICA S.p.a. and with the participation of other large companies (such as SELEX GALILEO), Research Institutes, and several SMEs. This project has received a significant contribution from a Research team of "Politecnico di Torino" and one of the SMEs involved in the Project, SPAIC S.r.l., a "Spin-off" company of Politecnico di Torino.

SMAT-F1's aim is to define, design and develop an Advanced Environment Monitoring System, based on Unmanned Air Systems (UAS). Within the UAS, Unmanned Aerial Vehicles (UAV) and their Ground Control Stations (GCS) are coordinated and managed by a SSC (Supervision and Coordination Station); a summarized conceptual overview of SMAT system is shown in Fig. 3.

The analysis of innovative solutions to improve performances of future UAV is a remarkable activity in the SMAT project. Some of these innovations have already been tested on some Alenia Aeronautica Technological Demonstrators.

Both technology and knowledge issues have been widely tested during the development of the SKY-Y, one of Alenia Aeronautica's technological demonstrators.

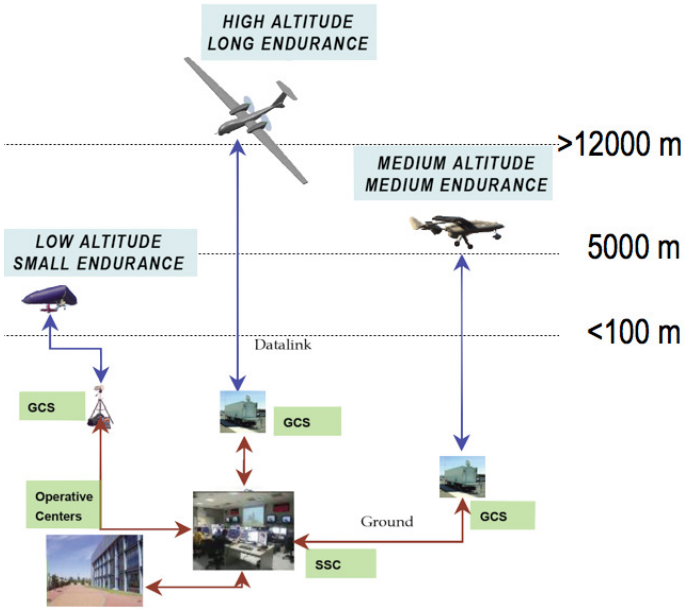


Fig 3. The SMAT System, an overview.

For example the High Altitude Long Endurance UAV (which flies at a higher altitude to Fig. 3, more than 12000 m), is not yet available, but will be developed starting from the SKY-Y (Table 1 and Fig. 4 collect some technical data and the design layout of the SKY-Y UAV).

The performances of such a demonstrator are reduced, if referred to the envisaged performances of the final product. Nevertheless studies for a derivative UAV are on going. The authors have been working on these improvements, in particular in Under-carriage Software Control Logics development.

Table 1. SKY-Y Technical Data

Dimensions	
Length	9.725 [m]
Span	9.937 [m]
Wing Area	10.785 [m^2]

Weights	
MTOW	1200 [kg]
OEW	800 [kg]
Max Fuel	250 [kg]
Max Payload	150 [kg]

Performances	
LOS Radius	185 [km]
Max Range	925 [km]
Altitude	>7600 [m]
Endurance	14 [h]

Propulsion
One DieselJet TDA 1.9 JTD 8 Valve Diesel Aviation Engine (Automotive derivative)

Payloads
EO/IR Sensor
Hyper-Spectral Sensor
Syntetic Aperture Radar
ESM/Elint

Fig. 4. Sky-Y: Design Layout.

During the development of SMAT-F1 project the authors and Alenia Aeronautica staff have been requested to define the control laws for a new Landing Gear system. This task has been performed by AeroSpace System Engineering Team (ASSET) of Politecnico di Torino, with the cooperation of S.P.A.I.C. S.r.l.

After just a few meetings it was clear that the procedure adopted to develop the control laws needed to be both structured and reproducible. This consideration led to an analysis of various methods. The authors have tested and refined them. The objective of the present paper is to discuss the procedure that was finally chosen in order to produce a reliable and effective FRD.

2. PROCEDURE

The method, developed and tested during the logic controller definition performed for the SMAT-F1 project, has three main steps. Firstly the structure of the algorithm (controller logics) is defined in general terms, by means of a fast customizable and easy to modify visual software. At the next stage the structure and sub-function development is refined, by means of a heavy duty programming language, ANSI C. Finally the code is implemented inside the so-called status model. In this way the controller logics can be tested and validated, while interacting with other systems.

2.1 Flowcharts

In the first phase of control logics design, the most effective technique is a mix of specialist knowledge, brainstorming, and troubleshooting. A scratch book may be a nice tool to use, but requires a “post-processing” phase to refine and comment on all the drafts. In order to avoid this “post-processing” phase another tool was used.

One of the most effective ways to express an idea or algorithm concept is to write it down by means of the flowchart language. The flow chart language is a diagrammatic representation, which gives a step-by-step solution to a given problem.

This graphic language can be enriched with colours or notes. Its fast and easy to adjust nature makes it a nice way to summarize the algorithm structure. Its visual grammar is well known (and there is a standard normative which describes it: ISO 5807:1985 Information processing – Documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts) and used in many applications. The flowcharts approach was used during the meetings and troubleshooting sessions.

There are many commercial products available for producing flowcharts, but even Small Office Home Office (SOHO) software (commercially available such as Microsoft PowerPoint or Apple Keynote or freely available such as Open Office or Google Documents) can do the job quite well. The more detail one wants to describe using the flowchart

technique, the more structured (but less effective and easy to read) plot he will obtain.

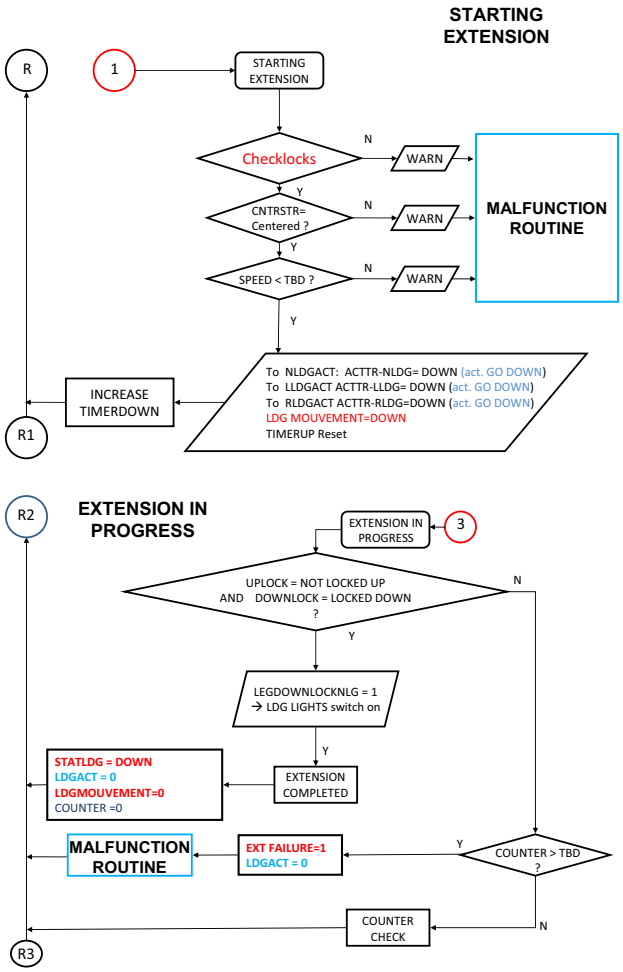


Fig. 5. Flowcharts: Some Examples.

Through the grammar of flowcharts it is possible to reproduce a wide range of conditional statements; from the basic single condition "if" lemmas, to more complex ones, such as iterative ("for", "while"), or time depending statements. This versatile visual language has proved to be the optimal choice in order to produce fast, easy to understand algorithm atoms.

There is a threshold beyond which this language becomes less and less effective; so the flowchart technique is better used to define the general algorithm structure, leaving the detail to a more focused tool. For this reason, once the algorithm structure has been defined, the specialists should freeze its configuration. At this point the specialists switch to the second step. The same flowcharts may be attached (with a brief description) as an Annex of the FRD document, giving the software developers a visual interpretation of the initial requirements list. They are still free to develop the software without any suggested path to follow. Few flowcharts and a simplified overview on a generic Landing Gear System have been reported in (Fig. 5 and Fig. 6).

Both flowcharts show a sub–function routine with a number of conditional checks. If every check is passed successfully the normal routine continues (and the required command sequence is performed). When receiving unexpected variables values, it invokes a "Malfunction Routine", which handles those events. The simplified Landing Gear system extension sequence starts from a couple of before-extension checks (on the Up-Locks sensors, on the angle of the nose wheel steering, and speed) and continues with the go down command to the actuators (Power Relays Arm so that actuators would be powered). The extension continues (extension in progress flowchart) until a count down timer expires or Down-Lock sensors report the landing gear to be in the Locked Down status. The components involved in the extension sequence are described in Fig. 6. When the Pilot commands the Landing gear to extend (this command is not shown), the actuator moves thanks to the powered Power relays. When Down-Lock relays reports Locked-Down Position it switches off the electrical actuator, and the extension sequence is complete.

2.2 Programming Language

Once the specialists freeze the algorithm's general configuration, the in-depth development of the whole main and sub-routines should take place. As discussed previously the flowchart approach in this case is less effective. A structured high level programming language (such as ANSI C, FORTRAN, ADA…) helps the specialists to develop and refine every detail of the algorithm. Every high level programming language is quite easy to use, and its English-like grammar helps the specialists during this development phase. The purpose of this stage is to produce a detailed and structured description of the control laws' behaviour. For this reason it is not required to run the code (and to build an executable file). During this phase the specialists may find some critical routines or adjust some key points of the algorithm. Each one of this event should be reported and described briefly, to be added in the FRD document.

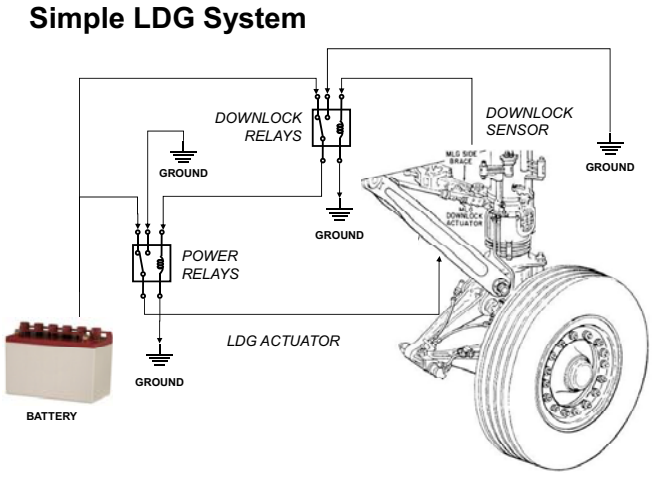


Fig. 6. Simplified Landing Gear (extension) System:

The code may use a lot of comments and logs in order to be easy to read and adjusted in future issues.

Even under those statements this phase still adds workload to the specialists. By the way the code that will be developed in this phase will be re-used during the next step. For example, during the cooperation between S.P.A.I.C. srl and Alenia Aeronautica, some 1800 lines of code were developed; the most of them has been then used in the status model.

2.3 Status model

The status model is a Matlab Simulink based model. The main idea is to create an embedded Matlab Function Block, the core, with the controller logics implemented as they have been defined in the previous steps. Each input of the system is then modelled outside (but linked with) the Matlab Embedded Function Block. The specialists using this approach can test (through the Simulink simulation) the system response to various inputs. Matlab Simulink has got a huge block library, which contains fully scriptable switches, relays, transfer functions blocks and much more. Those items can be used "out of the box", in order to model some simple subsystem components. Block oriented approach offers some features valuable for building customized systems. The basic object is the component: a reusable, self-contained entity that requires inputs and produces outputs. It is also possible to group and reuse customized components or even subsystems. This allows the specialists to build up a customized library which satisfy their needs. A component can be accessed (by other components) only through its inputs (external interface). Consequently, its internal (private) data and methods are hidden (encapsulated) from the other components (or Matlab Embedded Function Block). Even the icon of a component may be customized in order to recall the real component shape. The Block Oriented Approach is described in (Fig. 7). Once the model is complete, the specialists can test and validate their controller logics. By means of input link cuts or weird input behavior, the response to various sensors (input) failures can be verified.

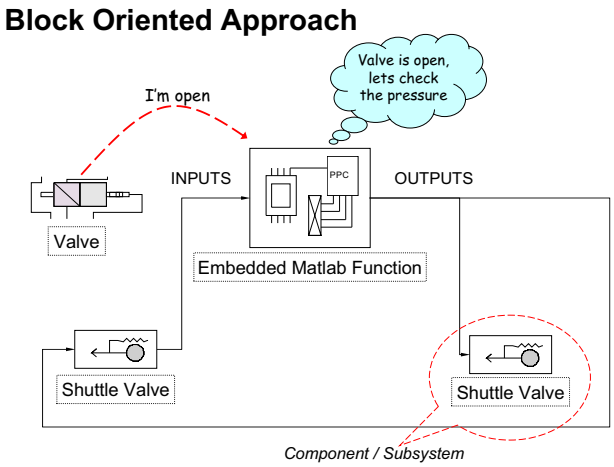


Fig. 7. Status model: Block Oriented Approach.

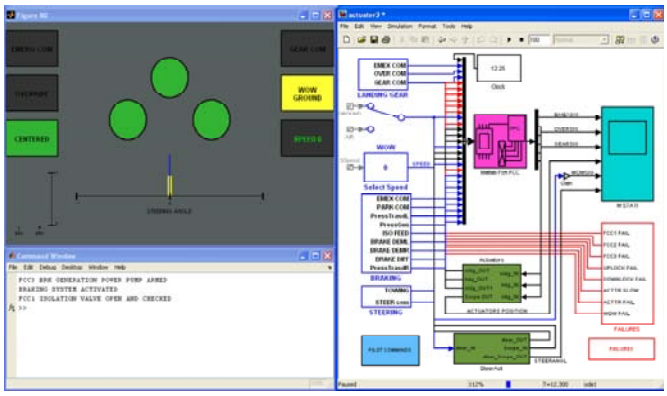


Fig. 8. Status model: Interface Example.

The modular nature of the block oriented approach allows the specialists to build an integrated (multi system) model. This feature enables validation and test of system behaviour under the effect of multiple subsystem inputs (and/or failures). The specialists may gather a collection of checklists and perform them before writing the FRD. Tests on "Single Failure Proof" or "Double Failure Proof" systems may also benefit of the status model.

The workload of the system logic controller development team will remain almost unchanged, and the effort required can be overcome by means of practice and the creation of a custom library.

An example of a multiple system status model is shown in (Fig. 8). The picture shows a possible arrangement of the status model interface. On the right half screen there is the main Simulink Model, containing the Matlab Core Function, all the inputs (and failures) the operator may invoke, and the various sub-systems models. On the upperleft side of the screen an indicator window is used, during the simulation, to let the operator know the actual system configuration. The lower left side of the screen is dedicated to the log window, this prompt is useful to produce all the text warning, status information, and time counters, that the operator should know about.

This interface shows all the information the operator receives and the actions he may invoke during the simulation ("online" actions and information).

Matlab Simulink has many useful tools that can be used to collect, summarize and save the most important informations (or variables state), in a time-history format, so that post processing analysis can be performed ("offline" actions and information).

The landing gear extension's sequence, described in the previous section is reported in (Fig. 9) using the "offline" time-history representation. The plot contains two crucial variables: the gear command signal status (a discrete command with a few milliseconds of consolidation time that

is produced when the pilot execute the gear down command), and the normalized actuators position. The plot shows that when the gear down command is received and consolidated, the extension sequence starts. The actuators (which have been modeled using a second order transfer function), change their state from UP to DOWN. The black dashed line indicates the uplock / downlock sensors activation threshold.

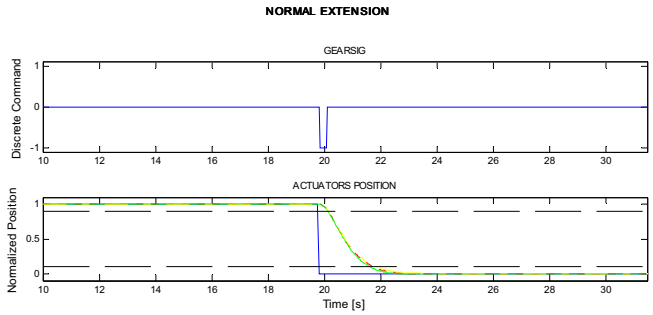


Fig. 9. Status model Plot Output Example. Variable's values are presented in blue, while actuators modelized position in green, yellow, and red.

The data collection and plot is a crucial tool both for the specialists and the software development team. The specialists may prepare a test cases database, run each test case, and use the "offline", time-history results of the most important sensors status (or subsystem position, speed, accelerations…) to verify the system behaviour. During this tests the specialists produce plots of the desired event sequence they expect from the system (a requirement representation). Using this method it is possible to produce plots that define only how the system is supposed to work, and leave complete freedom to the software developer to define their algorithms.

Once all test cases have been explored, and behaviour of faulted system verified, the specialists have all the elements required to produce a clear and effective FRD.

2.4 Functional Requirement Document

Writing an FRD is all about filling the gaps. The document should leave no room for anyone to assume anything not stated in the FRD. This means that through the requirement matrix (contained in the FRD), every function or operating mode of the system must be completely described. The specialists are able to produce such document by using the initial requirements list, and the ones that have been defined through the approach described in this paper. The plots describing the most relevant test cases may be attached as annex too.

3. CONCLUSIONS

A new Programmable Logic Controller development procedure has been defined and tested. This method tightens the gap between specialists and software development team, by mean of integration and validation of the general system logics. It proved to be effective, helping the specialists to release adequate FRD. The communication and data exchange between the specialists and the software development team may be supported by dynamic simulation video or test cases. The method described in this paper enhances the specialists' awareness on the entire system they are designing. This leads to an improved communication and data exchange between the system specialists and the software development team.

REFERENCES

Birbir, Y. and Nogay, S.H. (2008). Design and implementation of PLC-Based monitoring control system for three-phase induction motors fed by PWM Inverter, *International Journal of systems applications, engineering & development,* Vol. 2, Pag. 128 – 135.

Davidson, C.M. and McWhinne, J. (2000). Engineering the control software development process, *Factory 2000 - The Technology Exploitation Process, Fifth International Conference,* Vol. 1, Pag 247 – 25.

Documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts, ISO 5807:1985, (1985) *Information processing*, ISO Standards Handbook 1, International Organization for Standardization.

Gilberl, J.G. and Diehl, G.R. (1994). Application of Programmable Logic Controllers to Substation Control and Protection, *IEEE Transactions on Power Delivery,* Vol. 9, Pag. 384 – 388.

Ji, K., Dong, Y., Lee Y. and Lyoul, J. (2006). Reliability Analysis Safety Programmable Logic Controller. *SICE-ICASE International Joint Conference*.

Keller, J.P. (2006). Interactive control system design, *Control Engineering Practice,* Vol. 14, Pag. 177 – 184.

AKNOWLEDGEMENTS

The work has been performed through a close cooperation with ALENIA AERONAUTICA staff, with constant technical meetings and a continuous information exchange. In particular Authors wish to thank Eng. Maria Airoldi, Mr. Marco Mantovani and Eng. Alessandro Pasquino.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Task-Oriented Modelling of Rugged Terrain from Sparse Range Data ⋆

Dominik Belter[*], **Przemysław Łabęcki**[*], **Piotr Skrzypczyński**[*]

[*] *Institute of Control and Information Engineering, Poznan University of Technology, Poznań, Poland (e-mail: {db,pl,ps}@cie.put.poznan.pl)*

Abstract: Autonomous operation of a walking robot on rugged terrain requires to build a 3D map of this terrain. Different tasks of the walking robot control system can be supported by different representations of the environment, from a local grid-based elevation map to more abstracted, feature-based maps identifying traversable areas and obstacles. This paper presents a method for building of such a task-oriented representation of the environment from sparse 2D range measurements of the miniature Hokuyo 2D laser scanner. A procedure for local elevation map updating is described, then algorithms used for extraction of planar surfaces and selected features are presented. Experimental results are provided.

Keywords: walking robot, laser scanner, map building, plane fitting, edge extraction

1. INTRODUCTION

Mobile robots are increasingly employed in environments where the classic 2D maps are insufficient. In particular, walking robots require to build maps of the terrain, which serve the purpose of motion planning (Rusu *et al.* (2009)). Limited payload and computational resources of walking robots make the use of 3D laser scanners and stereovision problematic. Considering these limitations the Hokuyo URG-04LX miniature 2D laser scanner is applied in this work as a terrain sensor on a six-legged robot Messor. The scanner is tilted down, so the laser beam plane sweeps the ground ahead of the robot, enabling it to sense the terrain profile. Different geometric configurations of the sensing system were analysed, and the one that is best for terrain profile acquisition was chosen (Łabęcki *et al.* (2010)). Although the URG-04LX sensor is a small, lightweight, low-power device, and it is precise enough for the mapping task, in the chosen configuration it yields only 2D data about a terrain stripe located about half a meter in front of the robot. To obtain an environment representation the sparse range data have to be registered in a map using an estimate of the robot motion.

There are three main issues to deal with in a walking robot control system: foothold selection, robot stability, and path planning. This paper describes the development of a terrain modelling system which seeks to address these issues, and also provides an effective terrain visualization for a remote tele-operator. The foothold selection problem is crucial, because whenever the robot feet are placed improperly on the ground the risk of falling down is high. Because of that the core part of the proposed mapping system consists of a high-fidelity local elevation map that is optimized as a data source for foothold selection. Other robot functionalities: stability maintenance, path planning and teleoperation require a different model, which represents the terrain at larger scale, focusing on modelling particular obstacles (that have to be avoided), identifying traversable areas, and detecting such features as holes, ridges, and thresholds. The terrain representation takes a two level approach, in which the local grid-based terrain map is treated as a "virtual sensor" that collects the sparse range data over a certain area, while separate post-processing modules use this grid as an input to detect higher-level features. Thus, the modelling system is distributed and task-oriented, as particular feature extraction modules are data-driven, and work at different speeds required by their "parent" functionalities (Brzykcy *et al.* (2001)).

Currently, two higher-level modules are implemented: a plane segmentation module, and a local feature extractor based upon computer vision algorithms. The aim of the former is to extract the dominant surfaces in a scene and to find smaller obstacles based on their deviation from these surfaces. The plane-based segmentation keeps the amount of data reasonable, and allows building maps of larger areas that cannot be covered by a dense grid. The latter module is aimed at fast detection of ridges, ditches, and thresholds. The information on main surfaces and the protruding obstacles is not only useful to a path planner, but it helps also to maintain the robot stability, as information on the orientation of each surface in 3D is available. Moreover, both representations enrich the information presented to the remote operator, to whom correct and quick interpretation of the raw grid map can be a problem.

2. TERRAIN PERCEPTION

The URG-04LX scanner is mounted in the front of the Messor robot (Fig. 1a) at the nominal height of d_z=26cm, such that it aims forward and downward at the pitch angle of β=35° (Fig. 1b). The perception system design is a compromise between the requirements as to the field

⋆ This work is funded by MNiSW (Ministry of Science and Higher Education) grant no. N514 294635 in years 2008–2010.

of view and the accuracy of the perceived terrain profile. There are some sensor-specific requirements, related to the measurement errors characteristics of the Hokuyo URG-04LX scanner (Kneip *et al.* (2009)). The URG scanner is configured on the robot in such a way that the parts of terrain most interesting for the placement of robot feet, i.e. horizontal surfaces, are perceived with the incidence angle smaller than 40° to avoid specular reflections. The sensing system configuration ensures that objects which can be climbed by the Messor robot appear at distances falling into the range interval for which the URG-04LX sensor can be calibrated most effectively (Łabęcki *et al.* (2010)). An important advantage of the tilted sensor configuration is also reduction of mutual occlusions between the observed obstacles on an uneven terrain.

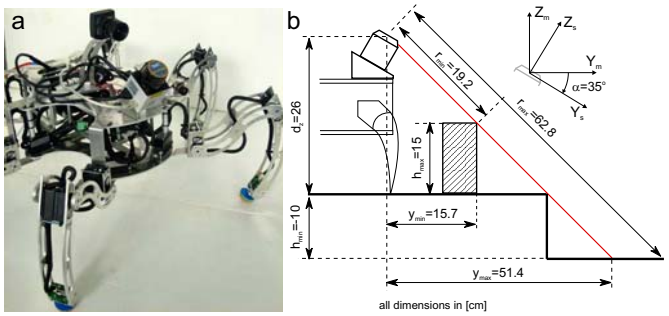


Fig. 1. Terrain perception system on Messor robot (a) and its geometry (b)

3. LOCAL 2.5D MAP OF THE TERRAIN

The local map that integrates the laser measurements moves with the robot always covering its surroundings. It is assumed that an estimate of the 6-dof robot pose is available. For a legged robot such an estimate is not easy to obtain. The URG data cannot be used for that purpose because of its sparse nature and the lack of overlapping between the new measurements and the recently perceived part of the terrain. For small local maps centered in the robot co-ordinates proprioceptive sensing exploiting an Inertial Measurement Unit (IMU) is enough, while for exploration of more extended areas a vision-based SLAM (Simultaneous Localization and Mapping) procedure is employed (Schmidt and Kasiński (2010)).

The local map serves mainly the purpose of foothold selection. For this task a grid-type map is preferred. It is easy to update in real-time, and may be directly used to select proper footholds (Belter *et al.* (2010)). To represent the 3D structure of the terrain an elevation map is used – a grid-type 2.5D map, where each cell holds a value that represents the height of the object at that cell (Krotkov and Hoffman (1994)). The map updating method is inspired by the algorithm of Ye and Borenstein (2004), developed for a wheeled vehicle with the Sick LMS 200 scanner. The terrain map consists of two grids of the same size: an elevation grid and a certainty grid. The elevation grid holds values that estimate the height of the terrain, while each cell in the certainty map holds a value that describes the accuracy of the corresponding cell's estimate of the elevation. Sensor-centered local grids of the size 100 × 100 cells with a cell size of 5 × 5 mm are used.

Laser range measurements are converted to 3D-points $\mathbf{p}_s$ in the scanner co-ordinate frame, then transformed to the map co-ordinates by using the robot pose estimate, which is assumed to be readily available, and finally projected onto the 2D grid map:

$$\mathbf{p}_m = \mathbf{T}_s^m \mathbf{p}_s, \tag{1}$$

$$\mathbf{T}_s^m = \mathrm{Rot}(X_m, \varphi_r)\mathrm{Rot}(Y'_m, \psi_r)\mathrm{Rot}(X''_m, \alpha), \tag{2}$$

where $\mathbf{p}_s{=}[x_s \;\; y_s \;\; z_s]^T$ and $\mathbf{p}_m{=}[x_m \;\; y_m \;\; z_m]^T$ are co-ordinates of the measured point in the sensor and the map frame, respectively. The homogeneous matrix $\mathbf{T}_s^m$ describes the transformation from the sensor frame to the map frame. This transformation consists of three rotations: the pitch φ_r and roll ψ_r angles of the robot trunk are shown in Fig. 2a and 2b, respectively, while the α angle (Fig. 2c) represents the constant pitch angle of the scanner with regard to (w.r.t.) the trunk.

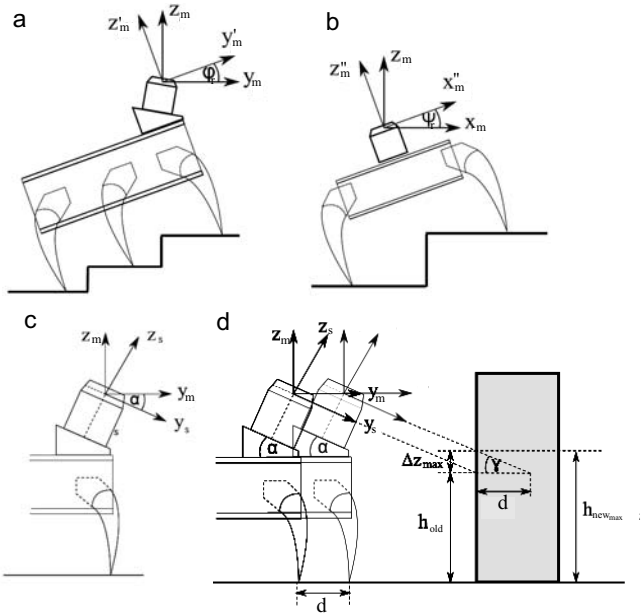


Fig. 2. Kinematic transformations of the measurements (a,b,c) and maximum change of elevation (d)

To make an assessment of the plausibility of a new range measurement a prediction of the maximum instantaneous change of elevation is used (Fig. 2d). For every two consecutive measurements r_k and r_{k+1}, the change of elevation $\Delta z_{\max}$ in a given cell of the map is computed taking into account the range measurement uncertainty, and the uncertainty of robot pose estimate:

$$\Delta z_{\max} = d_{(k,k+1)} \tan\gamma + \left|\frac{\partial z_m}{\partial r}\Delta r\right| + \left|\frac{\partial z_m}{\partial \psi_r}\Delta\psi_r\right| + \\ + \left|\frac{\partial z_m}{\partial \varphi_r}\Delta\varphi_r\right| + \left|\frac{\partial z_m}{\partial \alpha}\Delta\alpha\right|, \tag{3}$$

where z_m is the measured elevation of the observed point computed from (2), $d_{(k,k+1)}$ is a horizontal translation of the robot from k to $k+1$ time stamp, and γ angle is the total rotation of the trunk w.r.t. the X_m axis, which is obtained from the elements of the $\mathbf{T}_s^m$ matrix:

$$\gamma = \mathrm{atan2}\left(\frac{\cos\varphi_r \sin\alpha + \sin\varphi\cos\alpha\cos\psi_r}{-\sin\varphi_r\sin\alpha + \cos\varphi\cos\alpha\cos\psi_r}\right). \tag{4}$$

The values of $\Delta\varphi_r$, $\Delta\psi_r$ and $\Delta\alpha$ are maximum errors of the respective angles, while Δr is the laser scanner range measurement error, which is computed upon the sensor model (Łabęcki *et al.* (2010)). The value of $\Delta\alpha$ is 1° (a constant), but the values of $\Delta\psi_r$ and $\Delta\varphi_r$ depend on the accuracy of the IMU sensor used in the robot.

The maximum elevation change given by (3) is valid only if the robot is moving along a straight line. However, the walking robot often changes its orientation (the yaw angle θ_r) during motion because it has to put its feet at proper footholds. To enable computation of the $\Delta z_{\max}$ while changing the orientation, the idea given in (Ye and Borenstein (2004)) is extended to include also the turning motion. Assuming that the robot observes an obstacle of constant elevation, and turns by an angle of θ (Fig. 3a), the distance between the two points, p and p', being observed by the sensor is computed as $\Delta x = (r+d_s)\tan(\theta/2)\mathrm{sgn}(\theta)$ (Fig. 3b). Then, the instantaneous horizontal translation of the sensor is computed: $d_{(k,k+1)} = \Delta x \tan(\theta)$, and used in (3).

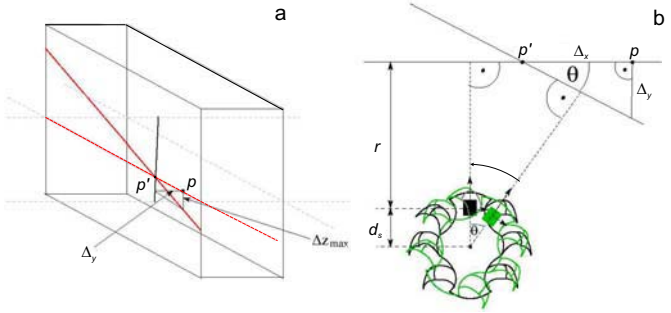


Fig. 3. Determination of the elevation change while turning (a,b)

A cell in the elevation map is denoted as $h^{[i,j]}$, and a cell in the certainty map as $c^{[i,j]}$. Whenever a new measurement is available cells in the certainty map are updated at first:

$$c^{[i,j]}_{(k+1)} = \begin{cases} c^{[i,j]}_{(k)} + a \text{ if } |h_{m(k+1)} - h^{[i,j]}_{(k)}| \le |\Delta z_{\max(k)}| \\ \quad \text{or } c^{[i,j]}_{(k)} = 0 \\ c^{[i,j]}_{(k)} \text{ otherwise,} \end{cases} \tag{5}$$

where a is the increment of the certainty value, and h_m is the elevation of the measured point, computed as $h_m = z_m + d_z + h_{\mathrm{ref}}$ taking into account the current height of the robot d_z (it is computed upon the known kinematics of the robot and the angles measured in joints, cf. Fig 1b), and the reference elevation h_{ref} at which the robot is located (obtained from the already created part of the map). Next, cells in the elevation map are updated:

$$h^{[i,j]}_{(k+1)} = \begin{cases} h_{m(k+1)} \text{ if } h_{m(k+1)} > h^{[i,j]}_{(k)} \\ h^{[i,j]}_{(k)} \text{ otherwise,} \end{cases} \tag{6}$$

In (Ye and Borenstein (2004)) the CAS (Certainty Assisted Spatial) filter is introduced, which employs the physical constraints on motion continuity and spatial continuity to remove corrupted values from the elevation map. However, during tests of the mapping system with the URG-04LX scanner on various obstacles the CAS filter failed to remove most of the elevation map artifacts due to "mixed pixels" – spurious range measurements that arise when a laser beam hits simultaneously two surfaces at different distances (Łabęcki *et al.* (2010)). Different spatial characteristics of the mixed range measurements in the ToF-based LMS 200 and the phase-shift-based URG scanner are a possible explanation of this behaviour (Skrzypczyński (2008)).

For that reason the CAS filter is not used in the presented system. However, to avoid erratic behaviour of the map-building procedures mixed measurements are eliminated at the pre-processing stage by two filtering algorithms described in details in (Łabęcki *et al.* (2010)). The weighted median filter proposed by Ye and Borenstein (2004) as the mechanism that fills in the missing data is used here as well. The output h_{wm} of the filter is assigned to each cell in the elevation map for which the corresponding cell in the certainty map $c^{[i,j]}$=0, what means that this cell was never updated:

$$h^{[i,j]} = \begin{cases} h^{[i,j]}_{\mathrm{wm}} \text{ if } c^{[i,j]} = 0 \\ h^{[i,j]} \text{ otherwise.} \end{cases} \tag{7}$$

This mechanism enables to fill in small portions of the elevation map that are invisible to the sensor due to occlusions.

4. FINDING STRUCTURES IN ELEVATION MAPS

4.1 The plane fitting algorithm

The aim of this algorithm is to identify areas of an elevation map that lie on a common plane. The grid map is treated as a set of points in 3D – the center of each cell is treated as a data point with the elevation value interpreted as the z co-ordinate. Because the elevation map may contain many outliers (e.g. due to small obstacles) the RANSAC (Random Sample Consensus) paradigm is applied to fit a plane to the set of noisy points.

For each set of points this method returns a fitted plane (i.e. the parameters of the plane), the points that fit to this plane within a given threshold (called inliers), and the points that are too far from the plane (outliers). An estimate of the plane parameters is calculated using a random sample of three points from the input set. To ensure that all of the three selected points are inliers, the selection is repeated several times and the candidate plane with the highest number of inliers is returned. The number of repetitions is determined observing a stop criteria based on a predefined percentage of outliers in the initial set. However, to prevent the procedure from taking too much time, another stop criteria is used in parallel, which is the maximum number of iterations. If the procedure cannot find a plane candidate that satisfies the maximum percentage of outliers criteria within the maximal number of iterations the plane fitting is considered not applicable to the given data set.

Whenever a plane candidate satisfying the above criteria is found, its parameters are refined by applying the total least squares method, considering the sum of squared distances to the plane. The estimation procedure assumes that the errors in locations of data points can be modelled by isotropic Gaussian noise. This step ensures that the plane is fitted optimally to all the inliers.

However, simple estimation with RANSAC and least squares returns only one plane that fits the provided set of

points. In general, only some of the points are inliers, the rest are outliers. Therefore, the presented algorithm uses a recursive approach, invoking the RANSAC procedure again with the outliers as input. Also, some computer vision methods are used to divide the outliers into groups. The general diagram of the algorithm is presented in Fig. 4

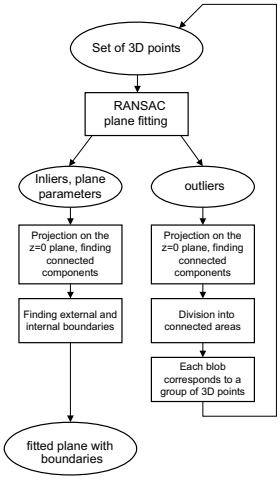


Fig. 4. Block diagram of the plane fitting algorithm

At first, all values from the elevation map (represented as 3D points) are subject to the RANSAC-based plane fitting method. The output is a single plane (that best fits this group of points), a set of inliers, and a set of outliers. Both groups of points are then separately projected onto the $z = 0$ plane. The grid of the projected points is then normalized, i.e. the x and y co-ordinates of the points are scaled to provide unitary spacing. Points adjusted in this way can be treated as a binary image, where each pixel corresponds to a point in 3D space. This step is followed by finding connected components (blobs) in both inlier and outlier images. Blobs in the inlier image are used to find external and internal boundaries of the plane elements. Blobs in the outlier image are used to divide the outliers to separate groups. Blobs with area smaller than a given threshold are considered spurious and discarded (a threshold of 5 pixels is used). Remaining blobs are treated as separate regions, and the whole algorithm is used again recursively, separately for each blob (i.e. the 3D points corresponding to the blob are used as input points to the RANSAC-based part). An example of the binary image of inliers that is a result of the first plane estimation from the elevation map shown later in the paper (cf. Fig. 8b) is presented in Fig. 5.

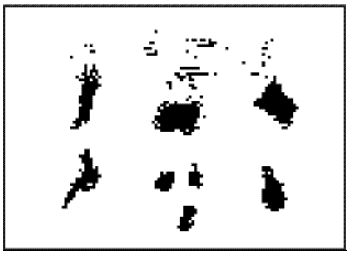


Fig. 5. Inliers projected onto the $z = 0$ plane

4.2 The vision–based obstacle extractor

The vision-based obstacle extraction algorithm treats the elevation grid map as a gray-scale image. Vision methods are then used to determine non-traversable areas, such as slopes and rifts.

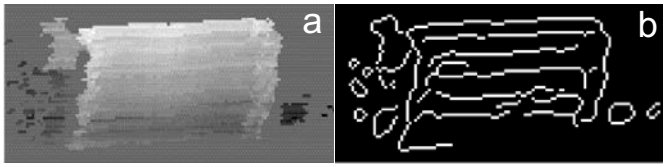


Fig. 6. Elevation map as an image (a) and edges extracted using the Canny detector (b)

The first step is filtering of the image. This step is necessary to avoid detecting measurement noise as actual obstacles. Therefore, a modified median filter is used. The goal of the modification is to retain edges of magnitude greater than a given value: if the differences between the median value and the extreme values within the filter window are greater than the threshold, the filter does nothing. Otherwise, it acts as a standard median filter, replacing the subject pixel, with the median value of the window. After filtering, the image is subject to the Canny edge extractor. This algorithm extracts strong as well as weak edges (Fig 6). To find edges that are non-traversable to the robot (i.e. the magnitude of the edge is greater than a given threshold), each pixel belonging to an edge is examined, along with its neighbors. If the difference between the edge pixel and any of its exact neighbors is greater than the threshold, the pixel is classified as non-traversable. To detect wider edges (e.g. ditches), similar differences within a radius of 2 pixels are computed, and twice the threshold is required to classify the pixel to a non-traversable obstacle. The block diagram of the algorithm is presented in Fig. 7

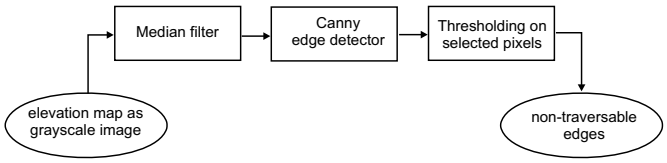


Fig. 7. Block diagram of the vision-based obstacle extraction method

5. EXPERIMENTAL RESULTS

Mapping experiments with the Messor robot were performed on a rocky terrain mockup (Fig. 8a) of size 2×2 m. The legged odometry and the on-board IMU were used to obtain an estimate of the robot pose. The obtained elevation maps correctly represent all encountered obstacles – an example is given in Fig. 8b. However, due to the imprecise robot pose estimates the obtained maps are sometimes skewed, and their surfaces are slightly wavy. The assessment of the map correctness is only qualitative, as there is no precise ground truth available for the terrain mockup. Obtaining such a ground truth requires scanning of the whole mockup with a precisely moved sensor, what was not possible so far. The certainty map (Fig. 8c) acquired by the robot reveals regions of very low certainty (pointed by arrows) behind some obstacles. These areas were never observed by the robot due to occlusions, and they are too large to be filled-in by the median filter. However, the certainty map allows the control system to avoid such unknown areas during walking.

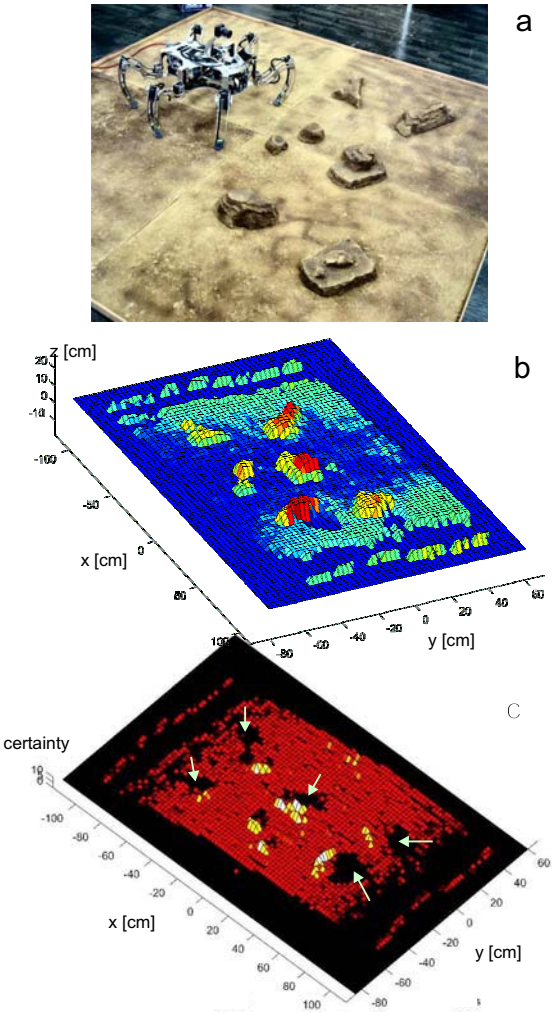


Fig. 8. Messor on the mockup (a) obtained elevation map (b) and certainty map (c)

The elevation map shown in Fig. 8b is a good data source for foothold selection, but it can be post-processed to obtain feature-based representations suitable to support other tasks of the robot control system. In this case extraction of planar surfaces allows to obtain the main traversable area and small planar patches representing obstacles protruding from this "ground" plane. These patches can be further examined w.r.t. their size and orientation in 3D to diagnose if they could provide stable support for the robot legs (Fig. 9a). The robot path planner also can benefit from this representation considering the "holes" created by protruding objects as 2D obstacles (Fig. 9b). With such a representation of the environment any of the well-known algorithms for fast 2D path planning can be employed. Also the vision-based feature extractor can provide information on the obstacles (Fig. 9c,d). Combining the information on non-traversable boundaries (Fig. 9d) with the information on planar areas the control system can figure out which areas are not accessible to the robot.

A good example of the usability of the proposed post-processing methods is given by a sloppy, hill-like terrain, which can be climbed by the walking robot (Fig. 10a). The vision-based boundary extraction provides information on the borders of the hill (Fig. 10b). If the robot puts

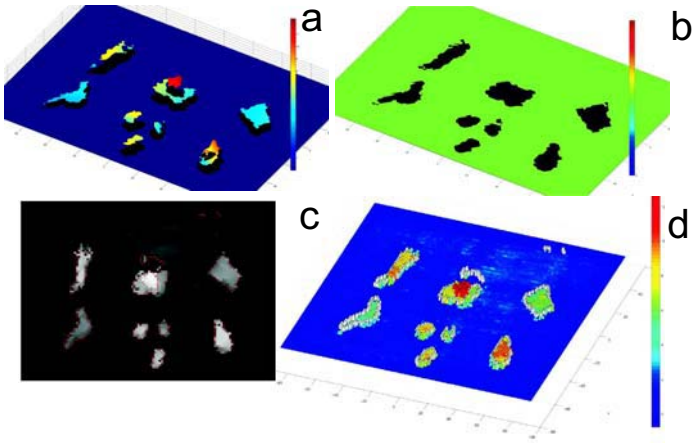


Fig. 9. Extracted planar surfaces (a) areas occupied by obstacles (b) edges found by Canny detector (c) and non-traversable boundaries (d)

accidentally one of its leg outside of such a border, it can fall down easily. On the other hand, the information on orientation of the hill surface (Fig. 10c) is important for maintaining the robot stability while climbing this object.

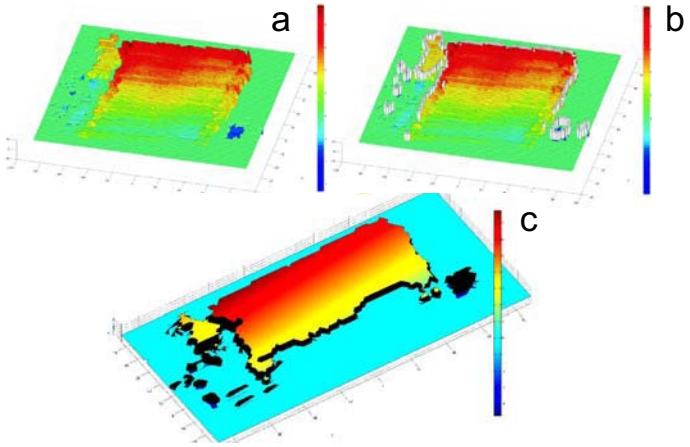


Fig. 10. Elevation map of the slopped obstacle (a) non-traversable boundaries (b) and main surfaces (c)

Some preliminary outdoor experiments were also performed (Fig. 11a), showing the ability of the presented map-building approach to acquire a map of the terrain with vegetation. In spite of using only the data from proprioceptive sensors for positioning, the proposed approach was able to create an elevation map, which identifies the main structures encountered by the robot, like the large root pointed out by the arrows (Fig. 11b). Although the elevation map quality is lower than in the indoor experiments, the plane-fitting method turned out to be robust enough to extract correctly the main surface (the grass-covered area) on which the robot is located (Fig. 11c).

6. DISCUSSION AND CONCLUSIONS

This paper exploits a number of known methods and algorithms, like the elevation map, the RANSAC algorithm, and the Canny detector, but shows that these elements can be combined and used in a new structure, resulting in a system that can solve the problem of rugged terrain modelling from sparse range data.

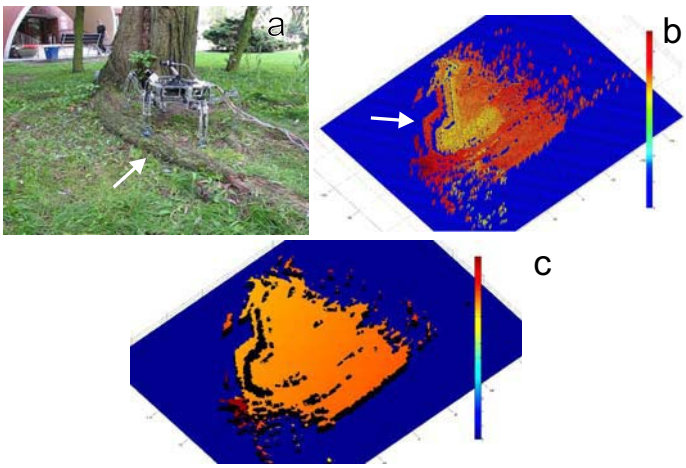


Fig. 11. Messor walking outdoors (a) obtained elevation map (b) and extracted planar surface (c)

The problem of rough terrain mapping has been considered many times in the robotics literature. The elevation map approach used here was pioneered by the work on Ambler (Krotkov and Hoffman (1994)). Grid-based approaches are used also on more recent walking robots, like Lauron IV. Roennau *et al.* (2009) use a time-of-flight range camera on the Lauron robot, while this paper shows that such a map can be built efficiently with a simpler and cheaper 2D scanner on a walking robot. Although an elevation map can be extended to describe holes in the environment, such like tunnels (Pfaff *et al.* (2007)), this is not applicable to the robot under study, because it cannot perceive objects not located at the ground level. Recent research in terrain modelling resulted in methods that do not assume a fixed discretization of the space, such as the work of Plagemann *et al.* 2008, which applies Gaussian process regression to the problem of rough terrain mapping. This off-line approach focuses on modelling large voids and discontinuities that can appear when using data from a long-range 3D sensor. In contrast, the system presented here uses a short-range sensor configured to minimize discontinuities in the obtained map. An example of mapping system tailored specifically to the application is given by Sheh *et al.* (2007), where geometric features (ruts, ridges) are extracted from range data for autonomous random stepfield traversal. This system also uses a time-of-flight 3D camera, and yields only a set of local features which mainly support the teleoperation task. No information on planar areas or general traversability of the terrain is extracted from the range data. Poppinga *et al.* (2009) propose and approach for accurate surface extraction from noisy 3D point clouds, which is implemented on an all-terrain tracked vehicle. However, this approach again assumes 3D data from a range camera and cannot be applied to the sparse range data from a 2D scanner that need a registration step in the elevation map.

The experimental results presented in this paper verified the proposed approach in both laboratory and real-world (outdoor) setups, however, the experiments are still of rather small scale, and will be repeated soon for more realistic scenarios. This should be possible after fully integrating the visual SLAM system on the Messor robot. Another direction of the further research is to show in experiments that the multi-aspect terrain model is actually beneficial to the motion planning procedures of the robot.

REFERENCES

Belter, D. Łabęcki, P., Skrzypczyński, P. (2010) Map-based Adaptive Foothold Planning for Unstructured Terrain Walking, *Proc. IEEE Int. Conf. on Robot. and Automat.*, Anchorage, pp. 5256–5261.

Brzykcy, G., Martinek, J., Meissner, A., Skrzypczyński, P. (2001) Multi-Agent Blackboard Architecture for a Mobile Robot, *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, Maui, pp. 2369–2374.

Kneip, L., Tache, F., Caprari, G., Siegwart, R. (2009) Characterization of the compact Hokuyo URG-04LX 2D laser range scanner, *Proc. IEEE Int. Conf. on Robot. & Automat.*, Kobe, pp. 1447–1454.

Krotkov, E., Hoffman, R. (1994) Terrain Mapping for a Walking Planetary Rover, *IEEE Trans. Robot. and Automat.*, 10(6):728–739.

Łabęcki, P., Rosiński, D., Skrzypczyński, P. (2010) Terrain Perception and Mapping in a Walking Robot with a Compact 2D Laser Scanner, in *Emerging Trends in Mobile Robotics* (H. Fujimoto *et al.*, eds.), Singapore, World Scientific, pp. 981–988.

Pfaff, P., Triebel, R., Burgard. W. (2007) An Efficient Extension to Elevation Maps for Outdoor Terrain Mapping and Loop Closing, *Int. Journal of Robotics Research*, 26(2):217–230.

Plagemann, C., Mischke, S., Prentice, S., Kersting, K., Roy, N., Burgard, W. (2008) Learning Predictive Terrain Models for Legged Robot Locomotion, *Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, Nice, pp. 3545–3552.

Poppinga, J., Vaskevicius, N., Birk, A., Pathak, K. (2008) Fast Plane Detection and Polygonalization in Noisy 3D Range Images, *Proc. IEEE/RSJ Int. Conf. on Intell. Robots and Systems*, Nice, pp. 3378–3383.

Roennau, A., Kerscher, T., Ziegenmeyer, M., Zoellner, J. M., Dillmann, R. (2009) Six-Legged Walking in Rough Terrain Based on Foot Point Planning, in: *Mobile Robotics: Solutions and Challenges* (O. Tosun *et al.*, eds.), Singapore, World Scientific, pp. 591–598.

Rusu, R. B., Sundaresan, A., Morisset, B., Hauser, K., Agrawal, M., Latombe, J.-C., Beetz, M. (2009) Leaving Flatland: Efficient Real-time Three-dimensional Perception and Motion Planning, *Journal of Field Robotics*, 26(10):841–862.

Schmidt, A., Kasiński A. (2010) The Visual SLAM System for a Hexapod Robot, *Computer Vision and Graphics* (L. Bolc *et al.*, eds.), LNCS 6375, Berlin, Springer, pp. 260–267.

Sheh, R., Kadous, M. W., Sammut, C., Hengst, B. (2007) Extracting Terrain Features From Range Images for Autonomous Random Stepfield Traversal, *Proc. IEEE Int. Workshop on Safety, Security and Rescue Robotics*, Rome, pp. 1–6.

Skrzypczyński, P. (2008) On Qualitative Uncertainty in Range Measurements from 2D Laser Scanners, *Journal of Automation, Mobile Robotics and Intelligent Systems*, 2(2):35–42.

Ye, C., Borenstein, J. (2004) A Novel Filter for Terrain Mapping with Laser Rangefinders, *IEEE Trans. Robot. and Automat.*, 20(5) pp. 913–921.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Flight Path Optimization Using Primitive Manoeuvres: a Particle Swarm Approach

L. Blasi*, S. Barbato*, M. Mattei*

**Dipartimento di Ingegneria Aerospaziale e Meccanica, Second University of Naples, Via Roma 29, 81031 Aversa, Italy*
Tel: 0039-081-5010289; e-mail: {luciano.blasi, simeone.barbato, massimiliano.mattei@unina2.it}.

Abstract: This paper studies the possibility to use Particle Swarm Optimization (PSO) techniques to perform two- and three-dimensional flight path optimizations compliant with operational constraints. Assuming a typical flight surveillance mission, such constraints are defined in terms of "target" and "no-fly" zones, fixed way-points and landing areas. It is well known that the success of flight path optimization techniques strongly depends on the trajectory parameterization adopted. In the proposed approach, flight paths are firstly divided into a finite number of segments; each segment is associated to an elementary manoeuvre chosen within a finite set and represented by means of a two-bit-coded number. This novel approach allows defining the sequence of manoeuvres through a reduced number of discrete-type variables that can be easily handled by the Particle Swarm optimizer. In addition to proper penalty functions, a linear obstacle avoidance model is introduced favouring the identification of feasible flight path. The nonlinear optimization problem is then formulated in terms of both single objective and multi objective cost function. Numerical results confirm that the proposed PSO-based path finding algorithm is particularly indicated to solve these kinds of mixed optimization problems.

Keywords: Flight path generation, Trajectory optimization, Particle Swarm Optimization, Genetic Algorithms, Flight Control.

1. INTRODUCTION

Planning optimal flight trajectories, consistent with mission objectives, operational scenario, and vehicle dynamics and performance, is a problem of interest both for civil and military applications with manned or unmanned vehicles. Parameters defining flight missions are usually related to regions to fly over, desired flight altitudes on targets. The operational scenario also provides constraints depending on take off and landing areas, no-fly zones or high-risk zones, the presence of mountains or adverse climatic conditions, minimum/maximum distance from base stations or cooperating vehicles. Finally, constraints related to the specific aircraft used, like maximum climbing rate, maximum and minimum speed, minimum turning radius, maximum fuel capacity etc., have to be satisfied too.

During the past decades, a lot of work has been carried out on the trajectory optimization for many kind of vehicles. The variational formulation is probably the most natural and rigorous one for this class of problems. However, the possibility to solve complex problems with variational methods is very poor. Many papers deal with numerical direct and indirect methods based on the solution of a Non Linear Programming (NLP) problem (Betts, 1998). By means of some approximations, feasible trajectories can be generated following a purely geometrical approach based on topological techniques creating a sequence of way points. This sequence can derive from probabilistic or potential methods (Cen et al., 2007). A classical geometric approach guaranteeing optimality conditions in terms of paths length and smooth trajectories compliant with curvature constraint was proposed by Dubins (1957) and refined in Anderson et al. (2005) and Chitsaz and LaValle (2007). An interesting technique, taking into account flight dynamics, is based on the so called "motion primitives" (Dever et al., 2006), where flight paths are defined through a sequence of trim conditions and manoeuvres. Due to the complexity and variety of the problem, non-conventional, nature-inspired optimization methods have shown their effectiveness and robustness in a wide range of optimization problems, taking advantage from some specific features such as the capability in easily handling mixed-type design variables accounting for a large number of constraint functions, and a parallel-like searching method leading to a greater effectiveness in finding global minimum within the design space (Hu et al., 2004). Among evolutionary computational techniques, optimization methods based on Swarm Theory (Particle Swarm Optimization, PSO) share many similarities with Genetic Algorithms (GA), having the great advantage of a simpler implementation (Eberhart and Shi, 1998; Raja and Pugazhenhi, 2009; Wang et al., 2006).

Potentials of PSO techniques in the field of flight path generation are further investigated in this paper. The objective is the development of a particle swarm-based procedure performing path optimization compliant with operational constraints.

Assuming a typical surveillance mission, environmental constraints are defined in terms of different "target" and "no-

fly" zones, fixed way-points and landing area. Flight paths are described through a set of continuous and discrete parameters varying within defined ranges. Each particle of the swarm is represented by a numerical combination of these parameters.

Trajectories, starting from a specified point with a given direction and ending on a selected landing area, are in practice made up of circular arcs and straight lines as for the Dubins curves proposed for free space trajectory generation in Dubins (1957). Such a sequence of circular arcs and straight lines has been associated to a finite number of elementary manoeuvres represented by means of binary number pairs. The proposed path finding algorithm is shown to have a high capability in identifying feasible paths in a constrained environment.

Both single-objective and multi-objective optimization procedures have been explored. The former was implemented so as to minimize the total flight path length; the latter also tries to maximize the trajectory length covered over a specified target area. Sensitivity studies with increasing problem complexity have been performed changing both number and position of "target" and "no-fly" zones. Computational time monitoring finally allowed making a preliminary assessment of PSO suitability for possible "on-line" flight path optimization problems.

2. PSO METHODOLOGY

The optimization technique based on Swarm Theory is a nonlinear method belonging to the class of evolutionary computational techniques that find solution through a probabilistic search process guided by a fitness function. It takes inspiration from the social behaviour of groups of simple creatures as swarms of bees that exhibit some form of collective intelligence based on information exchange. The searching for optimal solutions performed with PSO is obtained defining a population of particles, each one exploring the search space and communicating results to the rest of group. In this way, population evolution can be obtained through the cooperation among individuals (or particles). Each particle i has two state variables which are function of the time step k of the optimization process: current position $\mathbf{x}_i(k)$ and current velocity $\mathbf{v}_i(k)$. Particles also have a memory containing the previous particle best position or *personal best position* $\mathbf{p}_i(k)$ and the swarm best position or *global best position*, $\mathbf{g}_i(k)$.

At time step k+1, the particle position is updated according to the relation

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) + \mathbf{v}_i(k+1) \tag{1}$$

where $\mathbf{v}_i(k+1)$ is the i-th particle velocity, which is calculated as

$$\begin{aligned}\mathbf{v}_i(k+1) = \chi\Big\{\omega\,\mathbf{v}_i(k) + c_1\mathbf{r}_1 \otimes \big[\mathbf{p}_i(k) - \mathbf{x}_i(k)\big] \\ + c_2\mathbf{r}_2 \otimes \big[\mathbf{g}_i(k) - \mathbf{x}_i(k)\big]\Big\}\end{aligned} \tag{2}$$

In (2) $\mathbf{r}_1$ and $\mathbf{r}_2$ are vectors of random numbers uniformly distributed over the range [0,1]; the symbol $\otimes$ denotes a component wise vector product; c_1 and c_2, named *cognitive* and *social parameter* respectively, are scalar weights tuning $\mathbf{p}_i(k)$ and $\mathbf{g}_i(k)$ influence on the particle velocity; the *inertia weight*, ω, is a reducing factor for $\mathbf{v}_i(k)$ whereas the *constriction factor*, χ, is used to limit the particle velocity. Both ω and χ control swarm exploitation as well as exploration capability (Engelbrecht, 2005), heavily affecting convergence speed and effectiveness of the optimization task.

The effectiveness shown by PSO-based techniques in many single-objective optimization problems, combined with its capability in keeping information about the evolution of all the particles at the same time, has made the extension of PSO techniques to multi-objective optimization problems almost a natural progression (Multi-Objective Particle Swarm Optimization - MOPSO).

In a constrained multi-objective optimization task, we seek to simultaneously optimize D objectives $f_i(\mathbf{x})$, $i = 1,...,D$, depending on a vector $\mathbf{x}$ of K mixed continuous and discrete decision variables, subject to J equality/inequality constraints

$$c_j(\mathbf{x}) \le 0 \quad j = 1,...,J \tag{3}$$

If we assume, for the sake of simplicity, that all these objectives have to be minimized, the problem can be stated as:

$$\text{minimize} \quad f_i(\mathbf{x}) \quad i = 1,...,D \quad s.t. \tag{4}$$

$$c_j(\mathbf{x}) \le 0 \quad j = 1,...,J \tag{5}$$

A decision vector $\mathbf{u}$ is said to dominate a decision vector $\mathbf{v}$ if

$$f_i(\mathbf{u}) \le f_i(\mathbf{v}) \quad \forall i \in \{1,...,D\} \tag{6}$$

$$\exists j \in \{1,.....,D\} : f_j(\mathbf{u}) < f_j(\mathbf{v}) \tag{7}$$

Therefore, the aim of a multi-objective optimization problem is the identification of non-dominated solutions whose related objective vectors in the objective space are referred to as the *Pareto front*. Pareto optimality concept can be easily used to define a MOPSO fitness function that takes in account the degree of dominance of each solution among the population.

3. THE OPTIMIZATION PROBLEM FORMULATION

The PSO technique is now applied to the problem of minimizing the flight trajectory length in the presence of a certain number of no-fly zones and target areas in a two-dimensional space.
We assume that flight paths are made up of circular arcs and straight lines. All candidate trajectories start from a specified Starting Point (SP) with a fixed velocity vector (i.e. fixed speed and direction). The flying vehicle has to reach a certain landing zone centred at the Destination Point (DP), after flying over desired target regions in the presence of no-fly zone. The sequence of target areas to visit during the flight can be decided by the optimization process.

The basic idea is to model the sequence of circular arcs and straight lines into a sequence of four primitive manoeuvres which are represented by means of a two-bit-coded integer numbers. This novel formulation allows reducing the number of variables involved in the optimization process, taking full advantage of PSO capability in handling discrete variables.

A variant of the basic PSO algorithm has been implemented following the technique introduced by Kennedy et al. (2001) to cope with mixed variables applications.
In particular, the following "basic" manoeuvres are defined with the related binary code:
1) Turn right (0 1); 2) Turn left (1 0); 3) Straight flight (0 0); 4) Align aircraft nose to the target (1 1).
In order to identify a certain sequence of n manoeuvres (defined by an n-ple of binary number pairs or equivalently by two n-ple of binary number), we define two integer variables, namely *VP1m* and *VP2m*, varying in the range [0, 2^n-1] whose binary codification provides the two n-dimensional vectors representing the manoeuvres sequence. Table 1 shows an example of a n = 10 manoeuvres sequence defined on the basis of a particular choice of *VP1m* and *VP2m*.

Table 1. Example of manoeuvres sequence

VP1m= 346	*VP2m*= 715	
0	1	Turn right
1	0	Turn left
0	1	Turn right
1	1	Alignment to the target
0	0	Straight flight
1	0	Turn left
1	1	Alignment to the target
0	0	Straight flight
1	1	Alignment to the target
0	1	Turn right

The number of segments, n, is a parameter to be fixed a priori on the basis of the problem complexity.
For all the alignment to the target segments the connection between the (i-1)-th path segment, and the i-th one is obtained by means of a circular connection arc with a radius of curvature $r_t \geq r_{\min}$. As shown in Fig. 1, r_t is a linear function of the distance, d_i ,between the i-th path segment starting point and the nearest no-fly area with radius R.

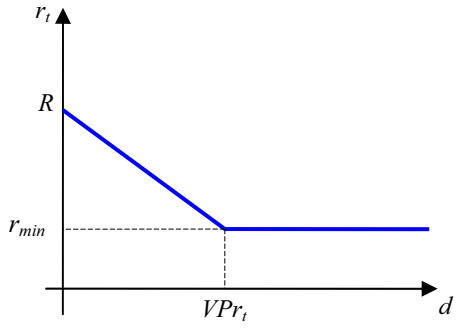


Fig. 1. Connection arc radius variation with avoid area distance d

During the optimization process, the line slope is tuned by the design variable VPr_t. The value of the minimum radius of curvature, r_{min} , is a user-defined parameter.

In case of a turn manoeuvre, a suitable radius of curvature has to be selected. We assume that the turn radius, r, is a linear function of the distance, d_i ,between the i-th path segment starting point and the nearest no-fly area as shown in Fig. 2. This way the nearest is the no-fly area, the tightest is the turn manoeuvre and vice versa. Turn radius is bounded below by the user-defined parameter r_{min} . During the optimization process, the line slope is tuned by design variables, *VP1r* and *VP2r*.

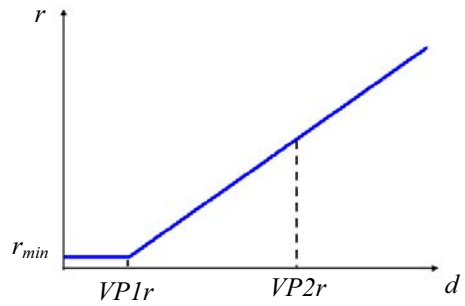


Fig. 2. Turn radius variation with avoid area distance d

The use of such a linear obstacle avoidance model favours the identification of feasible flight paths in addition to proper penalty functions degrading the fitness value whenever one or more constraints are violated. In order to better suit the line slope, we assume different variables for no-fly areas (superscript a) and target areas (superscript t).

If the operational scenario has more than one target area, the sequence of targets can be defined by means of a discrete-type design variable vector σ. Velocity, V , and the trajectory segments length, Δs , are also included in the vector of design variables. The design variables set representing a 2-D path turns out to be composed of the following variables:

$$\mathbf{x} = \left[VP1m, VP2m, VPr_t, VP1r^a, VP2r^a, VP1r^t, VP2r^t, \sigma^T, V, \Delta s \right]^T \quad (8)$$

4. NUMERICAL RESULTS: 2-D SINGLE-OBJECTIVE

Applications of the proposed PSO based methodology to single-objective trajectory optimization problems with different operational scenarios are briefly reported hereinafter.
Path length is selected as the objective function to be minimized. We assume n=10 (Scenario 1) and n=20 (Scenario2-4) as path segments maximum number. We consider circular targets and no-fly areas assuming a circular landing area centred at point $(5,5)$ with a radius of 0.1 km.

Scenario 1. We consider 2 no-fly areas centred at (2, 2.5) and (4, 4), having a radius of 0.5 and 1.0 km respectively. SP is

placed at (1.5, 1.5), the initial velocity vector (represented with an arrow in the following figures) is directed toward the centre of the nearest no-fly area. We select a population size of 100 particles, and a maximum of 100 iterations for PSO. Since PSO is a probabilistic-type technique, in order to assess final solution reliability, different optimization tasks have been performed starting from randomly generated swarms. Fig. 3 shows optimized paths obtained over 7 different runs. Optimum trajectories appear quite similar in terms of both shape and length. In particular, the best path measures 5.409 km whereas the average optimum path length is 5.413±0.0035 km.

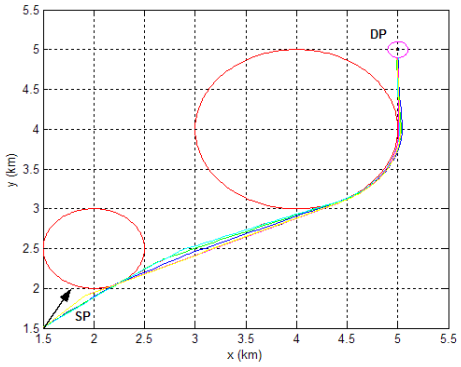


Fig. 3. Scenario 1: Optimized paths over 7 runs

Scenario 2. We now consider 24 circular no-fly areas, placed in a grid pattern resembling a sort of urban scenario (Fig.4). Each no-fly area has a radius of 0.3 km. SP is placed at $(1.5,1.5)$. The initial velocity vector is perpendicular to the x-axis. A population size of 100 particles and a maximum number of 100 iterations are chosen. Also for this scenario, in order to assess solution sensitivity to the initial population, 7 different optimization tasks have been performed. Optimized paths are superimposed in Fig. 4.

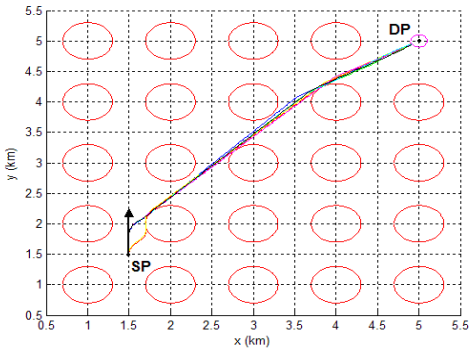


Fig. 4. Scenario 2: Optimized paths over 7 runs

Even with this more complex operational scenario, the optimization procedure provides quite similar solutions in terms of both shape and length for different runs. In particular the best path measures 4.991 km whereas the average optimum path length we obtained over 7 runs is 5.003±0.0067 km.

Scenario 3. This scenario is obtained by adding 3 way-points to Scenario 2, namely WP1 at point (1.5, 3.0), WP2 at point (3.0, 3.5), and WP3 at point (4.0, 1.5).

We minimize the path length, leaving the algorithm free to optimize the way-points sequence by using the discrete-type variable vector σ. We use a population size of 100 particles with a fixed number of 100 iterations.

As for previous scenarios, 7 runs with different starting conditions were performed (Fig. 5). As we can see the algorithm was always able to find the best way-point sequence by handling the discrete-type variable vector σ. In particular the best path measures 9.355 km whereas the average optimum path length we obtained over 7 runs is 9.640±0.147 km. Compared to Scenario 1 and 2, a higher scattering of optimum path lengths is observed which is however around 1.5 %.

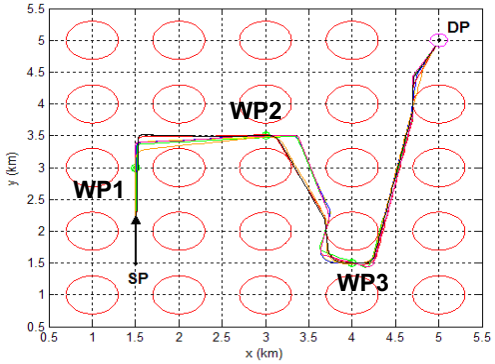


Fig. 5. Scenario 3: Optimized path over 7 runs with a free way-points sequence

5. NUMERICAL RESULTS: MULTI-OBJECTIVE CASE

A preliminary application of a MOPSO algorithm to a simple flight path optimization problem is reported. Beyond the path length minimization, a further objective has been considered, that is the maximization of the trajectory length covered over a specified target area. We consider Scenario 1, with the addition of a circular target area to fly over as much as possible, centred at (3.0, 2.8) and having a radius of 0.5 km. A population size of 500 particles is selected. The optimization task consists of 600 iterations. Fig. 6 shows the Pareto front obtained at the end of the optimization process whereas Fig. 7 shows the two optimized trajectories corresponding to the end points of the Pareto front.

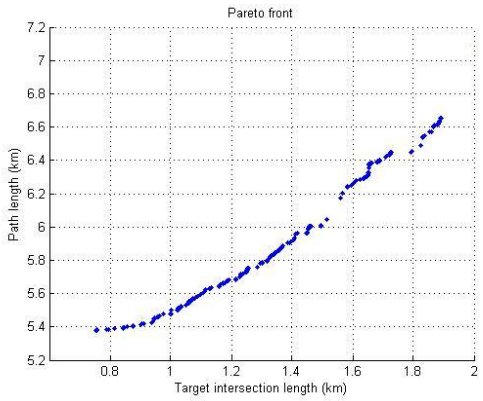
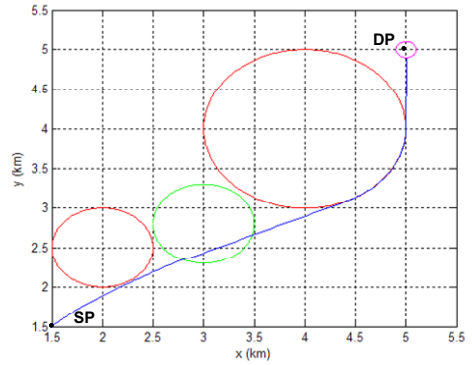
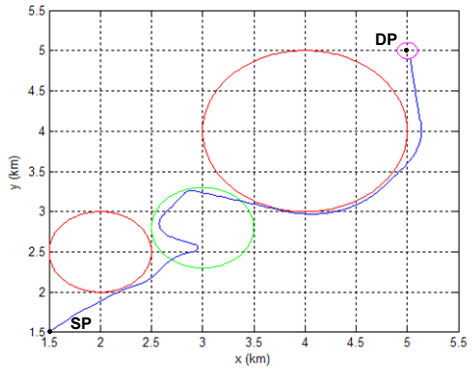


Fig. 6. Scenario 4: Pareto front

The flight path corresponding to minimum target intersection length (0.754 km), measuring 5.380 km, is shown in Fig. 7-a. As we can see, this solution is very close to the trajectory obtained in a previous application (see Scenario 1). Fig. 7-b shows the optimal solution on the opposite side of the Pareto front having a maximum target intersection length (1.886 km) and a path length of 6.636 km.



(a)



(b)

Fig. 7. Scenario 4: Optimized trajectory corresponding to the end points of the Pareto front: minimum target intersection length (a), maximum target intersection length (b)

A summary of population sizes, iteration numbers and computational times required for all test-cases is reported in Tab. 2. As we can see, gradually increasing the operational scenario complexity (number of no-fly areas and way-points) noticeable increment of the computational time is required to solve single-objective optimization tasks (Scenarios 1 to 3). On the other hand, multi-objective optimization procedure (Scenario 4) is the most demanding one in terms of required computational resources even if a very simple operational scenario is used.

Table 2. Required computational resources

Scenario	**Swarm size**	**Iterations**	**Computational Time** (Pentium 4, CPU 2.8 GHz, RAM 512 MB)
1	100	100	About 3 minutes
2	100	100	About 5.5 minutes
3	100	100	About 13 minutes
4	500	600	About 180 minutes

6. 3-D PROBLEM FORMULATION

To make a preliminary assessment on the algorithm capability to solve the optimum trajectory identification problem also in a three-dimensional environment, an extension of the formulation shown in paragraph 3 has been developed. We define four additional manoeuvres corresponding to two digits binary codes, describing possible changes in the flight path angle, γ:
1) Climb (1 0); 2) Descent (0 1); 3) No path angle change (0 0); 4) Aircraft alignment to the target (1 1).
We introduce two additional integer variables (*VP3m*, *VP4m*) whose binary codification provides the two *n*-dimensional vectors representing the γ manoeuvres sequence (see Table 3 for an example).

Table 3. Example of manoeuvres sequence (path angle)

VP3m= 523	*VP4m*= 27	
1	0	Climb
0	0	No path angle change
0	0	No path angle change
0	0	No path angle change
0	0	No path angle change
0	1	Descent
1	1	Alignment to the target
0	0	No path angle change
1	1	Alignment to the target
1	1	Alignment to the target

In case of a climb (or descent) flight phase a proper path angle has to be selected. Likewise the turn radius, we assume that the path angle, γ, is a linear function of the distance, d_i ,between the *i*-th path segment starting point and the nearest obstacle as shown in Fig. 8. This way the nearest is the obstacle, the steepest is the climb (or descent) path angle and vice versa. The flight path angle is bounded above by the user-defined parameter γ_{max}. During the optimization process, the line slope is tuned by two additional real design variables, *VP1a* and *VP2a*.

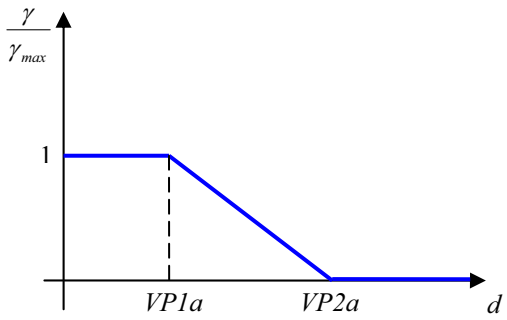


Fig. 8. Normalized path angle variation with obstacle distance *d*

Since in the 3-D case we use two couples of manoeuvre related variables (*VP1m*, *VP2m* and *VP3m*, *VP4m*), it is worth to notice that the binary sequence (1 1), provided by one of the two couples of variables, always takes priority over any binary sequence provided by the other two variables.

The design variables set representing a 3-D path turns out to be composed of the following variables:

$$\mathbf{x} = \left[VP1m, VP2m, VP3m, VP4m, VPr_l, VP1r^a, VP2r^a, VP1r^t, VP2r^t, VP1a, VP2a, \sigma^T, V, \Delta s \right]^T \quad (9)$$

7. NUMERICAL RESULTS: 3-D SINGLE-OBJECTIVE

Path length is selected as the objective function to be minimized. We assume n=10 as path segments maximum number. Proper penalty functions are defined degrading the fitness value whenever one or more constraints are violated.

Scenario 5. We consider two cylindrical no-fly areas assuming a spherical destination area centred at point $(5,5,0.4)$ with a radius of 0.1 km. SP is placed at (1.5, 1.5,0) whereas the initial velocity vector has a path angle γ=0 deg. and a yaw angle ψ=63 deg. We select a population size of 100 particles, and a maximum of 100 iterations for PSO.
Fig. 9 shows the trajectory obtained at the end of the optimization process.

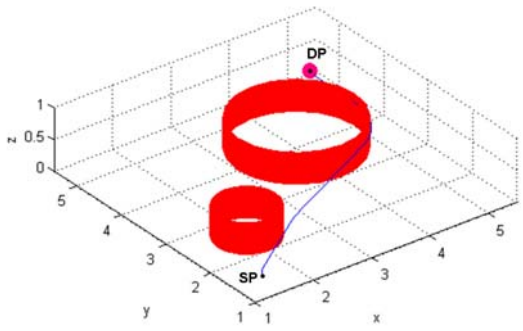


Fig. 9. 3-D trajectory optimization (Scenario 5).

Scenario 6. We consider only one cylindrical no-fly area assuming a spherical destination area centred at point $(0,0.6,0.95)$ with a radius of 0.1 km. SP is placed at (0, -0.5, 0) whereas the initial velocity vector has a path angle γ=0 deg. and a yaw angle ψ=0 deg. We select a population size of 100 particles, and a maximum of 100 iterations for PSO.
Fig. 10 shows the trajectory obtained at the end of the optimization process.

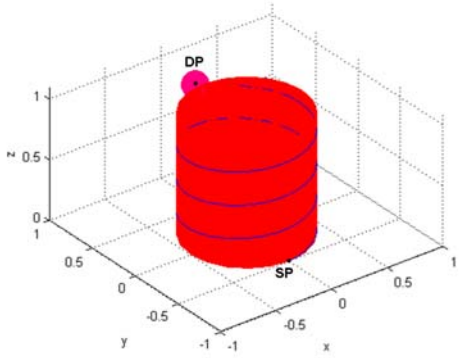


Fig. 10. 3-D trajectory optimization (Scenario 6).

As we can see in both cases the particle was able to reach the destination area defining a minimum length trajectory compliant with the operational constraints represented by no-fly areas.

REFERENCES

Betts, J. T. (1998). Survey of Numerical Methods for Trajectory Optimization. *Journal of Guidance, Control and Dynamics*, 21(2), 193–207.

Cen, Y., Wang, L. and Zhang, H. (2007). Real-time Obstacle Avoidance Strategy for Mobile Robot Based On Improved Coordinating Potential Field with Genetic Algorithm. *Proceedings of 16th IEEE International Conference on Control Applications*, Singapore, 415-419.

Dubins, L.E. (1957). On curves of minimal length with a constraint on average curvature, and with prescribed and terminal position tangents. *American Journal of Mathematics*, 79(3), 497-516.

Anderson, E., Beard, R. and McLain, T. (2005). Real-time dynamic trajectory smoothing for unmanned air vehicles. *IEEE Transactions on Control System Technology*, 13(3), 471–477.

Chitsaz, H., LaValle, S.M. (2007). Time-optimal Paths for a Dubins airplane. *Proceedings of 46th IEEE Conference on Decision and Control (CDC)*, New Orleans, LO, USA, 2379-2384.

Dever, C., Mettler, B., Feron, E., Popovic, J. and McConley, M. (2006). Nonlinear Trajectory Generation for Autonomous Vehicles via Parameterized Maneuver Classes. *Journal of Guidance, Control and Dynamics*, 29(2), 289-302.

Hu, X.B., Wu, S.F. and Jiang, J. (2004). On-line free-flight path optimization based on improved genetic algorithms. *Engineering Application of Artificial Intelligence*, 17(8), 897-907.

Eberhart, R.C. and Shi, Y. (1998). Comparison Between Genetic Algorithm and Particle Swarm Optimization. In Porto, V.W., Saravanan, N., Waagen, D., Eiben, A.E. (ed.), *Evolutionary Programming VII, 1447, Lecture Notes in Computer Science*, 611-616, Springer-Verlag, Berlin.

Raja, P. and Pugazhenhi, S. (2009). Path Planning for Mobile Robots in Dynamic Environments using Particle Swarm Optimization. *Proceedings of International Conference on Advances in Recent Technologies in Communication and Computing*, Kottayam, Kerala, India, 401-405.

Wang, L., Liu, Y., Deng, H. and Xu, Y. (2006). Obstacle-avoidance Path Planning for Soccer Robots Using Particle Swarm Optimization. *Proceedings of IEEE International Conference on Robotics and Biomimetics*, Kunming, China, 1233-1237.

Engelbrecht, A.P. (2005). *Fundamentals of Computational Swarm Intelligence*, 93-129. John Wiley & Sons, U.K.

Kennedy, J., Eberhart, R.C and Shi, Y. (2001). A Model of Binary Decision. *Swarm Intelligence*, 289-309. Morgan Kaufmann Publishers, San Francisco, CA.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

A Fault Detection Filter Design Method for Hybrid Switched Linear Parameter Varying Systems

G. Gagliardi * **A. Casavola** * **D. Famularo** * **G. Franzè** *

* *Università degli Studi della Calabria, Rende (CS), 87036, ITALY (e-mail:* {ggagliardi,casavola,famularo,franze}@deis.unical.it)

Abstract: In this paper a fault detection (FD) filter design method is proposed for hybrid switched linear parameter-varying (LPV) systems. The FD filter is designed as a bank of H_∞ Luenberger observers, achieved by optimizing frequency conditions which ensure guaranteed level of disturbance rejection and fault sensitivity. The switching signal is assumed to be known and satisfying a dwell-time prescription on the allowable switching sequences which ensures the asymptotical stability of the switched LPV system. The design method is recast as a Semidefinite Programming Problem in the observer bank gains. A FD threshold logic is also proposed in order to reduce the generation of false alarms. A practical example from lateral vehicle dynamics is provided to illustrate the effectiveness of the proposed technique.

Keywords: Hybrid Systems, Linear Parameter Varying Systems, Robust Fault Detection, Linear Matrix Inequalities.

1. INTRODUCTION

Fault Detection and Isolation techniques are important topics in systems engineering from the viewpoint of improving the system reliability. A fault represents any kind of malfunction in a plant that leads to anomalies in the overall system behavior. Such an event may happen due to process, sensors and/or actuator failures inside the plant.

During the last decade, model based fault detection (FD) technologies have attracted much attention (Chen and Patton [1999], Frank *et al.* [1997], Patton *et al.* [2000]). Starting from the rich theoretical results and increasing industrial applications it is well known that model-based fault detection can be approached as an output estimation problem and leads to a multi-objective design problem. In order to ensure a quick and reliable detection of faults, both the *robustness* of the FD systems to model uncertainties, unknown disturbance and its *sensitivity* to faults must be taken into consideration.

In the context of uncertain linear time-invariant (LTI) systems, a number of approaches have been proposed for the design of FD filters (see Frank *et al.* [1997], Patton *et al.* [2000], Rambeaux *et al.* [2000], Casavola *et al.* [2007] and references therein for a relevant bibliography on the matter). Nonetheless, a huge number of plants exhibit switching phenomena (see Dayawansa and Marlin [1999], Zefran and Burdick [1998]) which can be described by means of a hybrid model paradigm.

A hybrid model characterizes a system composed by both continuous and discrete components. The former are typically associated to physical variables and dynamics, the latter with logic devices, such as switches, digital circuitry, software code. Switched systems paradigms have in fact a lot of applications in control of mechanical systems, automotive industry, aircraft and air traffic control, electrical power converters and many other fields. To capture the evolutions of these systems, mathematical models need to combine, in one or another way, both continuous and discrete dynamics in all kinds of variations. However, they basically consist of a mix of differential or difference equations on one hand and automata, discrete-event models or Petri Nets on the other hand (van der Schaft and Schumacher [2000], Koutsoukos and Antsaklis [2003], Hamdi *et al.* [2009]).

Here, a model-based FD approach for hybrid systems is addressed by exploiting the theory of switching observers, whose literature is rich, e.g. see (Hamdi *et al.* [2009], Pettersson [2005, 2006], Alessandri and Coletta [2001], Alessandri *et al.* [2005]). Moving from the previous considerations, the purpose of this paper is to discuss a FD frequency design $\mathcal{H}_\infty/\mathcal{H}_-$ procedure for systems described by a class of linear time-varying models, whose structure jointly commutes according to an external switching signal $\sigma(t)$ which depends linearly on functions of measurable parameters $\theta(t)$ (H-LPV systems), Lim and Chan [2003]. We will suppose that both the measurable parameters and the switching signal are available at each instant time.

The FD filter consists here in a bank of time-varying Luenberger observers, doubly tuned with respect to both type of parameters, whose gains are to be found. It will be proved that the $\mathcal{H}_\infty$ disturbance decoupling requirement and $\mathcal{H}_-$ fault sensitivity performance can be easily turned into LMIs in the observer gains and that the *H-LPV* FD design problem is then solvable by means of standard semidefinite programming procedures. In order to ensure the stability to the overall FD setup, a *dwell-time condition* on the transition between two consecutive switching events is assumed. The proposed condition extends to the LPV case a previous result proved in a switched LTI framework by Abdo et al. [2010].

As a standard in FD schemes, a decision logic in charge to minimize the rate of false alarms generation is determined via frequency-domain conditions on *fictitious* doubly-indexed fault/disturbance vs. residual maps (transfer functions), each of them representing a specific transfer function of a LTI system corresponding to a vertex of the polytope family of plants.

Finally, a numerical experiment on a lateral vehicle dynamics with a faulty actuator is finally reported and described in details.

NOTATIONS

Given a symmetric matrix $A = A^T \in \mathbb{R}^{n\times n}$ we denote respectively with $\lambda_m(A)$ and $\lambda_M(A)$ the miniminum and maximum eigenvalues.

2. PROBLEM STATEMENT

Consider the class of multi-model linear possibly time-varying systems whose system matrices depend linearly on a time-varying parameter vector $\theta(t)$ and on a switching signal $\sigma(t)$. We will suppose also that both parameters may act as scheduling terms. This class of systems is described by

$$\begin{cases} \dot{x}(t) = A_{\sigma(t)}(\theta(t))x(t) + B_{\sigma(t)}(\theta(t))u(t) + E_{\sigma(t)}(\theta(t))f(t) + \\ \qquad G_{\sigma(t)}(\theta(t))d(t) \\ y(t) = C_{\sigma(t)}(\theta(t))x(t) + D_{\sigma(t)}(\theta(t))u(t) + F_{\sigma(t)}(\theta(t))f(t) + \\ \qquad H_{\sigma(t)}(\theta(t))d(t) \end{cases} \tag{1}$$

where

- $x(t) \in \mathbb{R}^n$ denotes the state, $u(t) \in \mathbb{R}^m$ the control input and $y(t) \in \mathbb{R}^p$ the measured output;
- $f(t) \in \mathbb{R}^{n_f}$ denotes the fault signal, $d(t) \in \mathbb{R}^{n_d}$ the exogenous disturbance;

In what follows the following conditions will be assumed:

- the fault signal $f(t)$ and the disturbance $d(t)$ are finite energy signals with radius norms belonging to the following proper subsets of L_2, i.e.
$$\Omega_f \triangleq \left\{ f(\cdot) \,|\, \exists \varepsilon_f > 0 \text{ s.t. } \sqrt{\int_0^\infty \|f(t)\|_2^2 \, dt} \le \varepsilon_f \right\},$$
$$\Omega_d \triangleq \left\{ d(\cdot) \,|\, \exists \varepsilon_d > 0 \text{ s.t. } \sqrt{\int_0^\infty \|d(t)\|_2^2 \, dt} \le \varepsilon_d \right\}$$
- the switching signal $\sigma(t) \in \{1,\dots,N\}$ characterizes the sudden transition of the plant structure and is supposed to be available at each time instant. Assuming N logical states $i \in \{1,\dots,N\}$ existing, $\sigma(t)$ is piecewise-constant taking one this integer values at each time instant, viz.
$$\sigma(t) \in \{1,\dots,N\} \tag{2}$$
- $\theta(t) \in \mathbb{R}^l$, is a possibly time-varying parameter which is known to belong to a given simplex
$$\Theta \triangleq \{\theta \in \mathbb{R}^l \mid \sum_{i=1}^{l} \theta_i = 1,\ 0 \le \theta_i \le 1,\ \ i = 1,\dots,l\}$$
and is supposed to be measurable. Notice that, the family of systems (1) consists of a finite set of possibly time-varying models. Then, the system matrices
$$\begin{bmatrix} A_{\sigma(t)}(\theta(t)) & B_{\sigma(t)}(\theta(t)) & E_{\sigma(t)}(\theta(t)) & G_{\sigma(t)}(\theta(t)) \\ C_{\sigma(t)}(\theta(t)) & D_{\sigma(t)}(\theta(t)) & F_{\sigma(t)}(\theta(t)) & H_{\sigma(t)}(\theta(t)) \end{bmatrix} = \sum_{j=1}^{l} \theta^j(t) \begin{bmatrix} A_i^j & B_i^j & E_i^j & G_i^j \\ C_i^j & D_i^j & F_i^j & H_i^j \end{bmatrix} \tag{3}$$
can be defined as a convex combination of the double-indexed matrices
$$\begin{bmatrix} A_i^j & B_i^j & E_i^j & G_i^j \\ C_i^j & D_i^j & F_i^j & H_i^j \end{bmatrix}, \quad \begin{matrix} i = 1,\dots,N, \\ j = 1,\dots,l \end{matrix}$$
being known constant matrices.

Given such a plant model we want to design a fault detection (FD) system such that:

(1) the effects of the process input signal and disturbances are minimized;
(2) the effects of the faults are properly enhanced;
(3) false alarm occurrences are minimized.

On the basis of the previous requirements such a filter must be sensitive with respect to failures, viz. capable to distinguish between faults and unknown disturbances. The design must be accomplished by means of a suitable residual evaluation function which must be close to zero in fault-free conditions and must deviate significantly when a failure occurs.

Thanks to the hypotheses that the plant parameter is measurable and the switching signal is directly available, the idea is to build a residual generator as a switched time-varying observer,

$$\begin{cases} \dot{\hat{x}}(t) = A_{\sigma(t)}(\theta(t))\hat{x}(t) + B_{\sigma(t)}(\theta(t))u(t) + \\ \qquad L_{\sigma(t)}(\theta(t))\,(y(t) - \hat{y}(t)) \\ \hat{y}(t) = C_{\sigma(t)}(\theta(t))\hat{x}(t) + D_{\sigma(t)}(\theta(t))u(t) \\ r(t) = y(t) - \hat{y}(t) \end{cases} \tag{4}$$

which can be characterized as a bank of N time-varying observers by considering all possible realizations of the switching signal. The filter gain $L_{\sigma(t)}(\theta(t))$, which is the design parameter of the diagnostic observer, has the following structure

$$L_{\sigma(t)}(\theta(t)) = L_i(\theta(t)) = \sum_{j=1}^{n_\theta} L_i^j \theta^j(t) \tag{5}$$

when the switching signal is equal to $\sigma(t) = i$. The gain matrices $L_i^j \in \mathbb{R}^{p\times n}$ are derived for each couple i, j so that the problem prescriptions are satisfied. When the residual generator (4) is applied to the plant (1), the estimation error $e(t) \triangleq x(t) - \hat{x}(t)$ and the residual $r(t)$ are governed by the following equation

$$\begin{cases} \dot{e}(t) = \left(A_{\sigma(t)}(\theta(t)) - L_{\sigma(t)}(\theta(t))C_{\sigma(t)}(\theta(t))\right)e(t) \\ \qquad + \left(E_{\sigma(t)}(\theta(t)) - L_{\sigma(t)}(\theta(t))G_\xi(\theta(t))\right)f(t) \\ \qquad + \left(F_{\sigma(t)}(\theta(t)) - L_{\sigma(t)}(\theta(t))H_{\sigma(t)}(\theta(t))\right)d(t) \\ r(t) = C_{\sigma(t)}(\theta(t))e(t) + G_{\sigma(t)}(\theta(t))f(t) + H_{\sigma(t)}(\theta(t))d(t) \end{cases} \tag{6}$$

The filter (4), for each discrete state $\sigma(t) = i$, must result asymptotically stable and designed so as to minimize the disturbance effects and enhancement of the fault sensitivity. In order to reduce the occurrence of false alarms, the obtained residuals are then processed by means of an appropriate index J_r, and then evaluated by using a decision logic which acts according to the following rules

$$J_r < J_{th}, \ \text{ for } f(t) = 0 \tag{7}$$

$$J_r \ge J_{th}, \ \text{ for } f(t) \ne 0 \tag{8}$$

Fault-detection decisions are based on the evaluation of the characteristics of the residual signals. To this end, the following frequency domain evaluation function is introduced

$$J_r \triangleq \left(\frac{1}{2\pi(\omega_s - \omega_i)} \int_{\omega_i}^{\omega_s} r^*(j\omega)\, r(j\omega)\, d\omega \right)^{\frac{1}{2}} \tag{9}$$

The frequency window $[\omega_i,\ \omega_s]$ is *a-priori* selected by the designer, even if a suitable choice could increase the robustness and the fault detection capabilities of the residual observer.

3. HYBRID LPV STABILITY CONDITIONS

In what follows, a dwell-time condition which ensures the stability of the overall switched system, provided that the individual LPV subsystems are quadratically stable, is presented. The result here outlined is a direct extension to the LPV case of similar conditions derived in Abdo et al. [2010] for switching LTI plants. The idea is that asymptotical stability can be proved if the switching rate is sufficiently slow and the transient effects occurring after each switch are dissipated. To this end, the following definitions of *dwell-time* and *average dwell-time* are of interest (see Liberzon *et al.* [1999], Hespanha *et al.* [1999], Morse [1996] for a detailed analysis on the matter).

Definition 1. ***Dwell-Time****. Let $t_1, t_2, \ldots, t_N$ be the switching time instants for the switching signal $\sigma(t)$. Then, the switching system has a dwell-time $\tau_d > 0$ if it satisfies $t_{i+1} - t_i \geq \tau_d$ for all i.*

A lower-bound on τ_d can be explicitly calculated from the exponential decay bounds derived from quadratic stability checks on the LPV matrices of the individual subsystems corresponding to each *i*-th logical state. The stability of *slow-switching* LPV systems can be ensured if the interval between any two consecutive switching is no smaller than τ_d. This result comes from the following Lemma which is a simple extension to the LPV case of a Lemma proved by Morse [1996], Abdo et al. [2010] for the switching LTI framework:

Lemma 1. Let $\{A_p(\theta) : p \in \mathbb{P},\ \theta \in \Theta\}$ be a closed, bounded set of real, $n \times n$ matrices such that, for a given value of p and θ, $A_p(\theta) \in \text{co}\{A_p^1, \ldots, A_p^l\}$. Suppose that for each $p \in \mathbb{P}$, the LPV system

$$\dot{x}(t) = A_p(\theta(t))x(t)$$

is quadratically stable and let a_p and λ_p be any finite, nonnegative and positive numbers, respectively, for which

$$\max_{j=1,\ldots,l} \left| e^{A_p^j t} \right| \leq e^{a_p - \lambda_p t}, \quad t \geq 0 \tag{10}$$

Suppose that τ_d is a number satisfying

$$\tau_d > sup_{p\in\mathbb{P}} \left\{ \frac{a_p}{\lambda_p} \right\} \tag{11}$$

For any admissible switching signal $\sigma(t) : [0, \infty) \to \mathbb{P}$ with dwell-time no smaller than τ_d, the state transition matrix of $A_{\sigma(t)}$ satisfies

$$|\Phi(t, \mu)| \leq e^{(a - \lambda(t-\mu))}, \quad \forall t \geq \mu \geq 0 \tag{12}$$

where

$$\begin{aligned} a &= \quad sup_{p\in\mathbb{P}} \{a_p\} \\ \lambda &= inf_{p\in\mathbb{P}} \left\{ \lambda_p - \frac{a_p}{\tau_d} \right\} \end{aligned} \tag{13}$$

Moreover

$$\lambda \in (0, \lambda_p], \quad p \in \mathbb{P} \tag{14}$$

Thus, if the switching signal $\sigma(t)$ "dwells" at each of its value $\sigma(t) = 1, 2, ..., N$ long enough for the norm of the state transition matrix A_p, to drop to at least τ_d time units, then the hybrid system $\ldots x(t) = A_{\sigma(t)}(\theta(t))x(t)$ is exponentially stable having a decay rate λ which is upper bounded by the smallest of the decay rates of the LPV system collection $\dot{x}(t) = A_p(\theta(t))x(t)$, $p \in \mathbb{P}$

Definition 2. ***Average Dwell-time****. Let $N_\sigma(T, t)$ be the number of discontinuities of the switching signal $\sigma(t)$ on the interval (t, T). Assume there exist two positive numbers N_o and τ_a such that*

$$N_\sigma(T, t) \leq N_o + \frac{T - t}{\tau_a}, \quad \forall T \geq t \geq 0 \tag{15}$$

where N_o is the chatter bound. Then, τ_a is the average dwell-time of $\sigma(t)$.

Note that the respect of either the *dwell-time* or the *average dwell-time* conditions implies in practice to keep active each observer of the bank for at least τ_d or τ_a time instants in the FD filter before a switch to a different observer of the bank could take place.

3.1 Multiple Lyapunov Functions Stability

The use of multiple Lyapunov functions is a useful tool for proving stability of a switched systems, (Hespanha [2004], Zhai *et al.* [2000, 2004]). Consider the hybrid switched linear system

$$\dot{x}(t) = A_{\sigma(t)}(\theta(t))x(t) \tag{16}$$

We assume that all LPV subsystems of (16) are quadratically stable. Note that the stability of all subsystems is not sufficient to ensure the stability for the whole system.

If we can find a positive λ_i such that $A_i^j + \lambda_i I$, $j = 1, \ldots, l$ [1], is still Hurwitz stable ($A_\sigma(\theta) = A_i(\theta)$, when $\sigma = i$) then there are symmetric positive definite matrices $P_1, \ldots, P_N$ such that

$$(A_i^j + \lambda_i I)^T P_i + P_i (A_i^j + \lambda_i I) < 0, \ j = 1, \ldots, l \tag{17}$$

By using the solution P_i of (17), one can ensure the stability of the switched system (16) by introducing a Multiple Lyapunov Functions (MLF) candidate

$$V_\sigma(t) = x(t)^T P_{\sigma(t)} x(t) \tag{18}$$

and exploiting the following properties (for a proof see e.g. Morse [1996]):

(1) $\dot{V}_i \leq -2\lambda_i V_i$
(2) there exists constant scalars $\alpha_2 \geq \alpha_1 > 0$ such that
$$\alpha_1 \|x\|^2 \leq V_i(t) \leq \alpha_2 \|x\|^2, \ \forall x \in \mathbb{R}^n, \forall i \in \{1, \ldots, N\}$$
(3) there exists a constant scalar $\mu \geq 1$ such that
$$V_i(t) \leq \mu V_j(t), \ \forall x \in \mathbb{R}^n, \forall i, j \in \{1, \ldots, N\}$$

The first property is a straightforward consequence of (18), while the second and the third hold for

$$\alpha_1 = \inf_{i\in\{1,\ldots,N\}} \lambda_m(P_i) \quad \alpha_2 = \sup_{i\in\{1,\ldots,N\}} \lambda_M(P_i) \tag{19}$$

$$\mu = \frac{\alpha_2}{\alpha_1} \tag{20}$$

The dwell-time can be computed as $\tau_d = \frac{ln\mu}{2(\lambda^* - \lambda)}$, whit $\lambda \in (0, \min_i(\lambda_i))$ and $\lambda^* \in (\lambda, \min_i(\lambda_i))$.

4. RESIDUAL GENERATOR DESIGN

The design of the residual observer (5) is accomplished by solving a multi-objective optimization problem. Starting from the discussion of the above Section, it is possible to restate the fault detection design problem as follows:

H-LPV-FD

[1] Note that it is always possible to choose λ_i because this quantity is upper bounded by $\max_{j=1,\ldots,l} \frac{1}{2} \left| \lambda_m(A_i^{T,j} + A_i^j) \right|$.

Given a Lyapunov function

$$V_\sigma(t) = x^T(t) P_{\sigma(t)} x(t), \tag{21}$$

find observer matrix gains $L_i^j \in \mathbb{R}^{n\times p}$, $i = 1,\ldots,N$, $j = 1,\ldots,l$ *and two positive scalars* α *and* β *solutions of the following LMI optimization problem*

$$\min_{L_i^j \in \mathbb{R}^{n\times p}} a_1\,\alpha^2 - a_2\,\beta^2 \tag{22}$$

s.t.

$$\int_0^\infty \left[\frac{dV_\sigma(t)}{dt}\right] dt < \int_0^\infty \alpha^2 \|d(t)\|_2^2 - \|r_d(t)\|_2^2\, dt \tag{23}$$

$$-\int_0^\infty \left[\frac{dV_\sigma(t)}{dt}\right] dt > \int_0^\infty \beta^2 \|f(t)\|_2^2 - \|r_f(t)\|_2^2\, dt \tag{24}$$

where a_1, a_2 are proper positive weights, $r_d(t)$ the fault-free residual evolution associated to

$$\begin{cases} \dot{e}_d(t) = \left(\hat{A}_\sigma(\theta) - L_\sigma(\theta)\hat{C}_\sigma(\theta)\right) e_d(t) + \\ \qquad\qquad \left(\hat{E}_\sigma(\theta) - L_\sigma(\theta)\hat{F}_\sigma(\theta)\right) d(t) \\ r_d(t) = \hat{C}_\sigma(\theta)\, e_d(t) + \hat{F}_\sigma(\theta)\, d(t) \end{cases} \tag{25}$$

whereas $r_f(t)$ the disturbance-free residual associated to

$$\begin{cases} \dot{e}_f(t) = \left(\tilde{A}_\sigma(\theta) - L_\sigma(\theta)\tilde{C}_\sigma(\theta)\right) e_f(t) + \\ \qquad\qquad \left(\tilde{G}_\sigma(\theta) - L_\sigma(\theta)\tilde{H}_\sigma(\theta)\right) f(t) \\ r_d(t) = \tilde{C}_\sigma(\theta)\, e_f(t) + \tilde{H}_\sigma(\theta)\, f(t) \end{cases} \tag{26}$$

The hatted ($\hat{\cdot}$) and tilded ($\tilde{\cdot}$) variables denote a representation where the respective residuals $r_d(t)$ and $r_f(t)$ have been processed by proper window filters $Q_d(s)$ and $Q_f(s)$ characterizing the disturbance and fault effects in specific frequency ranges of interest. Inequalities (23) and (24) characterize the usual trade-off between disturbance decoupling and minimum fault sensitivity achievements usually addressed in the $\mathcal{H}_\infty/\mathcal{H}_-$ approach.

The *H-LPV-FD* design problem can be turned into a Linear Matrix Inequality optimization procedure and the following Proposition reports explicitly the LMI conditions under which problem *H-LPV-FD* can be checked for admitting a solution:

Proposition 1. The inequalities (23) and (24) are satisfied if there exit a family of matrices $P_i = P_i^T$, $i = 1,\ldots,N$ *and matrix gains* $K_i^j \in \mathbb{R}^{n\times p}$, $i = 1,\ldots,N$, $j = 1,\ldots,l$ *such that the following* $2lN$ *linear matrix inequalities*

$$\begin{bmatrix} \mathrm{He}\left(P_i\hat{A}_i^j - K_i^j\hat{C}_i^j\right) + \hat{C}_i^{T,j}\hat{C}_i^j & P_i\hat{G}^i - K_i^j\hat{H}_i^j + \hat{C}_i^{T,j}\hat{H}_i^j \\ * & \hat{H}_i^{T,j}\hat{H}_i^j - \alpha^2 I \end{bmatrix} \preceq 0 \tag{27}$$

$$\begin{bmatrix} \mathrm{He}\left(P_i\tilde{A}_i^j - K_i^j\tilde{C}_i^j\right) + \tilde{C}_i^{T,j}\tilde{C}_i^j & P_i\tilde{E}^i - K_i^j\tilde{F}_i^j + \tilde{C}_i^{T,j}\tilde{F}_i^j \\ * & -\tilde{F}_i^{T,j}\tilde{F}_i^j - \beta^2 I \end{bmatrix} \preceq 0 \tag{28}$$

($He(X) := X + X^T$*),* $i = 1,\ldots,N$, $j = 1,\ldots,l$, *hold true with* $L_i^j = P_i^{-1}K_i^j$, $i = 1,\ldots,N$, $j = 1,\ldots,l$.

Remark 1 - It is worth pointing out that the set of linear matrix inequalities (27) and (28) has been obtained by assuming the joint availability of the switching signal $\sigma(t)$ and the plant parameter $\theta(t)$. Notice also that the existence, at each time instant t, of a family of observer gains for each mode i

$$L_i(\theta) = \sum_{j=1}^{l} \theta_j(t) L_i^j$$

solutions of *H-LPV-FD*, implies the observability of each doubly indexed couples $(\hat{A}_i^j, L_i^j)$ (disturbance rejection) and $(\tilde{A}_i^j, L_i^j)$, (fault sensitivity) $i = 1,\ldots,N$, $j = 1,\ldots,l$. □

5. THRESHOLD COMPUTATION

The threshold decision logic is based on the evaluation of the quantity J_{th} (eqs. (7) and (8)) in order to minimize the occurrence of false alarms. Such a term is derived in *fault free conditions* according to the time function residual evaluation

$$J_r(t) \triangleq \sqrt{\frac{1}{2\pi(\omega_s - \omega_i)} \int_{\omega_i}^{\omega_s} r_t^*(j\omega)\, r_t(j\omega)\, d\omega} \tag{29}$$

where $r_t(j\omega)$ denotes the Fourier Transform of the residual signal up to time t. From a computational point of view, $J_r(t)$ can be easily obtained by means of standard highly efficient Fast Fourier Transform (FFT) algorithms, available with MATLAB©. Given $J_r(t)$, the threshold can be computed by defining the following doubly indexed family of LTI systems

$$G_{ij}(s) \triangleq \hat{C}_i^j \left(sI - \left(\hat{A}_i^j - L_i^j\hat{C}_i^j\right)\right)^{-1} \left(\hat{G}_i^j - L_i^j\hat{H}_i^j\right) + \hat{H}_i^j, \quad i = 1,\ldots,N,\ j = 1,\ldots,l. \tag{30}$$

Due to the fact that the disturbance $d(t)$ is a finite energy signal we also have

$$J_{th} \triangleq \sqrt{\frac{1}{2\pi(\omega_s - \omega_i)}} \sup_{d\in\Omega_d} \max_{\substack{i=1,\ldots,N \\ j=1,\ldots,l}} \left\|G_{ij}(s)\right\|_\infty \|d\|_2 \tag{31}$$

Then, a computable upper-bound to J_{th} can be easily achieved by observing that

(1) the scalar α, which is part of the solution of the LMI procedure obtained from problem *H-LPV-FD* is an upper bound to
$$\left\|G_{ij}(s)\right\|_\infty \le \alpha, \ \ \forall i,j \tag{32}$$
(2) the set Ω_d is upper-bounded, in term of energy norm, by ε_d

The consequence is that

$$J_{th} \le \sqrt{\frac{1}{2\pi(\omega_s - \omega_i)}}\alpha\varepsilon_d \tag{33}$$

in healthy situations. Note finally that, the process input $u(t)$ does not appear in the residual generation function and it is, as a consequence, a decoupled input. It is then reasonable to consider a fixed threshold J_{th} instead of computing a more involved (and not necessary here w.r.t. $u(t)$) adaptive residual evaluation function.

6. SIMULATION RESULTS

Simulation studies on a vehicle lateral dynamical system (see Figure 1) are carried out to illustrate the effectiveness of the proposed method to design an H-LPV filters for robust fault detection purposes via LMIs.

The model used here is the so-called one-track model, aka bicycle model. One-track models are derived upon the assumption that the vehicle is simplified as a point mass with the center of gravity on the ground, which can only move along the x, y axes, and yaw around the z axis. By considering the vehicle side slip

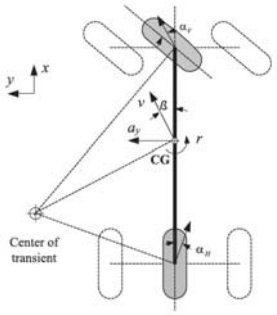


Fig. 1. Kinematics of one-track model

angle β and the yaw rate r to be the state variables, the lateral acceleration a_y and yaw rate r the output variables, the steering angle δ_L the input variable, the state space representation of the one track model is given by

$$\dot{x}(t) = Ax(t) + Bu(t) + Ef(t) + Gd(t), \;\; x(t) = \begin{bmatrix} \beta(t) \\ r(t) \end{bmatrix}$$
$$y(t) = Cx(t) + Du(t), \;\; y(t) = \begin{bmatrix} a_r(t) \\ r(t) \end{bmatrix}, \; u(t) = \delta_L(t) \tag{34}$$

where

$$A = \begin{bmatrix} -\frac{C_{\alpha_v} + C_{\alpha_H}}{mv} & \frac{l_H C_{\alpha_H} - l_V C_{\alpha_v}}{mv^2} - 1 \\ \frac{l_H C_{\alpha_H} - l_V C_{\alpha_v}}{I_z} & \frac{l_V^2 C_{\alpha_V} + l_H^2 C_{\alpha_H}}{I_z v} \end{bmatrix} \quad B = \begin{bmatrix} \frac{C_{\alpha_v}}{mv} \\ \frac{l_V C_{\alpha_v}}{I_z} \end{bmatrix}$$
$$C = \begin{bmatrix} -\frac{C_{\alpha_v} + C_{\alpha_H}}{m} & \frac{l_H C_{\alpha_H} - l_V C_{\alpha_v}}{mv} \\ 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} \frac{C_{\alpha_v}}{m} \\ 0 \end{bmatrix} \quad G = B$$

If the velocity v varies, the LTI model (34) is replaced with an LPV model in which the longitudinal velocity is a time varying parameter, (Rajamani [2006]). Furthermore, if we consider the velocity range $R = [10, 90]$ and we split the interval in three sub-ranges $R_1 = [10, 30)$, $R_2 = [30, 60)$ and $R_3 = [60, 90]$, the model (34) can be viewed as a three mode ($\sigma = 1, 2, 3$) hybrid switching linear parameter varying system. For each σ we assume that the longitudinal velocity is limited by the following constraints

$$v_\sigma^- \leq v_\sigma \leq v_\sigma^+ \quad \sigma = 1, 2, 3 \tag{35}$$

Moreover, we consider three categories of fault on the steering angle measurement δ_L, each one for a specific system mode $\sigma = 1, 2, 3$:

$$f_{\sigma=1}(t) = \{\, 0, \;\; t \leq 5s; \;\; 1, \;\; t > 5s;$$
$$f_{\sigma=2}(t) = \{\, 0, \;\; t \leq 25s; \;\; 1, \;\; t > 25s;$$
$$f_{\sigma=3}(t) = \{\, 0, \;\; t \leq 50s; \;\; 1, \;\; t > 50s;$$

The frequency window, where the residual signals $r(s)$ are evaluated, has been chosen equal to $[\omega_i, \omega_s] = [0, 25]$ rad/s. This choice corresponds to the following filters

$$Q_d(s) = \frac{\omega_s}{s^2 + 1.4\omega_s s + \omega_s^2}, Q_f(s) = \frac{s^2 + s + \omega_s}{s^2 + \omega_s s + \omega_s}$$

The dwell-time value, computed as in section (3.1), is $\tau_d = 1.4774$.

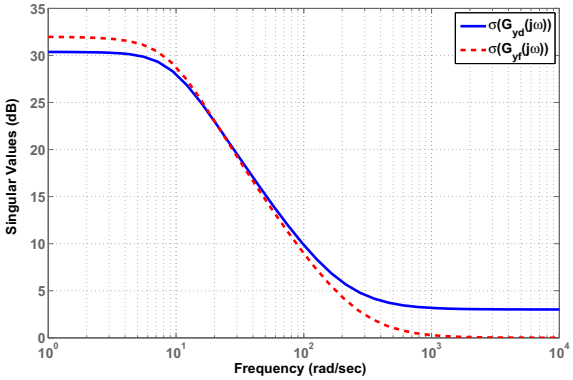


Fig. 2. Singular values plots of $G_{yd}(j\omega)$ and $G_{yd}(j\omega)$

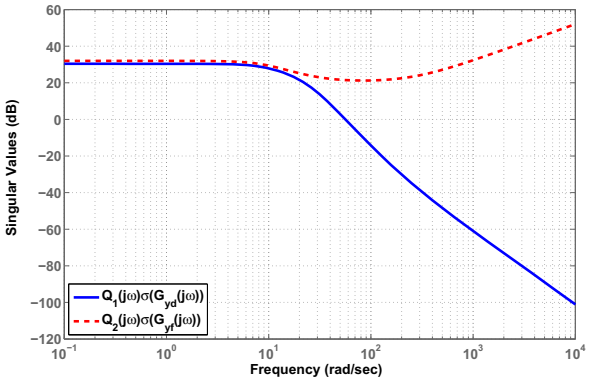


Fig. 3. Singular values plots of $Q_1(j\omega)G_{yd}(j\omega)$ and $Q_2(j\omega)G_{yd}(j\omega)$

The evolutions of system modes and relevant variables are reported in Figure 4. In this figure we can observe the effectiveness of the proposed solution.

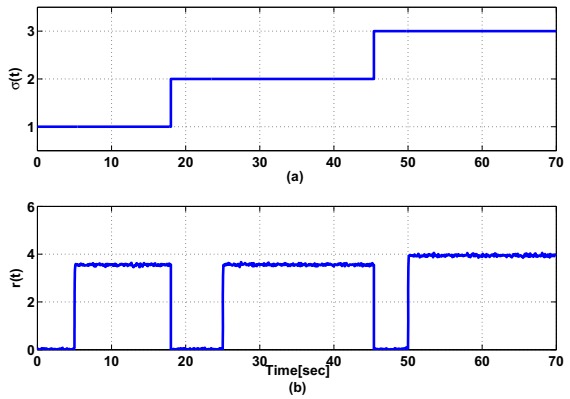


Fig. 4. Evolution of mode (a) and detection signal (b)

The threshold J_{th} (dashed line) and the $J_r(t)$ response (continuous line) are depicted in Figure 5. The filter seems to exhibit very good disturbance decoupling properties.

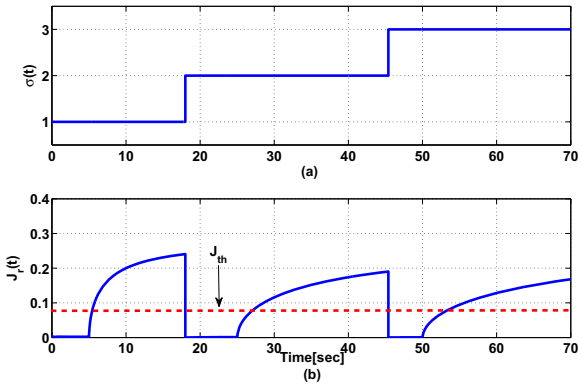


Fig. 5. Evolution of mode (a), threshold J_{th} (dashed line) and frequency-windowed norm $J_r(t)$ (continuous line) (b)

7. CONCLUSION

A novel Robust FD strategy for hybrid switched linear parameter-varying systems has been proposed. By taking advantage of the Multiple Lyapunov Functions stability concept and using congruence transformations, the FD design problem has been converted into a tractable LMI optimization problem. A fixed threshold logic has been proposed in order to to discriminate between real and false alarms. A numerical example showing the effectiveness of the proposed approach have been described in details where the results have shown good detection capabilities of the FD logic.

REFERENCES

Abdo A., Damlakhi W., Saijai J. and Ding S., "Design of Robust Fault Detection Filter for Hybrid Switched Systems", Proc. of the 2010 Conference on Control and Fault Tolerant Systems, Nice, France, 2010.

Chen J. and Patton R.J. "Robust Model-Based Fault Diagnosis for Dynamic Systems". Kluwer Academic Publishers: Boston, MA, USA, 1999.

Frank, P. M and Ding, X. Survey of robust residual generation and evaluation methods in observer-based fault detection systems. survey and some new results. *Journal of Process Control* **7(6)**,403–424, 1997.

Patton, R.J., Frank, P.M. and Clark, R.N. (Eds.) *Issues of Fault Diagnosis for Dynamic Systems*, Springer, 2000.

Rambeaux, F., Hamelin, F. and Sauter, D. Optimal thresholding for robust fault detection of uncertain systems. *International Journal of Robust and Nonlinear Control* **10**, 1155–1173, 2000.

Casavola A., Famularo D., Franzè G. and Sorbara M. "A fault-detection, filter-design method for linear parameter-varying systems", *Proc. IMechE - Journal of Systems and Control Engineering*, **221**, pp. 865-873, 2007.

Dayawansa W.P. and Marlin C.F., "A converse Lyapunov theorem for a class of dynamical systems which undergo switching" . IEEE Trans. Automat. Contr. vol. 44, pp. 751-760, 1999.

Zefran M. and Burdick J.W., "Design of switching controllers for systems with changing dynamics", Proc. of the 37th IEEE Conference on Decision and Control, pp. 2113-2118, 1998.

van der Schaft A. and Schumacher H. "An Introduction to Hybrid Dynamical Systems". Lecture Notes in Control and Information Sciences, vol. 251, Springer-Verlag, 2000.

Koutsoukos X.D. and Antsaklis P.J. "Hybrid Dynamical Systems: Review and Recent Progress". Software-Enabled Control: Information Technologies for Dynamical Systems, Wiley-IEEE Press, 2003.

Hamdi F., Manamanni N., Messai N., Benmahammed K. "Hybrid observer design for linear switched system via Differential Petri Nets". Nonlinear Analysis: Hybrid Systemsm Vol. 3, pp. 310-322, 2009.

Lim S. and Chan K. "Analysis of Hybrid Linear Parameter-Varying Systems". Proceedings of the American Control Conference, Denver, CO, USA, pp.4822-4827, 2003.

Liberzon D. and Morse A. S. "Basic Problems in Stability and Design of Switched Systems". IEEE Control Systems Magazine, Vol. 19, No. 5, pp. 59-70, 1999.

Decarlo R.A., Branicky M.S., Petterson S. and Lennartson B. "Perspectives and Results on the Stability and Stabilizability of Hybrid Systems". Proceeding of the IEEE, Vol. 88, No. 7, pp. 1069-1082, 2000.

Hespanha J.P.. "Uniform Stability of Switched Linear Systems: Extensions of LaSalles Invariance Principle". IEEE Transactions on Automatic Control. Vol. 49. pp, 470-482, 2004.

Liberzon D. "Stabilizing a linear system with finite-state hybrid output feedback". Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99), 1999.

Pettersson S. "Switched State Jump Observer For Switched Systems". Proceedings of the 16th IFAC World Congress, Prague, 2005.

Pettersson, S. "Designing Switched Observers For Switched System Using Multiple Lyapunov Functions and Dwell-Time". Preprints of the 2nd IFAC Conf. on Analysis and Design of Hybrid Systems (Alghero, Italy), 7-9 June 2006.

Alessandri A. and Coletta P. "Switching observers For Continuous-Time and Discrete-Time Linear Systems". Proceedings of the American Control Conference, Arlington VA, USA, June 25-27, pp. 2516-2521, 2001.

Alessandri A., Baglietto M., and Battistelli G. "Luenberger Observers for Switching Discrete-Time Linear Systems". Proceedings of the joint 44th IEEE Conference on Decision and Control - European Control Conference 2005, Seville, Spain, pp. 7014-7019, 2005.

Hespanha J. P. and Morse A. S. "Stability of Switched Systems with Average Dwell-Time". Proceedings of the 38th IEEE Conference on Decision and Control, Phoenix AR, USA, pp. 2655-2660, 1999.

Morse A. S. "Supervisory Control of Families of Linear Set Point Controllers - Part 1: Exact Matching". IEEE Transactions on Automatic Control, Vol. 41, No. 10, pp. 1413-1431, 1996.

Zhai G., Hu B., Yasuda K. and Michel A. N. "Piecewise Lyapunov Functions for switched Systems with Average Dwell-Time". Asian Journal of Control, Vol. 2, No. 3, pp. 192-197, 2000.

Zhai G., Lin H., Michel A. N. and Yasuda K. "Stability Analysis for Switched Systems with Continuous-Time and Discrete-Time Subsystems". Proceeding of the 2004 American Control Conference, Boston MA, USA, pp. 4555-4560, 2004

Rajamani R. "Vehicle Dynamics and Control", Springer Verlag, 2006.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Improvement of the Sensitivity of T² Quality Control Charts by Grouping of Variables

T. Friebel, R. Haber

Department of Process Engineering and Plant Design, Laboratory of Process Control,
Cologne University of Applied Science, D-50679 Köln, Betzdorfer Str. 2, Germany
e-mail: {Thomas.Friebel; Robert.Haber}@FH-Koeln.de
fax: +49/221/8275-2836

Abstract: With increasing number of variables the Hotelling's T^2 statistic can detect only larger failures in the variables. A new method is introduced for reducing the dimension of the Hotelling's statistic in order to detect smaller failures. The basic idea is to group some variables into a combined variable and to calculate the T^2 value from this grouped variable and from the remaining non-grouped variables. As the new calculated variable is not Gaussian distributed a proper static transformation is applied. Both uncorrelated and correlated data are dealt with. In the latter case principal component analysis is used before calculating T^2. Several simulations show the improvement of the new T^2 control chart.

Keywords: fault detection, T^2 control chart, sensitivity, grouping of variables

1. INTRODUCTION

Several (e.g. five) variables can be simultaneously monitored by using more (e.g. five) control charts. The advantage of this univariate analysis is fast calculation of control charts and simple parameterization. However, there are disadvantages like:

- The user has to check several control charts at the same time.
- Only the variances of the variables are taken into account, and the relations (covariances) between them are not considered.

Alternatively Hotelling's T^2 value or Mahalanobis distance D can be monitored as a single variable. If this value exceeds a prescribed limit then at least one of the variables exceeds its limit. The T^2 value is calculated from measurement vector $\mathbf{u}$, the mean vector $\overline{\mathbf{u}}$ and the covariance matrix $\mathbf{S}$ by (1), where k indicates the discrete time.

$$T_k^2 = D_k^2 = \left(\mathbf{u}_k - \overline{\mathbf{u}}\right)^T \mathbf{S}^{-1} \left(\mathbf{u}_k - \overline{\mathbf{u}}\right) \tag{1}$$

The mean value and the covariance matrix are usually estimated from a training data set. The quality of this estimation is relevant for the fault detection ability of the control chart. In practice it is often very difficult to separate a training data set which contains no disturbances or outliers. With the classical estimation method the estimation of variance and covariance is adversely affected by disturbances such as outliers, thus the sensitivity of the control chart decreases. [1] recommends median and MAD (Median Absolute Deviation) for robust estimation of these parameters. For the ongoing analysis it is assumed that the data has Gaussian distribution and the mean vector and the covariance matrix known, or they are estimated from a large amount of data. In this case the T^2 value of (1) underlies a χ^2 distribution. The control limit UCL for p number of variables can be calculated for a given confidence level α by

$$UCL = \chi^2_{\alpha;p} \tag{2}$$

The control limits are shown for different number of variables p and a probability $P = 0.9973$ in Fig. 1. The control limits are increased as the error of type I is increased with increasing number of variables p according to (3), see e.g. [2].

$$\alpha_{res} = 1 - (1-\alpha)^p \tag{3}$$

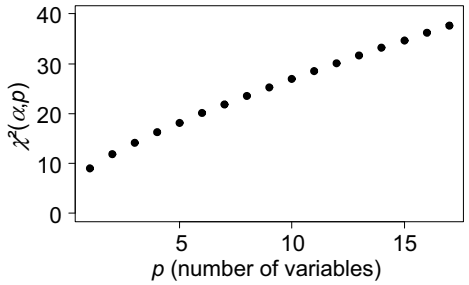


Fig. 1 Control limits for different number of variables p with a percentage $P = 99.73\%$

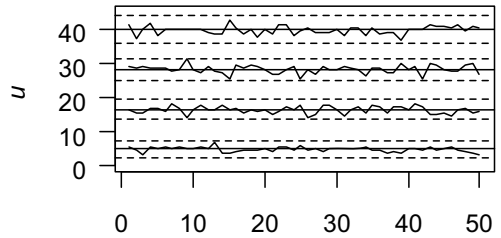
The Mahalanobis distance D is used as a measure of the deviation of the actual measurements from the fault-free state usually characterized by mean value. With increasing number of variables the detectable disturbance is always farther from the mean value. This means, the more variables are monitored the more difficult it becomes to detect small deviations from the normal state. Normally PCA is used and some (probably) not important principal components (PC's) are excluded from the analysis. For the choice which PC's contain relevant information several evaluation criteria such as

90% of the explained variance, Kaiser or elbow criterion are known from the literature. Disadvantage of these methods is, that deviations in the non-monitored PCs can not be detected. An alternative but more complicated way is to use *SPE* (Squared Prediction Error). From the important variables T^2 values are calculated and the other, unused variables are combined to *SPE*. As all the variables are considered, information is not lost. But it is necessary to observe two quality values, T^2 and *SPE*. A more detailed explanation can be found in [3]. However, evaluation of *SPE* is more difficult for a practitioner than using the T^2 control chart. The first method reduces the number of variables *p,* but looses some information. Here an alternative, new method is presented which provides easier calculation than the combined calculation of T^2 and *SPE* charts there is no information is loss. In this new procedure some variables are grouped to an additional, calculated variable. As the new variable not Gaussian distributed it is transformed to a Gaussian distribution. From this Gaussian variable and the other remaining, non-grouped variables the T^2 value is calculated and monitored in one control chart. As the number of the dimensions of the variables in the T^2 control chart is less than the total number of the variables, the failures in the non-grouped variables can be detected easier. Also the deviations in the grouped variables can be detected, even if they become smaller. Several methods exists how to interpret a T^2 signal. [2] and [4] try to decompose the variable T^2 into independent components. An overview of several methods can be found in [5]. But all these methods require a primary detection with T^2 control chart of an abnormal state. Therefore, it is important to improve the sensibility of the T^2 calculation as it is proposed in the present paper.

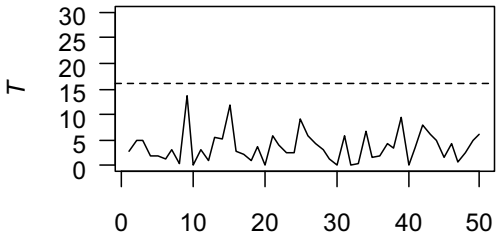
2. NEW METHOD FOR IMPROVEMENT OF THE SENSIBILITY OF T^2

2.1 Example

In Fig. 2a the measured data of four independent normally distributed variables are plotted. It can be seen that there is no value outside the control limits (dashed lines). The calculated T^2 values are drawn in Fig. 2b. None of the plots shows any abnormal condition. Fig. 3a shows almost the same data as in Fig. 2a. In Fig. 3a each variable is disturbed by a value of 4.03 times the standard deviation. All four values are above the upper control limits and are marked by circles. The remaining values are identical to those in Fig. 2a and lie within the control limits. The T^2 values calculated according to (1) are plotted in Fig. 3b. The disturbed values are marked again by circles. As it can be seen, they are located exactly on the control limit (dashed line) in all four cases. The control limit has a value of UCL = 16.25. That means all Mahalanobis distances D_k less than 4.03 are below the control limits. Therefore, for every sample a normal state is detected. In Fig. 3c a new T^2 value is shown. For this calculation the first two of the four variables are grouped to a new, normally distributed variable. The first two variables have a mean value of $\overline{u}_1$ = 5.00 and $\overline{u}_2$ = 16.66. The T^2 value is calculated by (1) from the new variable and the remaining non-grouped variables (variable 3 and 4). The last two variables

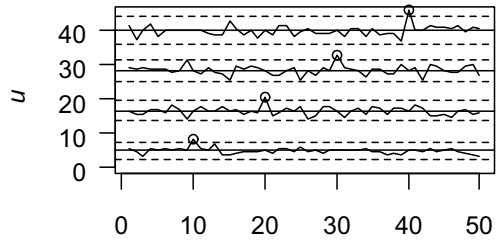


a) Shewhart control charts of the variables

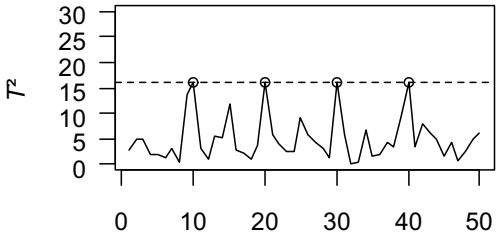


b) T^2 plot of the four variables

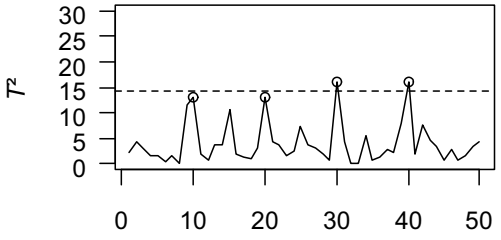
Fig. 2. Normal state without any failure



a) Shewhart control charts of the variables



b) T^2 plot of four variables (T^2 according to (1))



c) modified T^2 plot of four variables

Fig. 3. Four variables with always a single failure

have a mean value of $\overline{u}_3$ = 28.33 and $\overline{u}_4$ = 40.00. It can be seen, that the disturbances in the grouped variables at k = 10 and k = 20 are not detected. Only the disturbances in the non-grouped variables at k = 30 and k = 40 are above the control limit, with a value of UCL = 14.16. Therefore, the last two disturbances are recognized as abnormal. Now variable 1 and 2 have a Mahalanobis distance of D = 4.16 and for the variables 3 and 4 of D = 3.76. All new T^2 values with grouped variables are equal to the control limit UCL^* = 14.16. If the T^2 value is calculated by (1) without grouped variables, then the

last two disturbances at $k = 30$ and $k = 40$ with $T^2 = 14.16$ are below the control limit. The first two disturbances at $k = 10$ and $k = 20$ with $T^2 = 17.28$ are above the control limit $UCL = 16.25$. By comparing these two examples the following conclusions can be found:

- The residual subspace should be grouped.
- The sensitivity for detecting disturbances in the non-grouped variables is increased.
- The sensitivity for detecting disturbances in the grouped variables is decreased.
- A bad sensitivity for the residual subspace is indeed better than removing these variables with the calculation of T^2.

2.2 Principle of the new method

In the above example four variables are monitored. To increase the sensitivity of the control chart, $p' = 2$ variables are grouped. The T^2 value of the p' variables follow a χ^2 distribution, see Fig 4a. In the fault-free condition there is no deviation between the mean value $\bar{\mathbf{u}}$ and the measured value $\mathbf{u}$. This point $\mathbf{u} - \bar{\mathbf{u}} = 0$ is in the middle of a symmetric Gaussian distribution. The fault-free state of the T^2 value is also zero the point $T^2 = 0$ is at the left edge of a non-symmetric χ^2 distribution. Before using the T^2 value of the grouped variables as a new input variable in a second T^2 control chart there are some problems, which have to be solved:

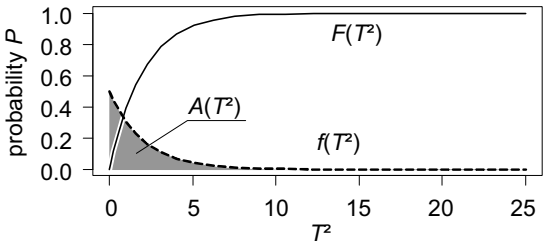
- For the transformation the fault-free condition of T^2 value must lie in the middle of a symmetric distribution. This means mean value of $f(T^2) = 0$.
- The transformed symmetric distribution of the grouped variables has to be a Gaussian distribution. If this condition is not fulfilled it is not allowed to use the grouped variable in a new T^2 control chart.
- The transformation to a Gaussian distribution should be exactly defined, because a reverse calculation should be possible if a disturbance is detected.

The presented new method can be used directly with the measured values $\mathbf{u}$ (uncorrelated variables) or with the score matrix $\mathbf{T}$ (correlated variables) after the transformation by PCA. To solve the first problem (1) is extended by a factor C which has to be calculated from the score matrix $\mathbf{T}_k$ of the p' grouped variables.

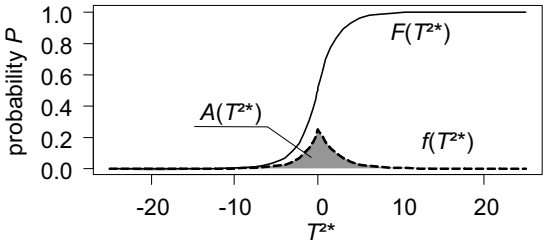
$$T_k^{2*} = \left(\mathbf{u}_k - \bar{\mathbf{u}}\right)^T \mathbf{S}^{-1} \left(\mathbf{u}_k - \bar{\mathbf{u}}\right) \cdot C_k \qquad (4a)$$

$$C_k = \prod_{ip=1}^{p} \frac{\mathbf{T}_{k;ip}}{\left|\mathbf{T}_{k;ip}\right|} \qquad (4b)$$

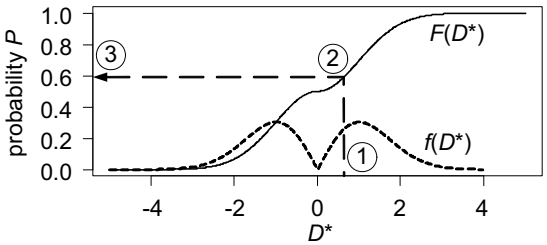
This factor in (4) is used to get a T_k^{2*} value which follows a symmetric distribution. The new symmetric density function of T_k^{2*} is defined in (5), see. Fig. 4b. The factor C gives for 50% of the T^{2*} values a positive sign and the other 50% get a negative sign. Therefore T^{2*} value is also defined in the negative region. Because the areas $A(T^2)$ and $A(T^{2*})$ under the density function in Fig. 4a and 4b have to be equal, the maximum of T^{2*} is only 50% of the maximum of T^2.



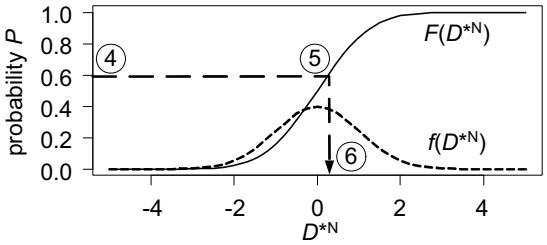
a) Probability density function f and the cumulative distribution function F of T^2



b) Symmetric distribution of T^{2*}



c) Mahalanobis distance D^*



d) Gaussian variable D^{*N}

Fig. 4. Probability functions during the transformation

$$f\left(T^{2*}\right) = \begin{cases} \chi^2_{0.5+2\alpha;p'} \cdot ½ & for\ C_k > 0 \\ \chi^2_{0;p'} \cdot ½ & for\ C_k = 0 \\ \chi^2_{2\alpha;p'} \cdot ½ & for\ C_k < 0 \end{cases} \qquad (5)$$

In the next step the Mahalanobis distance D^* is calculated by (6). The resulting functions are shown in Fig. 4c.

$$D_k^* = \begin{cases} \sqrt{T_k^{2*}} & for\ T_k^{2*} > 0 \\ 0 & for\ T_k^{2*} = 0 \\ -\sqrt{-T_k^{2*}} & for\ T_k^{2*} < 0 \end{cases} \qquad (6)$$

As D^* is not normally distributed, it is transformed by (7) to the Gaussian variable D^{*N}, as illustrated with arrows in Figs 4c and 4d. The transformation is done by using the sum func-

tions of D^* and D^{*N}. For a given D^* value the probability $P(D^*)$ is determined. By using (7) with the probability $P(D^{*N})$ the D^{*N} value is determined from a Gaussian sum function $F(D^{*N})$.

$$P\left(D_k^*\right) = P\left(D_k^{*N}\right) \tag{7}$$

The new grouped and transformed variable is used instead of the previously clustered variables. Now the T^2 value is calculated using (1) from the previously not used $p - p'$ variables and the new grouped and transformed variable that means altogether from $p^* = p - p' + 1$ variables. An overview of the whole transformation is shown in Tab. 1. Since $p^* < p$ the control limit UCL^* becomes smaller as if T^2 would have been calculated from all the variables according to (1). Therefore smaller deviations can be detected.

$$UCL^* = \chi^2_{\alpha;p^*} \quad < \quad \chi^2_{\alpha;p} = UCL \tag{8}$$

Tab. 1 Overview of the transformation steps

UCL					UCL^*	
$p = 4$		$p' = 2$			$p^* = 3$	
u_1	PC_4	T^2	T^2*	D^*	D^{*N}	T^2
u_2	PC_3					
u_3	PC_2	→			PC_2	
u_4	PC_1	→			PC_1	

2.3 Using the new method

In the above chapter the principle of the new method was explained. The next question is how the grouped variables are selected. As explained above, there are two possibilities; the first one is to use principal components and the second one is to use the measured variables. In the first case some common known procedures like Kaiser criterion can be used to define the PCs for the residual subspace. By using the presented new method the PCs of the residual subspace have to be grouped. In the second case it is possible to use physical information about the process and the measured data. For example some measurements of a machine are observed. Then it is possible to separate variables which indicate faults with high risk and some with lower risk. A variable which indicates a fault with high risk could be an acceleration measurement, because the disturbance increases very fast. A variable with lower risk could be the oil temperature, because the temperature increases slower than any mechanical unbalance. The variables with lower risk can be grouped as residual subspace in the new method. For example u_1 (or PC_4) and u_2 (or PC_3) are grouped. If a second T^2 value is calculated from the grouped and transformed variable D^{*N} and the PCs (PC_1 and PC_2) which are not used until now, no information is lost. If an abnormal state is detected, a reverse calculation can be done and the reason for the disturbance can be identified. Thereby disturbances in the non-grouped variables or principal components (T^2 → PC → variable) and also in the grouped variables or principal components (T^2 → D^{*N} → D^* → T^2* → PC → variable) can be identified.

2.4 Test of the new method

The new method was tested for both uncorrelated and correlated data. Here just one example is shown, how the fault detection sensitivity is increasing for the non-grouped variables. For the non-grouped variables (□) Fig. 5 shows the difference ΔD between the smallest detectable disturbance calculated by (1) and the new method. For the grouped variables (●) the difference between the new method and the calculation with (1) is shown.

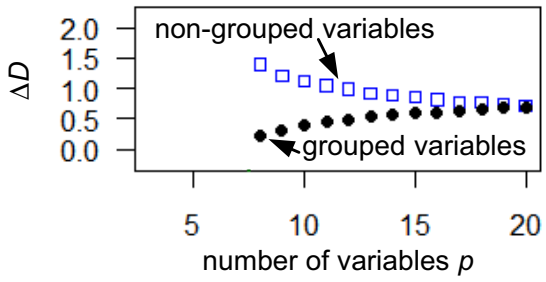


Fig. 5 Fault detection sensitivity with 7 grouped (not-correlated) variables

3. CONCLUSION

A new method was introduced for reducing the dimension of the Hotelling's statistic in order to detect smaller failures. The basic idea is to group some variables into a combined variable and to calculate the T^2 value from this variable and from the remaining variables. As the new calculated variable is not Gaussian distributed a proper transformation was applied to ensure a Gaussian distribution of the calculated variable. It was shown that those variables or PCs have to be grouped which span the residual subspace. Comprehensive simulations confirmed the presented method. These results and an application on sensor fault detection will be presented in a future work.

REFERENCES

1. T. Friebel, M. Stockmann, R. Haber, “Robust covariance matrix estimation for sensor monitoring by a two-dimensional control chart,” in 9th International Conference - Process Control, 2010, Kouty nad Desnou, Czech Republic.
2. D. C. Montgomery, “Introduction to statistical quality control,” John Wiley & Sons, 2008.
3. S. J. Quin, “Statistical process monitoring: basics and beyond,” in Journal of Chemometrics; Vol. 17, pp. 480-502, 2003.
4. R. L. Mason, N. D. Tracy, J. C. Young, “Decomposition of T² for Multivariate Control Chart Interpretation,” in Journal of Quality Technology, Vol.27, No. 2, pp. 99-108, 1995.
5. R. L. Mason, J. C. Young, “Multivariate Statistical Process Control with Industrial Application,” American Statistical Association, 2002.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

A Constrained Strategy to Control Plasma Shape in ITER

C.V.Labate[1], M. Mattei[2], D. Famularo[3], F. Koechl[4], V.Parail[5]

[1] DIMET, Università di Reggio Calabria ENEA/CREATE Association
IT (e-mail: carmelenzo.labate@unirc.it).
[2] DIAM, Seconda Università di Napoli - ENEA/CREATE Association
IT (e-mail: massimiliano.mattei@unina2.it).
[3] DEIS, Università della Calabria ENEA/CREATE Association
IT (e-mail: domenico.famularo@unical.it).
[4] Association EURATOM-ÖAW/ATI, Atominstitut
A (e-mail: Vassili.Parail@ccfe.ac.uk).
[5] EURATOM/CCFE Fusion Association, Culham Science Centre
UK (e-mail: Florian.Koechl@ccfe.ac.uk).

Abstract: When controlling plasma shape in a tokamak, the risk to drive the system out of its operating limits may become concrete. This paper presents the application of the CG (Command Governor) constrained control technique to the plasma shape control in ITER (International Thermonuclear Experimental Reactor). A primal internal loop controlling the minimum distances between the plasma and the tokamak wall chamber is firstly designed, then an external loop including the CG device modifies, if necessary, the reference signals to the primal controller, taking into account the operational constraints. The reference correction is accomplished through an on-line optimization procedure which embodies plasma model forecasts computed within a finite virtual time horizon as usual in model predictive paradigms. With respect with previous papers on the same argument this paper presents nonlinear simulation in which plasma current density profiles time histories are obtained using a detailed transport code available at JET.

Keywords: Tokamak, Nuclear Fusion, Constrained Control, Model Predictive Control, Multivariable Control.

1. INTRODUCTION

Plasma control represents a fundamental aspect in tokamak operations. It avoids plasma-wall contact, which would cause the loss of magnetic confinement and a too heavy thermal stress on mechanical structures. Desired plasma-wall clearance is obtained by regulating currents in a number of PF (Poloidal Field) coils, placed around the plasma chamber as shown in ITER (International Thermonuclear Experimental Reactor) poloidal section depicted in Fig. 1. Poloidal fields generated by such coils interact with plasma modifying its position, current and shape.

The feedback controller regulating currents in the PF coils generally has a quite simple structure based on multi-loop proportional integral derivative (PID) actions. Among MIMO control approaches, deeply investigated in the last decade to improve shape control performances, Wesson (1997). The most successful approaches are linear model based techniques Crisanti et al. (2003), Ariola et al. (1999), which however, in spite of good performance on the gap control, cannot take into account the possibility that undesired saturations can occur on some physical variables of interests. In facts during tokamak operations currents and voltage, but also induced electromagnetic fields and forces, and shape parameters must belong to prescribed ranges.

The presence of constraints is an important problem which modern control theory is trying to face with. Several approaches are presented in the literature on this topic. An attempt to deal with control inputs subject to operational limits is described in Ambrosino et al. (2001), whereas feedback control methodologies, as Anti-Windup (AW), Bumpless methods, AW/LQR, AW/H_2, Kothare at al. (1994), take into account the presence of saturations in an indirect way.

At the end of 90's, the increasing emphasis given to predictive control theory, also due to the availability of fast computing units (Diehl et al., 2005), highlighted new approaches based on invariant sets arguments and evolutions on a virtual time scale (Mayne et al.), whose peculiarity consists of an inherent capability to take directly into account, in the design phase, the presence of constraints. The control action is computed through the solution of a sequence of optimization problems with the objective to jointly maximize the control performance and enforce the satisfaction of the prescribed constraints.

Among these novel strategies, the so called *Command Governor* (CG) focuses exclusively on constraints fulfilment, leaving the control performance satisfaction (set point

tracking, disturbance rejection, robustness issues, etc.) to traditional regulation frameworks. In particular, the CG device is added as an external loop to a *primal* pre-compensated plant, characterized by stability and good tracking performance in the CG absence. At each time instant t_k, the CG computes a modified reference command which, if applied from t_k onward, does not produce constraints violations and, at the same time, represents the best approximation of the actual desired reference signal, according to an on-line constrained procedure on a receding horizon finite time interval. Many mature assessments of the CG *state of the art* for linear systems can be found in [8 - 9].

All the paper is developed with reference to the ITER tokamak which is a modern thermonuclear fusion reactor under construction in Cadarache (FR), designed as the largest nuclear fusion device ever built that should be able to achieve the ignition phase with a power ratio Q=10.

An important novelty introduced in this paper is the use of a plasma nonlinear model to assess closed loop performance with current a density profile evolution obtained by means of a detailed transport code available at JET (Joint European Torus).

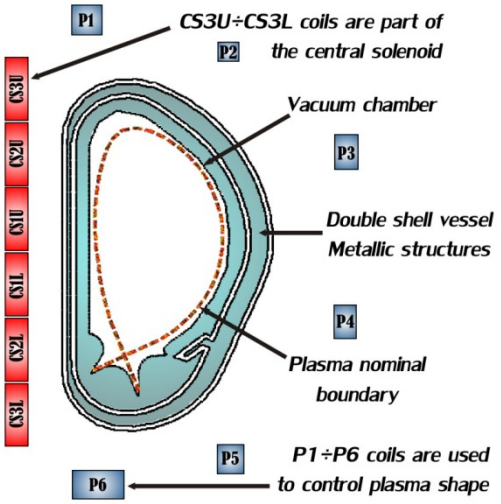


Fig.1: ITER Poloidal Section

2. PLASMA MATHEMATICAL MODELLING

The physical phenomenon to be controlled is governed by Maxwell's equations in their quasi-stationary form where the electric field can be assumed time independent and the current density is *divergence free.*

Moreover, in the time scale of interest for current, position, and shape control, because of the low plasma mass density, inertial effects can be neglected. As a consequence, in axial-symmetric geometry with cylindrical coordinates (r, ϕ, z), plasma momentum equilibrium equation becomes $J \times B = \nabla p$, rewritable in plasma region as the well known *Grad-Shafranov equation*, Wesson (1996):

$$\Delta^* \psi = r \frac{\partial}{\partial r}\left(\frac{1}{\mu_r r}\frac{\partial \psi}{\partial r}\right) + \frac{\partial}{\partial z}\left(\frac{1}{\mu_r}\frac{\partial \psi}{\partial z}\right) = -f\frac{df}{d\psi} - \mu_0 r^2 \frac{dp}{d\psi} \qquad (1)$$

$\psi(r,z)$ being the poloidal magnetic flux per radians, p the kinetic pressure profile, and f the poloidal current function profile, related to the poloidal current I_{pol} by the relation $f=\mu I_{pol}/2\pi$.

The partial differential equation problem is completed considering the interaction between plasma and surrounding passive structures and active coils. It can be written in the following form [1]:

$$\begin{cases} \Delta^* \psi = -f\dfrac{df}{d\psi} - \mu_0 r^2 \dfrac{dp}{d\psi} & \text{in plasma region} \\ \Delta^* \psi = -\mu_0 r j_{ext}(r,z,t) & \text{in conductors} \\ \Delta^* \psi = 0 & \text{elsewhere} \end{cases} \qquad (2)$$

with the initial and boundary conditions on ψ being zero at infinity and on the symmetry axis, j_{ext} being the toroidal current density in the external conductors and coils.

Solutions of Problem (2) can be numerically found by means of Finite Element Methods (FEMs) provided that the plasma boundary can be determined, the toroidal current density in the PF and the total plasma current are known, functions $p(\psi)$ and $f(\psi)$ are defined. $p(\psi)$ and $f(\psi)$ functions can be obtained running a suitable transport module, whereas the toroidal current density j_{ext} can be obtained as a linear combination of the PF circuit currents.

The time evolution of PF currents is then governed by a circuit equation driven by voltages in the active circuits. Also induced eddy currents in the metallic structures can be modelled as circuits driven by time derivatives of the plasma current and the active currents in coils.

The difficulty to use nonlinear FEM models for control design purposes, makes necessary a linearization procedure of the plasma response. We can finally approach to a linear plasma-circuit dynamics in the form:

$$\begin{cases} L\delta\dot{I} + R\delta I = \delta V + L_E \delta\dot{w} \\ \delta y = C\delta I + F\delta w \end{cases} \qquad (3)$$

where V is the vector of voltages applied to the circuits (zero for passive coils), R is the circuit resistance matrix, and L is the matrix of self and mutual inductances between the plasma, the coils, and the equivalent circuits of the passive structures, $I = [I_{PF}^T, I_E^T, I_{pl}]^T$ is the vector of PF, passive, and plasma currents respectively, and w is a vector of parameters describing plasma current density profile parameters typically assumed as external disturbances; y is a vector of outputs, C and F are output matrices, δz denotes the variation of the variable z with respect to a nominal condition.

The following sections are developed under the hypotheses that the presence of eddy currents can be neglected on the time scale of plasma shape control.

CREATE-NL model [10] is used to solve FEM plasma-circuit nonlinear equation and to produce linearized models

(3) for control design purposes. Time histories of profiles are obtained through a coupling between CREATE-NL and JETTO codes which is a recent development under validation at JET (Joint European Torus) laboratories.

3. THE CONTROL PROBLEM AND THE PRIMAL CONTROLLER

A so-called plasma controller in a tokamak has two main objectives: vertical stabilization and shape control.

The first one can is achieved by controlling the vertical speed of the plasma centroid. This control action stops vertical instabilities arising in highly elongated plasmas and is physically obtained by means of PF coil current producing radial fields variations. Since eddy currents are not considered in this paper such an action is not part of our discussions.

On the other hand, shape control can be achieved controlling a certain number of plasma-wall gaps. To this end the controller drives PF coil currents with an action which is the sum of a feedforward plus a feedback action. The feedfoward control action would be able to drive plasma through nominal expected conditions if modelling was perfect, and in absence of unexpected disturbances. The feedback action counteracts misalignments between expected and actual controlled output values.

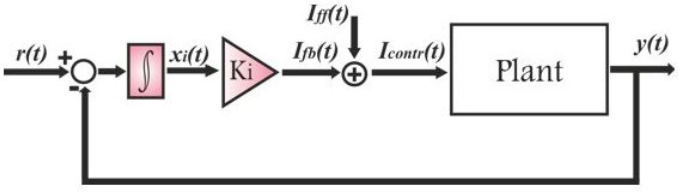


Fig. 2: Representation of the pre-compensated system, composed by the plant and the shape controller.

In our study, the feedforward action is obtained off-line by means of optimization tools based on plasma nonlinear models, whereas the feedback action is obtained by means of an integral action designed with a robust pole placement techniques forcing the closed loop time constant to be about 2-3s (Fig. 2)

Eleven voltages are independently driven to achieve the control action, corresponding to the P1÷P6, CS3U, CS3L, CS2U, CS2L, CS1 coils (CS1L and CS1U are connected in series).

4. THE COMMAND GOVERNOR FRAMEWORK

According to the scheme depicted in Fig. 3, a typical CG control scheme takes into consideration a discrete time pre-compensated linear time-invariant plant having the expression:

$$\begin{cases} x(t_{k+1}) = \Phi x(t_k) + Gg(t_k) + G_d d(t_k) \\ y(t_k) = H_y x(t_k) \\ c(t_k) = H_c x(t_k) + Lg(t_k) + L_d d(t_k) \end{cases} \quad (4)$$

where $t_k = t_0 + kT_s$, $k \in \mathbb{Z}_{0+}$, t_0 and T_s being the initial time instant and the sampling interval respectively; $x(t_k) \in \mathbb{R}^{n_x}$ is the state vector including plant and primal controller states; $g(t_k) \in \mathbb{R}^{n_r}$ is the command input vector which would coincide with the reference signal $r(t_k) \in \mathbb{R}^{n_r}$, if no constraints were present; signal $d(t_k) \in \mathbb{R}^{n_d}$ is an exogenous disturbance vector belonging to $\mathcal{D} = \{ d \in \mathbb{R}^{n_d} : Ud \le \bar{h} \}$, a closed convex and compact set, with $U \in \mathbb{R}^{n_u \times n_d}$, $n_u \ge n_d$, a full column rank matrix, and $\bar{h} = [\bar{h}_1 \quad \bar{h}_2 \ldots \bar{h}_{n_u}]^T \in \mathbb{R}^{n_u}$, a vector of nonnegative constraints ($\bar{h}_p \ge 0$, $p = 1, \ldots, n_u$); $y(t_k) \in \mathbb{R}^{n_r}$ is the output vector which is required to track $r(t_k)$; $c(t_k) \in \mathbb{R}^{n_c}$ is the vector to be constrained, viz. $c(t_k) \in \mathcal{C} \subset \mathbb{R}^{n_c}$, $\mathcal{C}$ being a closed and convex set: $\mathcal{C} = \{ c \in \mathbb{R}^{n_c} : Tc \le f \}$, with $T \in \mathbb{R}^{n_t \times n_c}$, $n_t \ge n_c$, a full column rank matrix, and $f \in \mathbb{R}^{n_t}$ a vector of constraints.

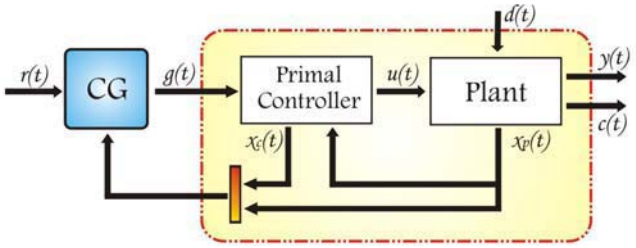


Fig. 3: Command Governor control scheme

The CG design problem consists of finding, at each time t_k, a command $g(t_k)$ as a memoryless function of the current state and the reference signal, in such a way that, under all possible disturbance sequences within $\mathcal{D}$, and compatibly with the constraints set $\mathcal{C}$ (i.e. $d(t_{k+h}) \in \mathcal{D}$, $c(t_{k+h}) \in \mathcal{C}$, $\forall h \ge 0$), $g(t_k)$ is the best approximation of $r(t_k)$ at time t_k. It is required that system (4) is asymptotically stable and offset-free (i.e. $H_y (I_{n_x} - \Phi)^{-1} = I_{n_r}$)

The contribute of external disturbances is taken into account using a *P-difference* argument. Starting from the constraints set $\mathcal{C}$ we have the recursion:

$$\mathcal{C}_0 := \mathcal{C} \sim L_d \mathcal{D}; \quad \mathcal{C}_h := \mathcal{C}_{h-1} \sim H_c \Phi^{h-1} G_d \mathcal{D}$$
$$\mathcal{C}_\infty := \bigcap_{i=0}^{\infty} \mathcal{C}_i \quad (5)$$

where the symbol $\sim$ indicates the following operation between two sets $\mathcal{A}$ and $\mathcal{B}$:

$$\mathcal{A} \sim \mathcal{B} = \{ a \in \mathbb{R}^n : a + b \in \mathcal{A}, \ \forall \ b \in \mathcal{B} \} \quad (6)$$

The set $\mathcal{C}_k$ turns out to be a suitable restriction of $\mathcal{C}$ such that, if the "disturbance free" component of $c(t_k)$, depending on the initial state and the input time sequence, belongs to $\mathcal{C}_h$, $c(t_k) \in \mathcal{C} \ \forall \ k \le h$ in the presence of disturbances.

The commands are instead considered by defining the convex and closed set $\mathcal{W}^\xi$ (assumed nonempty) which characterizes all constant inputs $\omega \in \mathbb{R}^{n_r}$ whose corresponding disturbance-free steady-state solutions of Eqs.

(4), $\bar{c}_\omega = H_c(I_{n_x} - \Phi)^{-1}G\omega + L\omega$ satisfy the constraints with a prescribed tolerance ξ. Once the role of disturbances and commands is clarified, the CG strategy consists in choosing, at each time step t_k, a constant virtual command $\omega \in \mathcal{W}^\xi$ such that the corresponding disturbance-free evolution from the measured state $x(t_k)$

$$\bar{c}(t_{k+h}, x(t_k), \omega) = H_c\left(\Phi^h x(t_k) + \sum_{i=0}^{h-1}\Phi^{h-i-1}G\omega\right) + L\omega \quad (7)$$

fulfils the restricted constraint sets $\mathcal{C}_h$, $\forall h > k$ (5), accounting for disturbance effects, and its distance from the reference $r(t_k)$ is minimal. Such a command is applied to the plant in the time interval $[t_k\ t_{k+1}[$, and the procedure is repeated at the next time t_{k+1} on the basis of the new measured state $x(t_{k+1})$. Consequently, if we denote with $\mathcal{V}(x(t_k)) \subset \mathcal{W}^\xi$ the set of all constant commands $\omega \in \mathcal{W}^\xi$, whose corresponding *c*-evolutions starting from an initial condition $x(t_k)$, at time t_k, satisfy the constraints also during transients (i.e. $\mathcal{V}(x(t_k)) := \{\omega \in \mathcal{W}^\xi : \bar{c}(t_{k+h}, x(t_k), \omega) \in \mathcal{C}_h,\ \forall\ h > 0\}$, and provided that $\mathcal{V}(x(t_k))$ is nonempty, closed and convex for all t_k, the CG command is the solution of the following constrained optimization problem:

$$\begin{aligned} &g(t_k) := \arg\min_{\omega \in \mathcal{V}(x(t_k))} J(r(t_k), \omega) \\ &\text{where} \quad J(r(t_k), \omega) := \|\omega - r(t_k)\|_\Psi^2 \end{aligned} \quad (8)$$

$\|x\|_\Psi^2 := x^T\Psi x$ being a weighted norm with Ψ a positive definite symmetric matrix.

In order to solve the optimization problem (8) in a finite time, let k^* be the *virtual horizon*, defined as the integer value such that, if $\bar{c}(t_{k+h}, x(t_k), \omega) \in \mathcal{C}_h$, $h \in \{0,1,...,k^*\}$, then $\forall h \geq 0$ $\bar{c}(t_{k+h}, x(t_k), \omega) \in \mathcal{C}_h$. k^* can be obtained by means of an off-line algorithm based on the following optimization problem:

$$\begin{aligned} &G_k(j) = \max_{x \in \mathbb{R}^{n_x}, \omega \in \mathcal{W}^\xi} T_j\,\bar{c}(t_k, x, \omega) - f_j^k \\ &\text{s.t.} \quad T_j\,\bar{c}(t_i, x, \omega) \leq f_j^i, \quad i = 0,\ldots,k-1 \end{aligned} \quad (9)$$

where T_j, $j=1,...,n_t$ denotes the j-th row of matrix T, and f_j^i $i = 0,\ldots,k-1$ formalize the P-difference operation defined in (5) and (6):

$$\begin{cases} f_j^0 = f_j - \sup_{d \in \mathcal{D}} T_j L_d d \\ f_j^1 = f_j^0 - \sup_{d \in \mathcal{D}} T_j H_c G_d d \\ f_j^{k-1} = f_j^{k-2} - \sup_{d \in \mathcal{D}} T_j H_c \Phi^{k-1-j} G_d d \end{cases} \quad (10)$$

The algorithm to derive the constraint horizon is the following:

Step 1. $k=1$;

Step 2. Find $G_k(j)$ solving Problem (9) $\forall\ \ j = 1,\ldots,n_t$

Step 3. *If* $G_k(j) \leq 0$, $\forall\ \ j = 1,\ldots,n_t$ *; then,* set $k^* = k$, and stop*; else*, $k = k+1$, go to *Step 2; end.*

According to the above algorithm, the optimization problem (9) is converted into a Quadratic Programming (QP) problem with a finite number of linear constraints to be solved on-line, that is:

$$\begin{aligned} &g(t_k) := \min_{\omega \in \mathcal{W}^\xi} J(r(t_k), \omega) \qquad s.t \\ &TH_c\Phi^h x(t_k) + T\sum_{i=0}^{h-1}\Phi^{i-h-1}G\omega + TL\omega \leq f^h, \quad h = 0,\ldots,k^* \end{aligned} \quad (11)$$

In the case that system (4) satisfies the offset-free and asymptotic stability assumptions and that $\mathcal{V}(x(t_k))$ is nonempty, the minimiser in (11) uniquely exists at each time. Moreover $\mathcal{V}(x(t_k))$ nonempty implies $\mathcal{V}(x(t_{k+h}))$ nonempty for all *h* along the trajectories generated by the CG command (*viability property*). Finally the constraints are always fulfilled and the overall closed loop system is asymptotically stable. In particular, $g(t_k)$ monotonically converges in finite time to either r or its best admissible approximation compatible with constraints.

A last remark is made on the computation of the set $\mathcal{W}^\xi$ representing a key ingredient in the CG numerical implementation. The set $\mathcal{C}_\infty$, which is mandatory to compute $\mathcal{W}^\xi$, can be numerically approximated with a convenient $\mathcal{C}_\infty^a(\varepsilon)$ such that $\mathcal{C}_\infty^a(\varepsilon) \subset \mathcal{C}_\infty \subset \mathcal{C}_\infty^a(\varepsilon) + \mathcal{B}_\varepsilon$, where $\mathcal{B}_\varepsilon$ represents a ball of radius ε (safety level) centred at the origin. Such a set is computable in a finite number of steps. Indeed, it can be shown that

$$\mathcal{C}_\infty = \mathcal{C}_k \sim \left(\sum_{i=k}^{\infty} H_c\Phi^i G_d \mathcal{D}\right) \quad (12)$$

The stability of matrix Φ and the boundedness of $\mathcal{D}$ imply the existence of two positive constants M and $\lambda \in (0,1)$, such that $\|\Phi^k\|_2 \leq M\lambda^k$ and $d_{\max} := \max_{d \in \mathcal{D}} \|d\|_2$, and this assures that, for all positive ε, there exists an index $k_\varepsilon > 0$ such that:

$$\sum_{i=k}^{\infty} H_c\Phi^i G_d \mathcal{D} \subset \mathcal{B}_\varepsilon \text{ for all } k > k_\varepsilon \quad (13)$$

Once the prescribed tolerance is fixed and M, λ and $d_{\max}$ are determined, due to the following inequality

$$d_{\max}\bar{\sigma}(H_c)\bar{\sigma}(G_d)M\sum_{i=k_\varepsilon}^{\infty}\lambda^i \leq \varepsilon \quad (14)$$

the value of k_ε can be computed as

$$k_\varepsilon = \frac{\ln(\varepsilon) + \ln(1-\lambda) - \ln(\bar{\sigma}(H_c)\bar{\sigma}(G_d)Md_{\max})}{\ln(\lambda)} \quad (15)$$

Now if we also consider the tolerance margin on the constraints fulfilment ξ, we have the following approximation of $\mathcal{C}_\infty$: $\mathcal{C}_\infty^{a\xi}(\varepsilon) = (\mathcal{C}_{k_\varepsilon} \sim \mathcal{B}_\varepsilon) \sim \mathcal{B}_\xi$ that can be used to compute $\mathcal{W}^{\xi}$ by solving the problem of determining all commands $\omega \in \mathbb{R}^{n_r}$ such that $\bar{c}_\omega \in \mathcal{C}_\infty^{a\xi}(\varepsilon)$ that brings finally to the following set:

$$\mathcal{W}^{\xi} = \left\{ \begin{array}{l} \omega \in \mathbb{R}^{n_r} : TH_c(I_{n_x} - \Phi)^{-1} G\omega + TL\omega \le \\ \qquad \le f^{k_\varepsilon} - (\xi + \varepsilon)\sqrt{T_j^T T_j} \quad j = 1,\ldots,n_t \end{array} \right\} \quad (16)$$

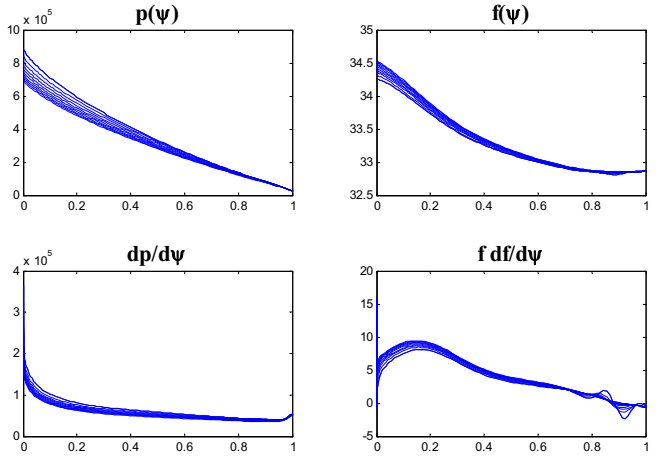


Fig 4. Example of current density profile parameter variation

5. NUMERICAL RESULTS

Numerical results are carried out with reference to a 15 MA inductive DT scenario at flat top.

Changes of the plasma shape are required to the control systems in the presence of profile variations generated by transport phenomena.

An example of the profile function evolution generated by JETTO code is shown in Figure 4.

The primal controller assumes five gaps as controlled outputs that are shown in Figure 5. This Figure also shows the change of boundary required to the plasma, whereas Figure 6 shows constraints on maximum allowable plasma displacement in terms of maximum and minimum gaps.

Modifications to the reference gaps introduced by the CG to enforce constraints are shown in Figure 7, whereas Figure 8 shows performance of the primal controller with respect to the gap references modified by the CG.

The designed CG takes into account a set of 60 constraints on the following output variables:

- poloidal field (PF) and central solenoid (CS) coil currents ($I_{P1} \div I_{P6}$, $I_{CS3L} \div I_{CS3U}$) and magnetic fields ($B_{P1} \div B_{P6}$, $B_{CS3L} \div B_{CS3U}$),
- plasma-wall gaps in correspondence of several predefined points in the plasma boundary,
- distance between first and second separatrices,
- vertical forces induced on CS coils.

Figure 9 shows the time behavior of some of the constrained variable affecting the CG re-calculation of the reference gaps.

Finally Figure 10 illustrates plasma snapshots during transients.

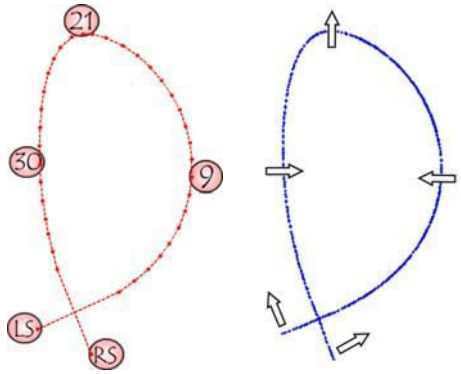


Fig. 5: Controlled gaps and required shape variation

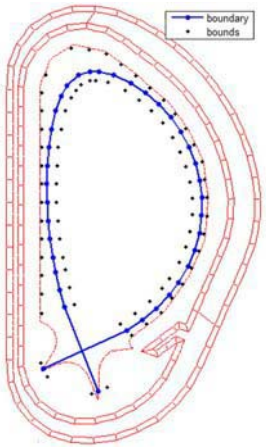


Fig 6: Constraints on maximum plasma displacements

REFERENCES

Wesson, J. (1997). *Tokamaks*, *2nd ed.* Clarendon Press, Oxford, U.K.

F. Crisanti, et al. (2003). Upgrade of the present JET Shape and Vertical Stability Controller. *Fusion Eng. Des., vol. 66–68*, pp. 803–807.

Ariola, M., Ambrosino, G., Lister, J.B., Pironti, A., Villone, F., and Vyas, P. (1999). A modern plasma controller tested on the TCV tokamak. *Fusion Technol., vol. 36, no. 2*, pp. 126–138.

Ambrosino, G., Ariola, M., Pironti, A., and Walker, M. (2001). A control scheme to deal with coil current saturation in a tokamak. *IEEE Transactions on Control Systems Technology, Vol 9, n.6*.

Kothare, M. V., Campo, P. J., Morari, M., and Nett, C. N. (1994). A Unified Framework for the Study of Anti-Windup Designs. *Automatica, Vol. 30, No. 12*, pp. 1869–1883.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., and Scokaert, P.O.M. (2009). Constrained model predictive control: Stability and optimality. *Automatica, Vol. 36*, pp. 789-814.

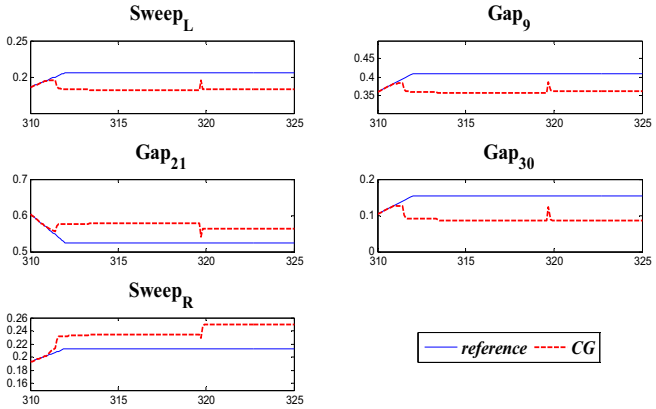


Fig 7: Original references on gaps compared with gap references modified by CG

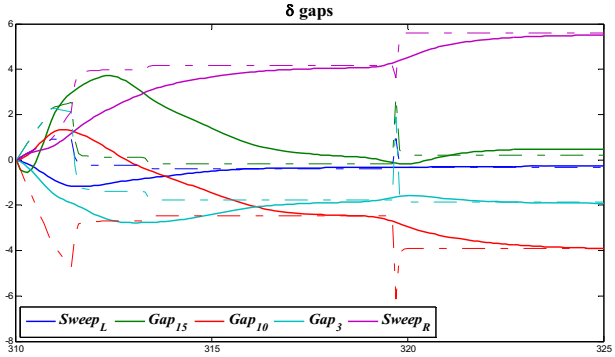


Fig 8: Controlled gaps time history compared with reference gaps produced by CG

Diehl, M., Bock, H. G., and Schloder, J. P. (2005). Real-Time Iterations for Nonlinear Optimal Feedback Control. *Proceedings of the Conference on Decision and Control, and the European Control Conference*, pp. 5871–5876, Seville, SP.

Casavola, A., Mosca, E., and Angeli, D. (2000). Robust Command Governors for Constrained Linear Systems. *IEEE Transaction on Automatic Control, Vol. 45, No. 11*, pp. 2071–2077.

Gilbert, E. G., and Kolmanovski, I. (1995). Discrete-Time Reference Governors and the Nonlinear Control of Systems with State and Control Constraints. *International Journal of Robust and Nonlinear Control, Vol. 5*, pp. 478–504.

R. Albanese and F. Villone (1998). The linearized CREATE-L plasma response model for the control of current, position and shape in tokamaks. *Nucl. Fusion, vol. 38, no. 5*, pp. 723–738.

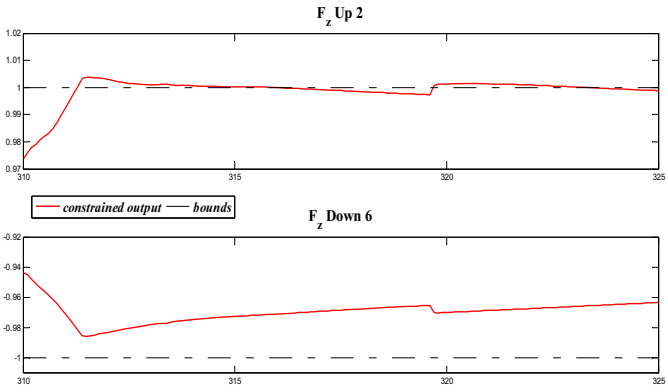


Fig 9: Example of constrained output forcing the CG to change gap references in real time

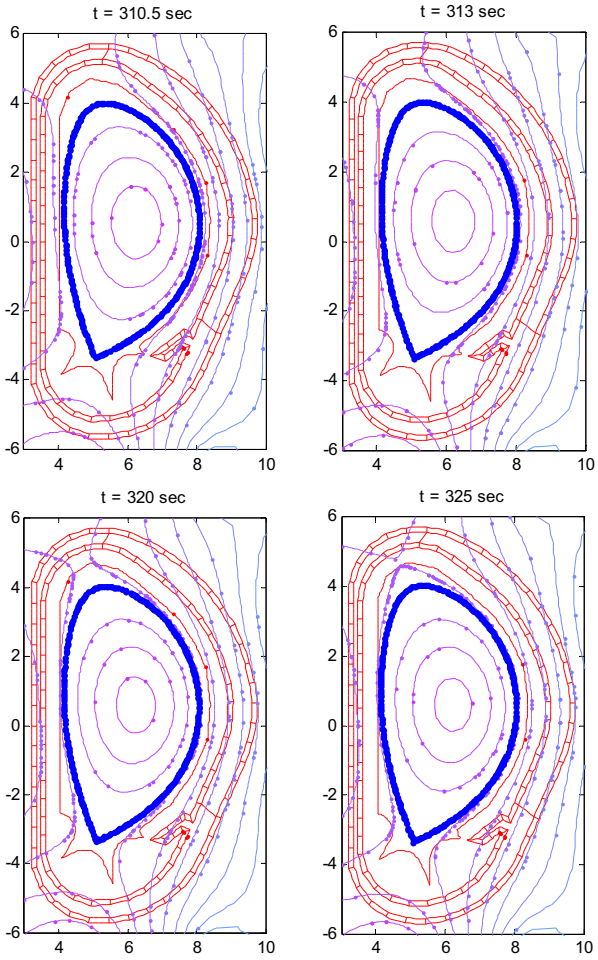


Fig. 10: Sequence of plasma shapes during transient

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fault Detection and Isolation of Wind Turbines: Application to a Real Case Study

Pep Lluís Negre * **Vicenç Puig** ** **Isaac Pineda** *

* *Alstom Wind S.L.U. - Innovation & Reliability,*
Roc Boronat, 78, 08005 Barcelona, Spain,
(e-mail: josep-lluis.negre-carrasco,isaac.pineda-amo@power.alstom.com).
** *Advanced Control Systems Group - Universitat Politcnica de Catalunya,*
Rambla Sant Nebridi, 10, 08222 Terrassa, Spain
(e-mail: vicenc.puig@upc.edu)

Abstract: The purpose of this paper is to design a Fault Detection and Isolation (FDI) system for wind turbines. With this aim, a robust fault detection based on an adaptive threshold generation is proposed. Real field data and system identification techniques are used to identify the nominal model as well as its uncertainty. The estimated output is computed from the nominal model and an observer that follows the so-called Luenberger scheme. The adaptive threshold is generated taking into account the model uncertainty. Since wind turbines are highly non-linear systems when operating in their whole range of operation, a Linear Parameter Varying (LPV) model is used. Finally, fault isolation is based on an algorithm that uses the residual fault sensitivity. Several fault scenarios are used to show the performance of the proposed approach.

Keywords: Fault detection, fault isolation, wind turbines, linear parameter varying systems, model error modelling.

1. INTRODUCTION

The future of wind energy passes through the installation of offshore wind farms. In such locations a non-planned maintenance is very costly. Therefore, a fault-tolerant control system that is able to maintain the wind turbine connected after the occurrence of certain faults can avoid major economic losses. A first step towards the implementation of a fault-tolerant system is to implement a Fault Detection and Isolation (FDI) that is able to detect, isolate, and if possible to estimate the fault (Isermann, 2006).

In this paper, a model-based FDI approach for wind turbines is proposed and applied to a commercial variable-speed, variable-pitch 3MW wind turbine of Alstom Wind S.L.U., named ECO100 (see Figure 1). This machine follows the standard of Danish concept: horizontal axis using a three bladed rotor design with an active yaw system keeping the rotor always oriented upwind. The ECO100 is II-A class IEC/EN-61400-1 with an ideal mean annual wind speed of 8.5m/s and the wind speeds cut-in and cut-off are respectively 3m/s and 25m/s (see Figure 2). The rotor velocity can vary between 7.94 - 14.3 r.p.m. and it has a swept area of 7980m^2. The tower is an hybrid 90m of height with the first 10m of concrete and the rest of steel.

Alstom Wind S.L.U. has provided a non-linear simulation model and real field data of ECO100 wind turbine that will be used in the stages of modelling and result validation. These data come from a big set of sensors installed along the wind turbine to collect time-domain measurements from the most important components.

To use any model-based technique it is necessary to obtain a model of the wind turbine. Since most of the techniques available in the literature utilize linear models, the more straightforward approach is to model the wind turbine in this way. However, a linear model will only be able to represent the non-linear wind turbine behaviour around a given operating point. To build a model that is valid along the whole operating range a Linear Parameter Varying (LPV) model will be used. Because the effectiveness of the FDI algorithm relies in its concordance with the reality, the plant model will be constructed using real faultless wind turbine data using system identification methods. Usually, the field data is presented in time series of 10 or 20 minutes. The low recording time causes that it is not possible to get the entire wind speed range (from 3m/s to 25m/s) in a single time series. Therefore, the form of the field data only allows the identification of the system in one wind speed operating

Fig. 1. Alstom ECO100 wind turbine.

point (the mean wind speed for each time series). Thus, several models have to be identified around single points along the full operation range as shown in Fig. 2.

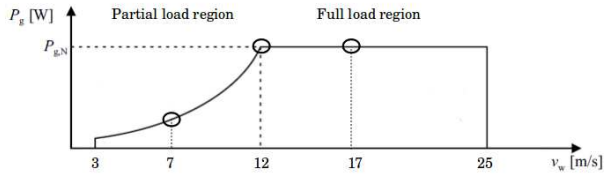


Fig. 2. Power curve range with three possible operating points separated 5m/s between them.

The innovation of this paper is to present the application of a new fault detection and isolation method for nonlinear systems that can be described as LPV models to a wind turbine. The fault detection methodology is based on comparing on-line the real system behavior of the monitored system obtained by means of sensors with the estimated behavior using an *LPV interval observer*. In the case of a significant discrepancy (residual) is detected between the LPV model and the measurements obtained by the sensors, the existence of a fault is assumed. Due to the effect of the uncertain parameters, the outputs of LPV models are bounded by an interval to avoid false alarms in the detection module. Analyzing in real-time how the faults affect to the residuals using the residual fault sensitivity, it is possible, to isolate the fault, and even in some cases it is also possible to determine its magnitude.

2. WIND TURBINE FAULTS

As in any FDI system, the set of faults to be detected and isolated should be pre-established. With this aim, the set of wind turbine faults that cause the major economic losses (in statistical terms) should be the ones that the FDI system should be designed for. One relevant information available about wind turbine fault statistics is the technical report published by UpWind in 2009 (Faulstich and Hahn, 2009) that analyses the faults in the main wind turbine components and related with the different machine typologies. This report indicates that the electrical subsystems fail more often than the mechanical ones, while mechanical subassemblies experience longer downtimes after the failure. But, it is interesting to note that, by examining this failure database, the components of the electrical and control systems fails more often than 2 years and half. In opposite, for example, a failure in the gearbox occurs only every 19 years (see Fig. 3). Similar results were obtained in the study of Ribrant and Bertling (2006) where a statistical analysis about wind turbine failures was done with data of Sweden, Finland and Germany wind companies.

Analysing these reports, it is clear that the control system is one of those responsible for the greatest number of failures in wind turbines. The sensors are one of the most important parts of the control system since the control actions are directly related to input reference sensors. Thus, it is very reasonable designing a FDI system that takes into account the faults in the sensors to prevent the control malfunctions. Besides the two control actuators governing the *Generator torque* and the *Blade pitch angle* are also susceptible to faults. The faults in actuators can be easily mitigated by fault-tolerant control techniques. Therefore, it is reasonable to include these components in the list of considered faults (Table 1).

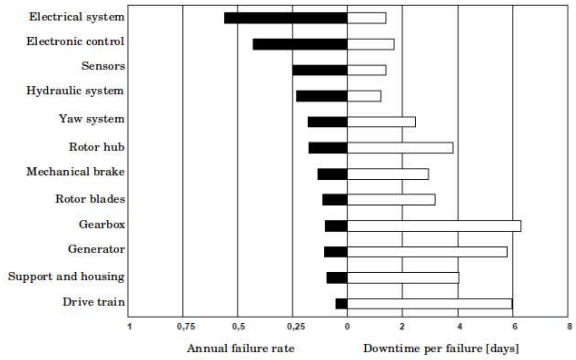


Fig. 3. Reliability statistics for main wind turbine systems.

Fault	Signal name	Signal type
f_1	Electrical power	sensor
f_2	Generator speed	sensor
f_3	Generator torque	actuator
f_4	Blade pitch angle	actuator
f_5	Wind speed	sensor

Table 1. List of considered faults

3. FAULT DETECTION USING LPV INTERVAL OBSERVERS

3.1 LPV representation

Let us consider that the nonlinear system (in our case the wind turbine) to be monitored can be described by the following LPV representation:

$$\begin{aligned} x(k+1) &= A(\tilde{\vartheta}_k)x(k) + B(\tilde{\vartheta}_k)u_0(k) + F_a(\tilde{\vartheta}_k)f_a(k) \\ y(k) &= C(\tilde{\vartheta}_k)x(k) + D(\tilde{\vartheta}_k)u_0(k) + F_y(\tilde{\vartheta}_k)f_y(k) \end{aligned} \quad (1)$$

where $u_0(t) \in \Re^{n_u}$ is the real system input, $y(t) \in \Re^{n_y}$ is the system output, $x(t) \in \Re^{n_x}$ is the state-space vector, $f_a(t) \in \Re^{n_u}$ and $f_y(t) \in \Re^{n_y}$ represents faults in the actuators and system output sensors, respectively. $\tilde{\vartheta}_k := \vartheta(k)$ is the system vector of time-varying parameters of dimension n_ϑ that change with the operating point scheduled by some measured system variables p_k $(p_k := p(k))$ that can be estimated using some known function: $\vartheta_k = f(p_k)$. However, there is still some uncertainty in the estimated values that can be bounded by:

$$\Theta_k = \{\vartheta_k \in \Re^{n_\vartheta} \mid \underline{\vartheta_k} \leq \vartheta_k \leq \overline{\vartheta_k}\}, \quad \vartheta_k = f(p_k) \quad (2)$$

This set represents the uncertainty about the exact knowledge of real system parameters $\tilde{\vartheta}_k$.

The system (1) describes a model parametrized by a scheduling variable denoted by p_k. In this paper, the kind of LPV system considered are those whose parameters vary affinely in a polytope (Apkarian et al., 1995). In particular, the state-space matrices range in a polytope of matrices defined as the convex hull of a finite number of matrices N. That is,

$$\begin{pmatrix} A(\tilde{\vartheta}_k) & B(\tilde{\vartheta}_k) & F_a(\tilde{\vartheta}_k) \\ C(\tilde{\vartheta}_k) & D(\tilde{\vartheta}_k) & F_y(\tilde{\vartheta}_k) \end{pmatrix} \in Co\left\{ \begin{pmatrix} A_j(\vartheta^j) & B_j(\vartheta^j) & F_{a,j}(\vartheta^j) \\ C_j(\vartheta^j) & D_j(\vartheta^j) & F_{y,j}(\vartheta^j) \end{pmatrix} \right\}$$

$$:= \sum_{j=1}^{N} \alpha_j(p_k) \begin{pmatrix} A_j(\vartheta^j) & B_j(\vartheta^j) & F_{a,j}(\vartheta^j) \\ C_j(\vartheta^j) & D_j(\vartheta^j) & F_{y,j}(\vartheta^j) \end{pmatrix} \quad (3)$$

with $\alpha_j(p_k) \geq 0$, $\sum_{j=1}^{N} \alpha_j(p_k) = 1$ and $\vartheta^j = f(p^j)$ is the vector of uncertain parameters of j^{th} model where each j^{th}

model is called a vertex system and it is assumed according property (2) that: $\vartheta^j \in [\underline{\vartheta^j},\ \overline{\vartheta^j}]$.

Consequently, the LPV system (1) can be expressed as follows:

$$\begin{aligned} x(k+1) &= \sum_{j=1}^{N} \alpha^j(p_k) \left[A_j(\vartheta^j)x(k) + B_j(\vartheta^j)u_0(k) + F_{a,j}(\vartheta^j)f_a(k)\right] \\ y(k) &= \sum_{j=1}^{N} \alpha^j(p_k) \left[C_j(\vartheta^j)x(k) + D_j(\vartheta^j)u_0(k) + F_{y,j}(\vartheta^j)f_y(k)\right] \end{aligned} \tag{4}$$

Here A_j, B_j, C_j and D_j are the state space matrices defined for j^{th} model. Notice that, the state space matrices of system (1) are equivalent to the interpolation between LTI models, for example: $A(\tilde{\vartheta_k}) = \sum_{j=1}^{N} \alpha^j(p_k) A_j(\vartheta^j)$.

The polytopic system is scheduled through functions $\alpha^j(p_k)$, $\forall\ j \in [1, \ldots, N]$ that lie in a convex set

$$\Psi = \Big\{ \alpha^j(p_k) \in \mathbb{R}^N, \alpha(p_k) = \left[\alpha^1(p_k), \ldots, \alpha^N(p_k)\right]^T, \\ \alpha^j(p_k) \geq 0,\ \forall j,\ \sum_{j=1}^{N} \alpha^j(p_k) = 1 \Big\}. \tag{5}$$

3.2 LPV Interval Observer

The system described by (1) is monitored using a LPV interval observer with Luenberger structure considering parameter uncertainty given by $\vartheta^j \in [\underline{\vartheta^j},\ \overline{\vartheta^j}]$. In the following, we consider only strictly proper systems such that $D = 0$. Consequently, the LPV interval observer can be written by extending the representation of Meseguer et al. (2006) for LTI models as:

$$\begin{aligned} \hat{x}(k+1) &= \sum_{j=1}^{N} \alpha^j(p_k) \left[A_{0,j}(\vartheta^j)\hat{x}(k) + B_j(\vartheta^j)u(k) + L_j y(k)\right] \\ \hat{y}(k) &= \sum_{j=1}^{N} \alpha^j(p_k) \left[C_j(\vartheta^j)\hat{x}(k)\right] \end{aligned} \tag{6}$$

where $A_{0,j}(\vartheta^j) = A_j(\vartheta^j) - L_j C_j(\vartheta^j)$, $u(k)$ is the measured system input vector, $\hat{x}(k)$ is the estimated system state vector, $\hat{y}(k)$ is the estimated system output vector and L_j is the observer gain that has to be designed in order to stabilize the observer given by (6) for all $\vartheta^j \in [\underline{\vartheta}^j, \overline{\vartheta}^j]$. Each observer gain matrix $L_j \in \Re^{n_x \times n_y}$ is designed to stabilize each vertex j^{th} and to guarantee a desired performance ($A_{0,j}$) regarding fault detection for $\vartheta^j \in [\underline{\vartheta^j},\ \overline{\vartheta^j}]$ (Chilali and Gahinet, 1996).

3.3 Observer input/output form

The system in (1) can be expressed in input-output form using the shift operator q^{-1} and assuming zero initial conditions as follows:

$$y(k) = y_0(k) + G_{f_a}(q^{-1}, \tilde{\vartheta_k}) f_a(k) + G_{f_y}(q^{-1}, \tilde{\vartheta_k}) f_y(k) \tag{7}$$

where:

$$y_0(k) = G_u(q^{-1}, \tilde{\vartheta_k}) u_0(k) \tag{8}$$

$$G_u(q^{-1}, \tilde{\vartheta_k}) = C(\tilde{\vartheta_k})(qI - A(\tilde{\vartheta_k}))^{-1} B(\tilde{\vartheta_k}) + D(\tilde{\vartheta_k}) \tag{9}$$

$$G_{f_a}(q^{-1}, \tilde{\vartheta_k}) = C(\tilde{\vartheta_k})(qI - A(\tilde{\vartheta_k}))^{-1} F_a(\tilde{\vartheta_k}) \tag{10}$$

$$G_{f_y}(q^{-1}, \tilde{\vartheta_k}) = F_y(\tilde{\vartheta_k}) \tag{11}$$

Alternatively, the observer described by Eq. (6) can be expressed in input-output form by [1]:

$$\hat{y}(k) = \sum_{j=1}^{N} \alpha^j(p_k) \left[G^j(q^{-1}, \vartheta^j)u(k) + H^j(q^{-1}, \vartheta^j)y(k)\right] \tag{12}$$

where:

$$G^j(q^{-1}, \vartheta^j) = C_j(\vartheta^j)(qI - A_{0,j}(\vartheta^j))^{-1} B_j(\vartheta^j) \tag{13}$$

$$H^j(q^{-1}, \vartheta^j) = C_j(\vartheta^j)(qI - A_{0,j}(\vartheta^j))^{-1} L_j \tag{14}$$

The effect of the uncertain parameters ϑ_k on the observer temporal response $\hat{y}(k, \vartheta_k)$ can be bounded using an interval satisfying [2]:

$$\hat{y}(k) \in \left[\underline{\hat{y}}(k), \overline{\hat{y}}(k)\right] \tag{15}$$

in a non-faulty case. Such interval is computed independently for each output (neglecting couplings between outputs):

$$\underline{\hat{y}}(k) = \min_{\vartheta_k \in \Theta} \left\{ \sum_{j=1}^{N} \alpha^j(p_k) \left[G^j(q^{-1}, \vartheta^j)u(k) + H^j(q^{-1}, \vartheta^j)y(k)\right] \right\}$$

$$\overline{\hat{y}}(k) = \max_{\vartheta_k \in \Theta} \left\{ \sum_{j=1}^{N} \alpha^j(p_k) \left[G^j(q^{-1}, \vartheta^j)u(k) + H^j(q^{-1}, \vartheta^j)y(k)\right] \right\}$$

subject to the observer equations given by (6). Such interval can be computed using the algorithm based on numerical optimization presented in Puig et al. (2005).

3.4 Adaptive thresholding

Fault detection is based on generating a nominal residual comparing the measurements of physical variables $y(k)$ of the process with their estimation $\hat{y}(k)$ provided by the associated system model:

$$r(k) = y(k) - \hat{y}(k) \tag{16}$$

where $r(k) \in \Re^{n_y}$ is the residual set and $\hat{y}(k)$ is the prediction obtained using the nominal LPV model. According to Gertler (1998), the computational form of the residual generator, obtained using (12), is:

$$r(k) = \sum_{j=1}^{N} \alpha^j(p_k) \left[-G^j(q^{-1}, \vartheta^j)u(k) + \left(I - H^j(q^{-1}, \vartheta^j)\right) y(k)\right] \tag{17}$$

Alternatively, the residual given by (17) can be also expressed in terms of the effects caused by faults using its internal or unknown-input-effect form (Gertler, 1998). This form, obtained combining (7), (12) and (16), is expressed as:

$$\begin{aligned} r(k) = r_0(k) + \sum_{j=1}^{N} \alpha^j(p_k) \Big[&\left(I - H^j(q^{-1}, \vartheta^j)\right) \left(G^j_{f_y}(q^{-1}, \vartheta^j) f_y(k)\right. \\ &\left. + G^j_{f_a}(q^{-1}, \vartheta^j) f_a(k)\right)\Big] \end{aligned} \tag{18}$$

where

$$\sum_{j=1}^{N} \alpha^j(p_k) G^j_{f_y}(q^{-1}, \vartheta^j) = G_{fy}(q^{-1}, \tilde{\vartheta_k})$$

$$\sum_{j=1}^{N} \alpha^j(p_k) G^j_{f_a}(q^{-1}, \vartheta^j) = G_{fa}(q^{-1}, \tilde{\vartheta_k})$$

$$r_0(k) = \sum_{j=1}^{N} \alpha^j(p_k) \left[-G^j(q^{-1}, \vartheta^j)u(k) + \left(I - H^j(q^{-1}, \vartheta^j)\right) y_0(k)\right] \tag{19}$$

Notice that, the expression (19) represents the non-faulty residual. Comparing (17) and (19), it should be noticed that both

[1] In the following, for simplicity and with abuse of notation, transfer functions are used for LPV systems, although computations are performed entirely using the state space representation: $G^{(j)} \triangleq \left[\begin{array}{c|c} A_0^{(j)}(\vartheta^j) & B^{(j)}(\vartheta^j) \\ \hline C^{(j)}(\vartheta^j) & 0 \end{array}\right]$

[2] In the remainder of the paper, interval bounds for vector variables should be considered component wise.

$r_0(k)$ and $r(k)$ are affected in the same way by the observation gain L.

When considering model uncertainty, the residual generated by (16) will not be zero, even in a non-faulty scenario. To cope with the parameter uncertainty effect a passive robust approach based on adaptive thresholding can be used (Horak, 1988). Thus, using this passive approach, the effect of parameter uncertainty in the residual $r(k)$ (associated to each system output $y(k)$) is bounded by the interval:

$$r(k) \in [\underline{r}(k), \overline{r}(k)] \tag{20}$$

where:

$$\underline{r}(k) = \underline{\hat{y}}(k) - \hat{y}(k) \text{ and } \overline{r}(k) = \overline{\hat{y}}(k) - \hat{y}(k) \tag{21}$$

being $\hat{y}(k)$ the nominal predicted output, $\underline{\hat{y}}(k)$ and $\overline{\hat{y}}(k)$ the bounds of the predicted output (15) using observer (6). The residual generated by (21) can be expressed in input-output form using (12) as:

$$\underline{r}(k) = \min_{\theta \in \Theta} \left\{ \sum_{j=1}^{N} \alpha^j(p_k) \left[\Delta G^j(q^{-1}, \vartheta^j) u(k) + \Delta H^j(q^{-1}, \vartheta^j) y(k) \right] \right\} \tag{22}$$

$$\overline{r}(k) = \max_{\theta \in \Theta} \left\{ \sum_{j=1}^{N} \alpha^j(p_k) \left[\Delta G^j(q^{-1}, \vartheta^j) u(k) + \Delta H^j(q^{-1}, \vartheta^j) y(k) \right] \right\} \tag{23}$$

where:

$$\begin{aligned} \Delta G^j(q^{-1}, \vartheta^j) &= G^j(q^{-1}, \vartheta^j) - G^j(q^{-1}, \vartheta_0^j) \\ \Delta H^j(q^{-1}, \vartheta^j) &= H^j(q^{-1}, \vartheta^j) - H^j(q^{-1}, \vartheta_0^j) \end{aligned}$$

being ϑ_0^j the nominal parameters.

Then, a fault is indicated if the residuals do not satisfy the relation given by (20), or alternatively, if the measurement is not inside the interval of predicted outputs given by (16)-(16).

Fig. 4 summarizes the robust fault detection scheme proposed. The main signals that appear in the picture are the following: the controller actions u, the measured outputs y, the estimated outputs $\hat{y}$, he residual r, the observer correction c and the fault f. y examining the *residual generation* block, one can see that, in addition to the nominal model, includes an observer scheme. The nominal model is used to estimate the outputs that more closely fits the current wind turbine outputs. The observer is placed in order to avoid drifting between the estimated and measured outputs that would cause erroneous fault detection. On the other hand, *residual evaluation* part is responsible for generating the threshold taking into account the model uncertainty. These limits take into account the uncertainty in the modelling stage (by using a model of the error) and make the FDI system robust.

4. FAULT ISOLATION USING LPV FAULT SENSITIVITIES

4.1 Fault signature matrix

Fault isolation consists in identifying the faults affecting the system. It is carried out on the basis of fault signatures, (generated by the detection module) and its relation with all the considered faults, $f(k)$=$\{f_a(k), f_y(k)\}$. Robust residual evaluation presented in Section 3.4 allows obtaining a set of *fault signatures* $\phi(k) = [\phi_1(k), \phi_2(k), \ldots, \phi_{ny}(k)]$, where each fault indicator is given by:

$$\phi_i(k) = \begin{cases} 0 \text{ if } r(k) \notin [\underline{r}(k), \overline{r}(k)] \\ 1 \text{ if } r(k) \in [\underline{r}(k), \overline{r}(k)] \end{cases} \tag{24}$$

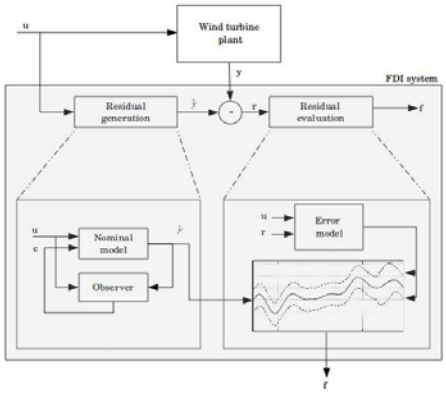


Fig. 4. Complete model-based FDI scheme designed in this research.

The standard FDI fault isolation method is based on exploiting the relation defined on the Cartesian product of the sets of considered faults:

$$FSM \subset \phi \times f, \tag{25}$$

where *FSM* is the theoretical fault signature matrix (Gertler, 1998). One element of such matrix $FSM_{i\ell}$ will be equal to one, if the fault $f_\ell(k)$ is affected by the residual $r_i(k)$. In this case, the value of the fault indicator $\phi_i(k)$ must be equal to one when the fault appears in the monitored system. Otherwise, the element $FSM_{i\ell}$ will be zero.

In this work, it is proposed to use of information provided by the *fault residual sensitivity* in the design of the diagnosis system in order to increase fault isolability.

4.2 LPV Fault residual sensitivity

In general, the occurrence of a fault signal can be caused by different faults. Therefore what allows distinguishing one fault from the others are the fault signal dynamic properties that should be different for each different fault. According to (Gertler, 1998), these theoretical dynamic properties are described by the *residual fault sensitivity* that can be expressed as follows:

$$S_f = \frac{\partial r}{\partial f} \tag{26}$$

which is a transfer function that describes the effect on the residual, r, of a given fault f. The expression of residual sensitivity is obtained using the residual internal form given by (18). Thus, the sensitivity changes with the operating point parametrized by scheduling variable p_k as the LPV system (1).

The residual (18) can be re-written as follows:

$$r(k) = r_0(k) + S_{f_y}(q^{-1}, \tilde{\vartheta}_k) f_y(k) + S_{f_a}(q^{-1}, \tilde{\vartheta}_k) f_a(k) \tag{27}$$

where S_{f_y} is the sensitivity of the output sensor fault and S_{f_a} is the sensitivity of the actuator fault.

LPV Residual sensitivity of an output sensor fault

Analyzing the residual internal form given by (27), and considering the fault residual sensitivity definition given by (26), the residual sensitivity for the case of a output sensor fault f_y is given by a matrix S_{f_y} of dimension $n_y \times n_y$ whose expression is:

$$S_{f_y}(q^{-1},\tilde{\vartheta}_k)=\sum_{j=1}^{N}\alpha^j(p_k)\left[\left(I-H^j(q^{-1},\vartheta^j)\right)G^j_{fy}(q^{-1},\vartheta^j)\right]$$
$$=\begin{bmatrix} S_{fy_{1,1}}(q^{-1},\tilde{\vartheta}_k) & \cdots & S_{fy_{1,ny}}(q^{-1},\tilde{\vartheta}_k)\\ \vdots & \ddots & \vdots \\ S_{fy_{ny,1}}(q^{-1},\tilde{\vartheta}_k) & \cdots & S_{fy_{ny,ny}}(q^{-1},\tilde{\vartheta}_k)\end{bmatrix} \quad (28)$$

where the element of this matrix located at the i^{th}-row and in the ℓ^{th}-column. $S_{fy_{i,\ell}}$ describes the sensitivity of the residual $r_i(k)$ regarding the fault $f_{y_\ell}(k)$ affecting the output sensor.

LPV Residual sensitivity of an actuator fault

Applying the analysis procedure used in the output sensor case, the residual sensitivity of an actuator fault f_a is given by a matrix S_{fa} of dimension $n_y \times n_u$:

$$S_{f_a}(q^{-1},\tilde{\vartheta}_k)=\sum_{j=1}^{N}\alpha^j(p_k)\left[\left(I-H^j(q^{-1},\vartheta^j)\right)G^j_{fa}(q^{-1},\vartheta^j)\right]$$
$$=\begin{bmatrix} S_{fa_{1,1}}(q^{-1},\tilde{\vartheta}_k) & \cdots & S_{fa_{1,nu}}(q^{-1},\tilde{\vartheta}_k)\\ \vdots & \ddots & \vdots \\ S_{fa_{ny,1}}(q^{-1},\tilde{\vartheta}_k) & \cdots & S_{fa_{ny,nu}}(q^{-1},\tilde{\vartheta}_k)\end{bmatrix} \quad (29)$$

where each row of this matrix is related to one component of the residual vector $r(k) = \{r_i(k) : i = 1, 2, \ldots, n_y\}$ while each column is related to one component of the actuator fault vector $f_a = \{f_{a,\ell} : \ell = 1, 2, \ldots, n_u\}$.

4.3 Fault isolation algorithm

Figure 5 shows the scheme of the fault diagnosis algorithm proposed in this paper. The detection module has been alredy explained in Section 3. The result of this module applied to the residual $r(k)$ produces an *observed fault signature* $\phi(k)$. The observed fault signature is then supplied to the fault isolation module that will try to isolate the fault so that a fault diagnosis can be produced.

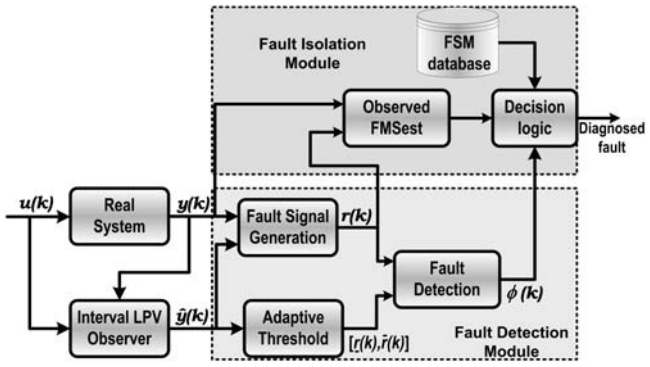


Fig. 5. Block diagram of the fault diagnosis system.

In this paper, a new fault isolation approach is proposed that makes use of the fault estimation provided the residual fault sensitivity (26). More precisely, assuming that $\left(S_f(q^{-1},\tilde{\vartheta}_k)\right)^{-1}$ exists [3], the expression of the fault estimation is given by:

$$\hat{f}_{\varrho,\ell}(k)=\left(S_{f_{\varrho,\ell}}(q^{-1},\tilde{\vartheta}_k)\right)^{-1}r_i(k) \quad (30)$$

[3] If $\left(S_f(q^{-1},\tilde{\vartheta}_k)\right)^{-1}$ is non-square and can be tackled using the left pseudo-inverse

where $i \in [1, \ldots, n_y]$ and being $\hat{f}_{\varrho,\ell} = \left\{\hat{f}_{y,\ell}, \hat{f}_{a,\ell}\right\}$, $\forall\ \ell \in [1, \ldots, n_y, 1, \ldots, n_u]$. This relation considers the influence of each fault $f(k)$ on the each residual $r(k)$. Notice that, the sensitivity expression changes with the operating point and consequently the fault estimation is parametrized by scheduling variable p_k.

Using the fault estimation (30), a new *FSM* matrix (called *fault signature matrix FSMest*) can be defined as shown in Table 2. This fault signature matrix is evaluated at every time instant.

$\hat{f}_{\varrho,\ell}$	$f_{y,1}$	$\cdots$	f_{y,n_y}	$f_{a,1}$	$\cdots$	f_{a,n_u}
$r_1(k)$	$\hat{f}_{r_1 f_{y,1}}$	$\cdots$	$\hat{f}_{r_1 f_{y,n_y}}$	$\hat{f}_{r_1 f_{a,1}}$	$\cdots$	$\hat{f}_{r_1 f_{a,n_u}}$
$r_2(k)$	$\hat{f}_{r_2 f_{y,1}}$	$\cdots$	$\hat{f}_{r_2 f_{y,n_y}}$	$\hat{f}_{r_2 f_{a,1}}$	$\cdots$	$\hat{f}_{r_2 f_{a,n_u}}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$r_{ny}(k)$	$\hat{f}_{r_{ny} f_{y,1}}$	$\cdots$	$\hat{f}_{r_{ny} f_{y,ny}}$	$\hat{f}_{r_{ny} f_{a,1}}$	$\cdots$	$\hat{f}_{r_{ny} f_{a,nu}}$

Table 2. Fault signature matrix based on the fault estimation (*FSMest*) with respect to $r_i(k)$

Each fault hypothesis corresponds to each ℓ^{th}-column of *FSMest* matrix of Table 2. The fault hypothesis corresponding to ℓ^{th}-column is accepted if all the fault estimation values are equal. More precisely, assuming that the system is just affected by one fault $f(k)$ at a time t_0, the isolation process is done by finding the fault that presents a fault estimation with a minimum distance with respect to the average of fault estimation hypothesis being postulated as a diagnosed fault:

$$\min\left\{d_{f_{y,1}},\ldots,d_{y,n_y},d_{f_{a,1}},\ldots,d_{a,n_u}\right\} \quad (31)$$

where the distance is calculated using the Euclidean distance between vectors:

$$d_{f_{\varrho,\ell}}=\sqrt{\left(\hat{f}_{r_1 f_{\varrho,\ell}}(k)-\hat{f}^m_{f_{\varrho,\ell}}(k)\right)^2+\cdots+\left(\hat{f}_{r_{ny} f_{\varrho,\ell}}(k)-\hat{f}^m_{f_{\varrho,\ell}}(k)\right)^2} \quad (32)$$

where:

$$\hat{f}^m_{f_{\varrho,\ell}}(k)=\frac{\sum_{i=1}^{n_y}\hat{f}_{r_i f_{\varrho,\ell}}(k)}{n_y}$$

for $f_{\varrho,\ell} = \{f_{y,\ell}, f_{a,\ell}\}$, $\forall\ \ell \in [1, \ldots, n_y, 1, \ldots, n_u]$.

Finally, in order to prevent false alarms when the fault signals appears in different time instants, Puig et al. (2005) proposes a solution that consists in not allowing an isolation decision until a prefixed waiting time (T_w) has elapsed from the first fault signal appearance. This time T_w can be calculated from the largest transient time response from non-faulty situation to any faulty situation. The value T_w must be evaluated once the first residual is activated. This interval of time is maximum when the fault is the minimum isolable fault. Such a fault will be determined in the next section.

5. APPLICATION TO A REAL WIND TURBINE

5.1 Wind turbine model for FDI

For FDI purposes the measured variables used for the control of the wind turbine were considered. Fig. 6 illustrates the basic control scheme of the wind turbine Alstom ECO100. The inputs of the plant are the wind speed υ, which can be divided in mean wind speed $\overline{\upsilon}$ and the turbulent part $\widetilde{\upsilon}$, the generator torque reference τ_{ref} and pitch angle reference β_{ref}. On the other

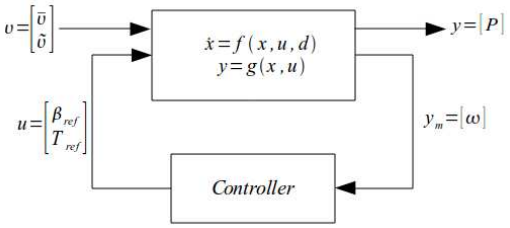


Fig. 6. Control scheme for the ECO100 wind turbine.

hand, the outputs of the plant are the angular generator speed ω and the electrical power P.

The GH Bladed model [4] of the ECO100 wind turbine was delivered by Alstom. This is an encrypted model that does not allow to use the model equations explicitly. However, this model can be used to extract information about the structure of the model to be used when identifying a LPV model for fault detection. GH Bladed allows users to linearize its internal non-linear model at any wind turbine operating point.

The linear model obtained with GH Bladed around a given operating point contains all the dynamics and can be a higher order model (upper than 40 states). Therefore, it is useless to be used for designing a FDI system. These techniques usually require relatively low order models but high accuracy. With this aim, a set of lower models relating each output signal with the considered input signals are obtained using the Hankel model reduction technique (see Figure 7). The order of each model is presented in Table 3.

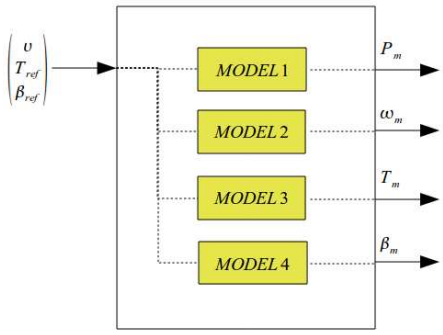


Fig. 7. Bank of models for FDI

Model	Output	Order
1	Electrical power	4
2	Generator speed	4
3	Generator torque	2
4	Blade pitch angle	3

Table 3. Order of the FDI models

Once the structure of the models have been determined, the parameters of the nominal models has been estimated around each considered operating using real data coming from a real ECO100 wind turbine and the MATLAB identification toolbox. Figure 8 shows the result of model prediction for the electrical power output after parameters have been calibrated at different operating points. Once the parameters around each operating point have obtained, the scheduling functions $\tilde{\vartheta}_k = f(p_k)$ for the LPV parameters are approximated by polynomials whose coefficients are estimated following the procedure described in Bamieh and Giarre (2002) using data taken at different operating points $p_k \in [\underline{p}, \overline{p}]$.

[4] GH Bladed is a program for wind turbine modelling highly validated which provides very accurate non-linear simulation models

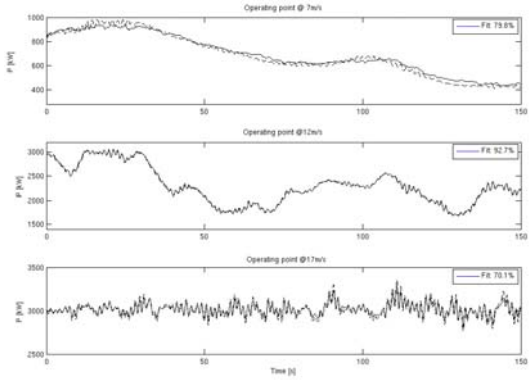


Fig. 8. Electrical power model prediction

5.2 Model error modelling

The uncertainty model around each operating point is obtained by Model Error Modelling (MEM) techniques proposed by Reinelt et al. (2001). The basic idea of MEM is to use the nominal model identified in the previous section (denoted G_0), and a collection of measured field data (y, u) to identify an error model as follows:

(1) Compute the residual $\epsilon = y - G_0 u$.
(2) Consider the "error system", with input u and output ϵ, and identify a model G_e for this system. This is an estimation of the error due to undermodeling, the so-called MEM.

Identification of the MEM from residual data can be seen as a separation between noise and unmodeled dynamics. In fact, G_e is an estimation of the dynamic system ΔG, such that:

$$\epsilon(k) = \Delta G u(k) + e(k) \tag{33}$$

If not knowledge about the structure of this model exists and in the absence of specific suspected non-linearities, it is reasonable to test non-linear neural networks black boxes (Ljung (1999)) to identify the error model. It means that the Eq. (33) can be rewritten as:

$$\epsilon(k) = \widetilde{f}(u(k)) + e(k) \tag{34}$$

where $\widetilde{f}$ is a non-linear function that can be modeled f.e. using a neural network NNFIR model.

Figure 9 shows the prediction provided by nominal and error models once have been calibrated using real data for the electrical power output (Model 1 in Table 3).

5.3 FDI system design

The methods presented in the previous sections allow the FDI system to detect faults but not to isolate them. There are some faults that can cause the activation of more than one residual as for example in the case of the wind sensor. Since the wind

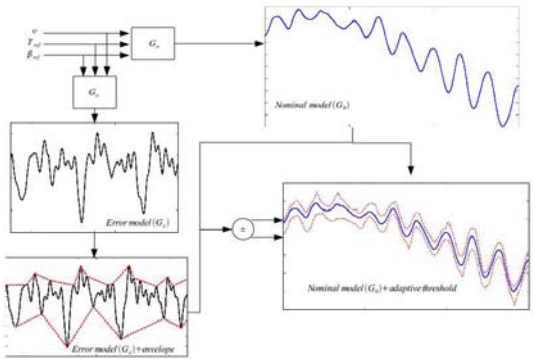


Fig. 9. Prediction provided by nominal and error models

speed is an input of all models in the FDI scheme, a fault in this sensor can cause an activation of all residual signals. There is no way to determine if a fault in such signal will induce a massive residual activation because it depends on the type and magnitude of the fault. The results described in Mesenguer et al. (2010) enables FDI system to identify the source of fault by analysing the residual sensitivities. The Table 4 indicates to the FDI system which is the transfer function that determines the time evolution should have the residual signal for each fault. Then, for example, if there is a fault in the wind speed sensor, the residual of model 1 (with electrical power output) must have the same shape than the indicated by the sensitivity $S_{fu}(q)$.

Fault	Signal name	Signal type	Sensitivity
f_1	Electrical power	output sensor	$S_{fy}(q)$
f_2	Generator speed	output sensor	$S_{fy}(q)$
f_3	Generator torque	actuator	$S_{fa}(q)$
f_4	Blade pitch angle	actuator	$S_{fa}(q)$
f_5	Wind speed	input sensor	$S_{fu}(q)$

Table 4. Sensitivity analysis of each fault according its type.

5.4 Results

Let us consider, for example, an abrupt fault scenario in wind speed input sensor. In such case, the fault was only detected by the residual corresponding to generator speed output sensor (see Figure 10). Without the use of the residual sensitivity analysis, the FDI system would have assigned the fault to generator speed sensor and the reconfiguration action taken by the fault-tolerant control would be wrong. Note that two faults could have affected to the generator speed residual: fault f_2 (generator speed sensor fault) and f_5 (wind speed sensor fault).

Therefore, the corresponding two residual sensitivity functions have to be analysed and shown in Figure 11. In this figure, the time evolution of the residual sensitivity for model 2 with generator speed output is illustrated. The left graph is the fault sensitivity to an output sensor fault, i.e. a fault in the generator speed sensor. The right graph is the fault sensitivity to an input sensor fault, i.e. a fault in the wind speed sensor. Both curves are different, thus the faults are isolable using residual fault sensitivities.

Let us consider now the two possible fault scenarios described in the list above with an abrupt fault (fixed value to 0) in both cases. Figure 10 illustrates the FDI internal signals for the corresponding model when these faults occur. The figure shows the generator speed output sensor fault (abrupt fault)

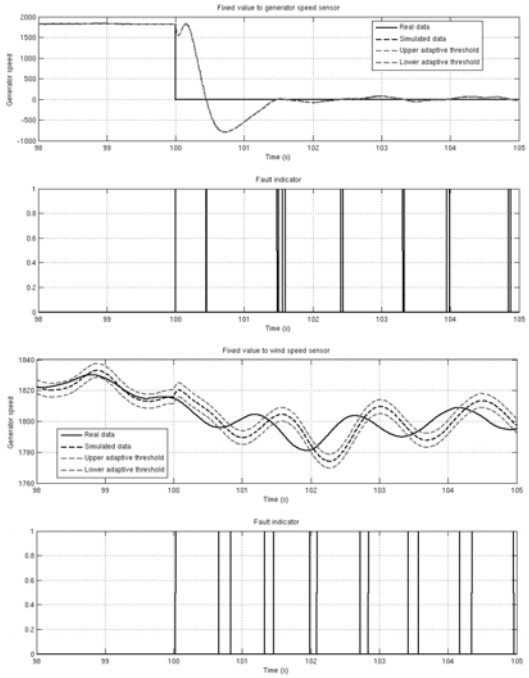


Fig. 10. Fault detection results using model 2 in case of fault f_2 (two upper plots) and f_5 (two lower plots), respectively.

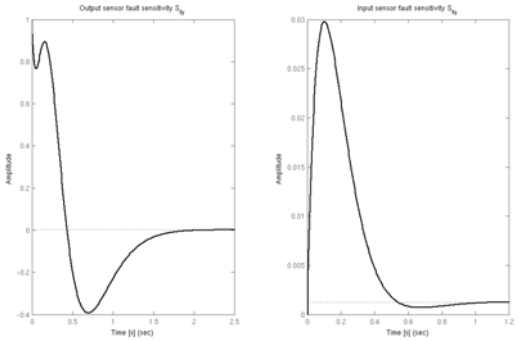


Fig. 11. Time evolution of the residual sensitivity for generator speed.

in the top plot, and the wind speed input sensor fault (abrupt fault) in the bottom plot. It is obvious that both scenarios cause different behaviours in the estimated output: when an abrupt fault is done in generator speed output sensor, the estimated output is not able to fit the real faulty measurement. Thus the fault indicator is always active after the fault. When an abrupt fault is given in wind speed input sensor, not only the input of the model is affected because the LPV model uses this signal for the parameter variation. Therefore, after the transient is difficult to determine what happens with the estimated output. However, the signal time evolution in the transient zone can be analysed in order to find out the fault origin. By examining Figure 11 and Figure 10 it is easy to see that the faults are clearly isolable since the estimated output follows the shape of the corresponding residual sensitivity in the transient. For the case of wind speed input sensor fault, the transient has less amplitude than in the generator speed output sensor fault. This effect is also visible in Figure 11 where the residual sensitivities have different amplitude scale being much lower the one corresponding to wind speed input sensor.

6. CONCLUSIONS

In this paper, a FDI system has been developed for wind turbines. The study is based on a ECO100 model (a variable-speed, variable pitch 3MW real wind turbine) using real data provided by the company Alstom Wind S.L.U. The wind turbine faults considered are chosen based on published statistical studies of wind turbine faults and are related to elements used for the control system that include sensor and actuator faults. Models for FDI have been constructed using system identification methods using real wind turbine field data to achieve the maximum matching with the reality. Additionally, the uncertainty is taken into account to be robust against modelling errors and signal noise. This goal is achieved by using techniques of model error modelling that allow finding a model for the uncertainty. The fault detection has been addressed through LPV interval observers. The fault isolation task has been implemented using the concept of residual fault sensitivities. This concept has been exploited to provide additional information to the relationship between residuals and faults. Moreover, it allows obtaining the possible fault estimation for each residual signal. Additionally the minimum detectable and isolable fault has been presented. This information is important to evaluate the limits of fault diagnosis method. For faults smaller than minimum detectable/isolatable fault, this methodology can not detect or isolate the fault, respectively. Satisfactory results have been obtained in several fault scenarios in the considered wind turbine.

REFERENCES

Apkarian, P., Gahinet, P., and Becker, G. (1995). Self-scheduled H_∞ control of linear parameter-varying systems: a design example. *Automatica*, 31(9):1251 – 1261.

Bamieh, B. and Giarre, L. (2002). Identification of Linear Parameter Varying Models. *International Journal Robust Nonlinear Control*, 2(12):841–853.

Chilali, M. and Gahinet, P. (1996). H_∞ design with pole placement constraints: an LMI approach. *IEEE Transactions on Automatic Control*, 41(3):358–367.

Faulstich, S. and Hahn, B. (2009). Comparison of different wind turbine concepts due to their effects on reliability. *Project Upwind*.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, New York.

Horak, D. T. (1988). Failure detection in dynamic systems with modelling errors. *Journal of Guidance, Control, and Dynamics*, 11(6):508–516.

Isermann, R. (2006). *Fault Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer, New York.

Ljung, L. (1999). Model validation and model error modeling. *Linkopings universitet*.

Meseguer, J., Puig, V., and Escobet, T. (2006). Observer gain effect in linear interval observer-based fault detection. *Sixth IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, 6.

Mesenguer, J., Puig, V., Escoert, T., and Saludes, J. (2010). Observer gain effect in linear interval observer-based fault detection. *Process Control*.

Puig, V., Schmid, F., Quevedo, J., and Pulido, B. (2005). A new fault diagnosis algorithm that improves the integration of fault detection and isolation. *44th IEEE Conference on Decision and Control, and the European Control Conference 2005. Seville, Spain*, pages 3809–3814.

Reinelt, W., Garulli, A., and Ljung, L. (2001). Model error modelling in robust identification. *Linkopings universitet*.

Ribrant, J. and Bertling, L. (2006). Reliability performance and maintenance - a survey of failures in wind power systems. *KTH School of Electrical Engineering*.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Second-order sliding modes and soft computing techniques for fault detection

Milan Rapaić, Zoran Jeličić *
Alessandro Pisano and Elio Usai **

* *Computing and Control Dept., Faculty of Technical Sciences, Univ. of Novi Sad, Serbia (e-mail: {rapaja,jelicic}@uns.ac.rs)*
** *Dept. of Electrical and Electronic Engineering (DIEE), Univ. of Cagliari, Cagliari, Italy (e-mail: {pisano,eusai}@diee.unica.it).*

Abstract: This paper outlines some results concerning the combined application of second-order sliding-mode and soft-computing techniques in the framework of fault-detection problems. A method for estimating the discrete state of an LTI affine switched system is developed to that end. Simple controller/observer tuning formulas are constructively developed along the paper by Lyapunov analysis. Simulation and experimental results confirm the expected performance.

Keywords: Distributed parameter systems. Sliding mode control. Uncertain Systems.

1. INTRODUCTION

In the framework of FDI, faults in dynamical systems are usually modeled by uncertain exogenous signals entering the system dynamics (i.e., unknown inputs). Within this area, powerful results have been achieved in the context of the so called model based FDI by using several types of Unknown-Input Observers (UIOs) (Simani et al. (2002)). Here we follow a different direction, by representing the faults by means of abrupt changes in the system dynamics. This choice naturally leads to consider a **switched** dynamics as the mathematical model of the system under investigation (see e.g. (Wang et al. (2007))).

More precisely in this paper we deal with a problem of fault detection for affine linear switched dynamics. The system's current mode of operation (location) is not known, and is wanted to be reconstructed. The state vector is assumed to be available for measurements. We assume that some of the locations correspond to faulty modes of operation and a nonlinear observer stack is constructed which allows to identify the occurrence of the faulty behaviour, and to insulate it. The structure of the suggested scheme is a stack of second-order sliding-mode observers, each one producing a scalar residual signal. From an appropriate processing of the delivered residuals (residual evaluation) the actual mode of operation can be identified. We compare two methods for the residual evaluation: i.) standard thresholding, and ii.) soft-computing technique by means of Artificial Neural Networks (ANNs) and Support Vector Machine (SVMs).

Although simple residual tresholding may be quite effective from a purely theoretical point of view, in an industrial application the plant under consideration is often nonlinear and sometimes affected by a number of exogenous signals. Thus nonzero residual values from *all* of the observers will commonly be present. The measurements will be also usually strongly corrupted by noise, and several residual signals might become approximately equal at the same time, since due to measurement noise it is often impossible to accurately detect the smallest one. All of the above problems may be overcome by utilization of a suitable robust classifier. In recent years several classification techniques based on different soft computing methodologies emerged. These include artificial neural networks (ANNs) and support verctor machines (SVMs) (Kecman (2001)),(Scholkopf and Smola (2002)).

In the present work, making reference to affine switched dynamics with uncertain discrete state, we present a sliding mode based approach to discrete state recostruction that makes use of soft computing techniques at the residual evalulation stage. Section II describes a method for discrete state identification in affine switched systems by means of a stack of sliding mode obserber along with a standard threshold-based residual evaluation technique. The method in tested in Section 3 by numerical simulations. The alternative methods for residual evaluation, using neural networks, are illustrated in the Section 4. Section 5 presents an experimental application of the suggested methods to a laboratory size hydraulic process where certain faults in a centrifugal pump are detected. It follows from the obtained results that soft computing-based residual evaluation technique outperform the simple thresholding method. Section 6 presents some concluding remark.

2. DISCRETE-MODE IDENTIFICATION FOR SWITCHED DYNAMICS

Consider the linear affine switched system

$$\dot{x}(t) = A_{j(t)}x(t) + B_{j(t)}u(t) + F_{j(t)}; \quad j(t) \in \{1, 2, ..., q\} \quad (1)$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and where the so-called "commutation signal" (or "discrete state") $j(t)$ determines the actual system dynamics among the possible q

⋆ The authors gratefully acknowledge the financial support from the FP7 European Research Projects "PRODI - Power plants Robustification by fault Diagnosis and Isolation techniques", grant no.224233.

"operation modes" represented by the triplets (A_i, B_i, F_i), $i = 1, 2, ..., q$.

Consider the next expression for the piecewise constant commutation signal

$$j(t) = j_k, \quad t_{k-1} \leq t < t_k, \quad k = 1, 2, ..., \infty \tag{2}$$

where $t_0 = 0$ and t_k are the "switching times" at which the discrete state is changing.

Let the next **dwell-time** restriction holds for the switching sequence

$$t_k - t_{k-1} \geq \Delta, \quad k = 1, 2, ..., \infty \tag{3}$$

The dwell time restrictions inhibits the occurrence of the so-called "Zeno phenomenon" for the considered switched dynamics, namely the occurrence of infinitely fast changes in the system modes of operation.

Some specific operation modes in the set $\{1, 2, ..., q\}$ are supposed to correspond to **faulty conditions** for system (1) that need to be detected for real-time monitoring and fault diagnosis purposes. The pair $(x(t), u(t))$ is supposed to be accessible for measurements. The task is to reconstruct the unknown discrete state $j(t)$. The logic that drives the mode switchings can be either driven by internal system's variables or driven by an external supervisor, in any case it is unknown to the designer. Then, the identification of the correct mode after the switching times will require a transient. This transient should be faster that the Δ value involved in dwell time restriction (3), otherwise the estimation will be use-less.

A parallel stage containing q observers, one for each of the possible modes of operation, is suggested:

$$\dot{\hat{x}}_i(t) = A_i x(t) + B_i u(t) + F_i + v_i(t), \quad i = 1, 2, ..., q \tag{4}$$

where $v_i(t)$ is the injection input for the $i-th$ observer, to be designed.

Denote the observation error for the i-th observer as

$$e_i = \hat{x}_i - x_i \tag{5}$$

Then the switched and fractional order error dynamics will be given by

$$\begin{aligned} \dot{e}_i(t) = (A_i - A_{j(t)})x(t) + (B_i - B_{j(t)})u(t) + \\ +(F_i - F_{j(t)}) + v_i(t) \end{aligned} \tag{6}$$

It can be then separated the error dynamics of the "**correct**" observer (i.e., that having the index i which matches the current mode of operation $j(t)$):

$$\dot{e}_i(t) = v_i(t), \quad i = j(t) \tag{7}$$

and the error dynamics of the remaining "**wrong**" observers:

$$\begin{aligned} \dot{e}_i(t) = (A_i - A_{j(t)})x(t) + (B_i - B_{j(t)})u(t) + \\ +(F_i - F_{j(t)}) + v_i(t), \quad i \neq j(t) \end{aligned} \tag{8}$$

Denote

$$\Delta A_i^j = A_i - A_{j(t)} \tag{9}$$

$$\Delta B_i^j = B_i - B_{j(t)} \tag{10}$$

$$\Delta F_i^j = F_i - F_{j(t)} \tag{11}$$

and

$$\varphi_i^j(x, u, t) = \Delta A_i^j x(t) + \Delta B_i^j u(t) + \Delta F_i^j \tag{12}$$

then (8) is rewritten as

$$\dot{e}_i(t) = \varphi_i^j(x, u, t) + v_i(t), \quad i \neq j(t) \tag{13}$$

Concerning the state- and input-dependent functions $\varphi_i^j(x, u, t)$ entering the dynamics (13) of the wrong observers, in order to guarantee the identifiability of the correct mode they should not be identically zero. Then it is made the next

Assumption A_1

$$\left\| \varphi_i^j(x, u, t) \right\| \neq 0, \quad \forall i, j = 1, 2, ..., q, \;\; i \neq j \tag{14}$$

The above assumption A_1 should be understood as a constraint on the dynamics of the switched system and in particular on the resulting $(x-u)$ time evolutions. In other words it could be said that the manifolds $\varphi_i^j(x, u, t) = 0$ should not contain admissible $x(t) - u(t)$ trajectories of the switched system.

Functions $\varphi_i^j(x, u, t)$ are also supposed to be smooth enough according to the next

Assumption A_2 There is a constant Φ such that

$$\left\| \frac{d}{dt} \varphi_i^j(x, u, t) \right\| \leq \Phi, \quad \forall i, j = 1, 2, ..., q \tag{15}$$

Clearly, the above Assumption A_2 implies a bounded, although arbitrarily large, admissible domain for the evolution of the (x, u) trajectories in the respective space. This gives **semi-global** validity to the presented discrete mode observer.

The design of the observer injection terms is carried out as follows

$$\sigma_i = \hat{x}_i - x \tag{16}$$

$$v_i = v_{1i} + v_{2i} \tag{17}$$

$$v_{1i} = -k_1 \sigma_i - k_2 |\sigma_i|^{1/2} \text{sign}(\sigma_i) \tag{18}$$

$$\dot{v}_{2i} = -k_3 \text{sign}(\sigma_i) \tag{19}$$

It is worth to note that the dynamics of σ_i is:

$$\dot{\sigma}_i = \begin{cases} v_i(t) & i = j(t) \\ \varphi_i^j(x, u, t) + v_i(t) & i \neq j(t) \end{cases} \tag{20}$$

It can be defined a unique set of tuning rules for the gains of the q observers. Consider the next inequalities involving the tuning coefficients:

$$k_1 > 2\Phi \quad k_2 > 0 \quad k_3 > \Phi \sqrt{k_1} \tag{21}$$

The main idea behind the proposed observer structure is that, after a finite transient starting at any switching times, the injection input $v_{2i}(t)$ will be identically zero for the *correct* observer and will be separated from zero for the *wrong* observers. This can be obtained, by virtue of Assumption A_1, if the **finite-time convergence to zero of** σ_i **and** $\dot{\sigma}_i$ is provided for all the q observers $(i = 1, 2, ..., q)$.

Let the maximal finite transient duration be denoted as T. Then, in that case, provided that $T < \Delta$, the next relationship directly derives by the achieved conditions $\sigma_1 = \dot{\sigma}_1 = \sigma_2 = \dot{\sigma}_2 = ... = \dot{\sigma}_q = 0$:

$$v_{2i}(t) = \begin{cases} 0 & i = j(t) \\ -\varphi_i^j(x,u,t) & i \neq j(t) \end{cases} \quad t_{k-1} + T \leq t \leq t_k \quad (22)$$

On the basis of (20), and by taking into account the Assumption A_1 as well, it can be developed a simple method for estimating the actual discrete state $j(t)$ by comparing the norms of the the observer signals $v_{21}(t)$, $v_{22}(t)$, ..., $v_{2q}(t)$ looking for the closest to zero:

$$\hat{j}(t) = \arg\min_i \|v_{2i}(t)\| \quad (23)$$

The proposed scheme for the identification of the discrete state in the switched system (1) is summarized in the next:

Theorem 1 Consider system (1), fulfilling the Assumptions A_1 and A_2, and the observer stack (4), (16)-(19) with the observer gains chosen according to (21). Then, there is $T > 0$ such that the discrete state estimation (23) will satisfy the next relation

$$\hat{j}(t) = j(t), \qquad t_{k-1} + T \leq t \leq t_k, \quad k = 1, 2, ... \quad (24)$$

Proof of Theorem 1 By combining (7) and (8), the dynamics of the error variables e_i is given by:

$$\dot{e}_i(t) = \begin{cases} v_i(t) & i = j(t) \\ \varphi_i^j(x,u,t) + v_i(t) & i \neq j(t) \end{cases} \quad (25)$$

By (2), during the first switching interval $t \in (0, t_1)$ the actual mode is $j(t) = j_1$. Then (20) specializes as

$$\dot{\sigma}_{j_1} = v_{j_1}(t) \quad (26)$$

$$\dot{\sigma}_i(t) = \varphi_i^j(x,u,t) + v_i(t), \quad i = 1, 2, ..., q, \quad i \neq j_1 \quad (27)$$

Considering (17)-(19) into (26) yields

$$\dot{\sigma}_{j_1}(t) = -k_1\sigma_{j_1} - k_2|\sigma_{j_1}|^{1/2}\text{sign}(\sigma_{j_1}) + v_{2,j_1}(t) \quad (28)$$

$$\dot{v}_{2,j_1} = -k_3\text{sign}(\sigma_{j_1}) \quad (29)$$

$$\dot{\sigma}_i(t) = -k_1\sigma_i - k_2|\sigma_i|^{1/2}\text{sign}(\sigma_i) + v_{2,i}(t) + \varphi_i^j(x,u,t) \quad (30)$$

$$\dot{v}_{2,i} = -k_3\text{sign}(\sigma_i) \quad i = 1, 2, ..., q, \quad i \neq j_1 \quad (31)$$

By introducing the new coordinates

$$z_i(t) = v_{2,i}(t) + \varphi_i^j(x,u,t) \quad (32)$$

one can augment and rewrite (30)-(31) as

$$\dot{\sigma}_i = -k_1\sigma_i - k_2|\sigma_i|^{1/2}\text{sign}(\sigma_i) + z_i(t) \quad (33)$$

$$\dot{z}_i = -k_3\text{sign}(\sigma_i) + \frac{d}{dt}\varphi_i^j(x,u,t), \quad i = 1, 2, ..., q, \quad i \neq j_1 \quad (34)$$

To prove the finite time convergence to zero of σ_i and z_i (and, hence, of $\dot{\sigma}_i$) the same Lyapunov function as that used in (Moreno et al. (2009)) is considered:

$$V_i = 2k_3|\sigma_i| + \frac{1}{2}z_i^2 + \frac{1}{2}\left(k_1|\sigma_i|^{1/2}\text{sign}(\sigma_i) + k_1\sigma_i - z_i\right)^2 \quad (35)$$

which can be rewritten as follows

$$V_i = \xi_i^T H \xi_i \quad (36)$$

$$\xi_i = \begin{bmatrix} |\sigma_i|^{1/2}\text{sign}(\sigma_i) \\ \sigma_i \\ z_i \end{bmatrix} \quad H = \begin{bmatrix} (4k_3 + k_2^2) & k_1k_2 & -k_2 \\ k_1k_2 & k1^2 & -k_1 \\ -k_2 & -k_1 & 2 \end{bmatrix} \quad (37)$$

By evaluating the derivative of (36)-(37) along the trajectories of system (33)-(34), and considering the tuning rules (21), it can be found two positive constants γ_1 and γ_2 such that

$$\dot{V}_i \leq -\gamma_1 V_i - \gamma_2\sqrt{V_i} \quad (38)$$

which easily implies, by simple application of the comparison Lemma, that V_i tends to zero in a finite time.

Let $T > 0$ be the finite transient time. By increasing the Φ constant in the tuning formulas, the transient time can be made as small as desired (Levant (2003),Polyakov et al. (2009)), and in particular such that $T << \Delta$.

The next conditions are thus achieved:

$$v_{2j_1}(t) = 0 \qquad T < t < t_1 \quad (39)$$

$$v_{2i}(t) = -\varphi_i^j(x,u,t), \quad i = 1, 2, ..., q, \quad i \neq j_1 \quad (40)$$

In light of the assumption A_1, the residual-based estimation logic (23) provides the reconstruction of the discrete state after the transient time T, i.e.

$$\hat{j}(t) = j_1, \qquad 0 + T \leq t \leq t_1 \quad (41)$$

At the time moment $t = t_1$ the discrete state will be changing. A new transient of length T is activated for the observer error dynamics, at the end of which the next conditions will be in force:

$$v_{2j_2}(t) = 0 \qquad t_1 + T < t < t_2 \quad (42)$$

$$v_{2i}(t) = -\varphi_i^j(x,u,t), \quad i = 1, 2, ..., q, \quad i \neq j_2 \quad (43)$$

Thus the estimation logic (23) still provides the reconstruction of the discrete state after the transient time, i.e.

$$\hat{j}(t) = j_2, \qquad t_1 + T < t < t_2 \quad (44)$$

By iteration on the successive switching intervals, condition (24), and so Theorem 1, is proven. □

Remark 1. The logic (23) appears not completely effective, since in some case it can happen that Assumption A_1 is violated and, as a result, also the "wrong" residuals $v_{2i}(t)$ $(i \neq j(t))$ can occasionally cross the zero value.

On the other hand, only the correct residual $(i = j(t))$ can stay at (or, more realistically, close to) zero for long time intervals. Hence the next averaged residuals can be considered

$$R_i(t) = \int_{t-\delta}^{t} \|v_{2i}(\tau)\| d\tau. \quad (45)$$

where δ is a small time delay (the width of a receding horizon window of observation for the residuals $\|v_{2i}\|$), along with the corresponding modified discrete state evaluation strategy

$$\hat{j}(t) = \arg\min_i R_i(t) \quad (46)$$

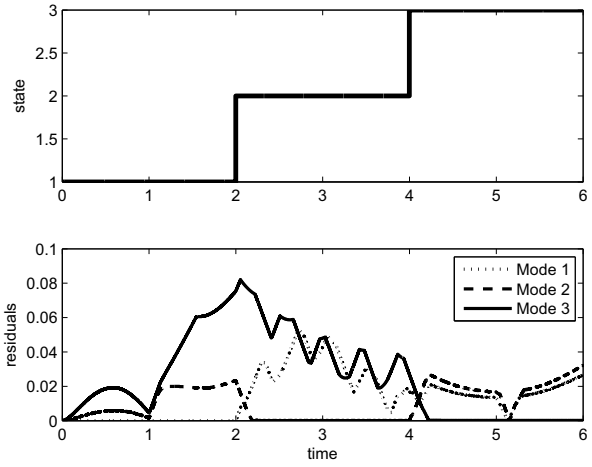


Fig. 1. Discrete state estimation test. The residuals $\|v_{21}\|$, $\|v_{22}\|$, $\|v_{23}\|$.

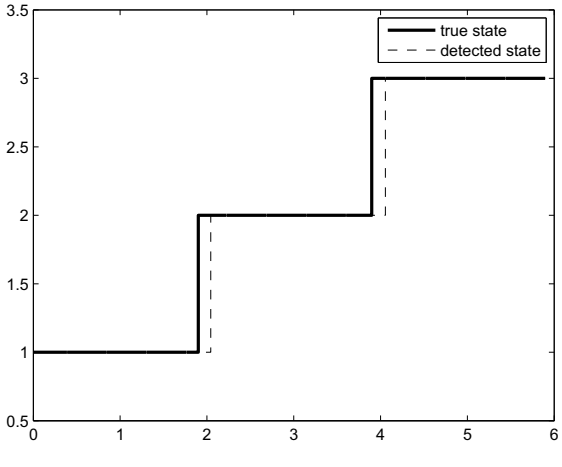


Fig. 2. Discrete state estimation test. Actual and estimated discrete state.

3. SIMULATION RESULTS

Now let us consider the discrete state estimation problem for the affine switched system (1) with $q = 3$ distinct submodels defined by the matrix triplets

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad F_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{47}$$

$$A_2 = \begin{bmatrix} 0 & 1 \\ -2 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{48}$$

$$A_3 = \begin{bmatrix} -6 & -4 \\ 1.5 & -1 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad F_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{49}$$

The discrete state is changed according to the next rule

$$j(t) = \begin{cases} 1 & 0 \leq t < 2 \\ 2 & 2 < t \leq 4 \\ 3 & 4 < t <\leq 6 \end{cases} \tag{50}$$

The parallel stage of observers (4), (16)-(19) have been implemented with the gains $k_1 = 0.01$, $k_2 = 0.01$ $k_3 = 0.1$. The discrete state and the residual signals $\|v_{21}\|$, $\|v_{22}\|$, $\|v_{23}\|$ corresponding to the different observers are presented in Figure 1. It can be noted that the "correct" residuals tend to zero in the corresponding time intervals, while the "wrong" residual keep always separated from zero except some isolated time instant (the residual of mode 1 becomes zero around $t = 5.5$, however promptly leaving the zero value). To cope with this fact, the modified residual evaluation strategy (45)-(46) has been implemented, with the length of the time window chosen as $\delta = 0.1s$. The actual and discrete state are depicted in the Figure 2 which confirms the satisfactory performance of the discrete mode observer.

4. RESIDUAL EVALUATION VIA SOFT COMPUTING TECHNIQUES

Many different topologies of neural networks have been investigated in literature. Single layer feed-forward networks (FFN) are utilized in the present work. The structure of such network is presented in Fig. 3. If the network inputs are denoted as x_i ($i \in \{1, \ldots, ni\}$), network output as y, and the outputs of the nodes within the hidden layer as o_j ($j \in \{1, \ldots, nh\}$) then

$$y = f_o(\sum_{j=0}^{nh} w_j o_j) + y_0, \tag{51}$$

$$o_j = f_h(\sum_{i=0}^{ni} \theta_{ji} x_i) + o_{j0}, \tag{52}$$

where f_o and f_h are the activation functions of the output and hidden layer, respectivelly. In principle, these can be arbitrary mappings. In the current work, both are selected to be equal to the *logarithmic sigmoid function* (commonly abbreviated as *logsig*)

$$logsig(x) = \frac{1}{1 + exp(-x)}. \tag{53}$$

Real parameters w_j and θ_{ji} are the so called *weights*, while y_0 and o_{j0} are *biases*. The values of these parameters are set in the training phase of the neural network implementation. During training, the size (number of nodes in the hidden layer) of the neural network is fixed, and the parameter values are set by minimization of classification error by means of a suitable nonlinear optimization technique.

SVMs represent a more recent development in the field of soft computing. The overal structure of a SVM is identical to the one presented in Fig 3, however the network output is computed differently, as

$$y = sign\left(\sum_{j=1}^{M} y_j \alpha_j k(\mathbf{x}_j, \mathbf{x}) + b\right), \tag{54}$$

with $\mathbf{x}_j$ denoting the j-th input pattern of the training set and y_j is its class label (0 or 1). $\mathbf{x}$ denotes the input vector being evaluated (classified). b and α_j are bias and weights, all obtained during the training phase. $k(\cdot)$ is the so called *kernel-function* which should be determined prior to SVM training. Common kernel functions are *polynomial kernel* $k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y})^d$ (d is some pre-specified positive integer) and *rbf kernel* $k(\mathbf{x}, \mathbf{y}) = \exp(-||\mathbf{x} - \mathbf{y}||^2/c)$ (c is some

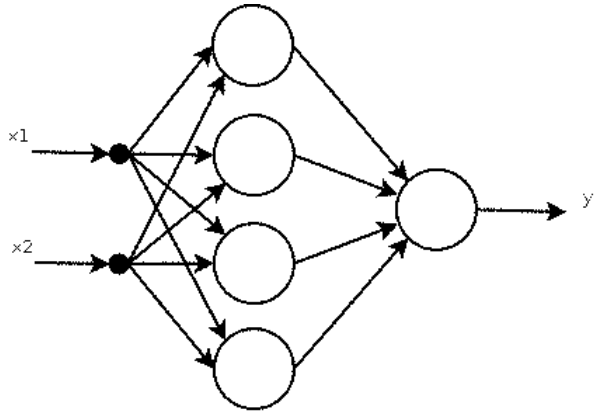


Fig. 3. Feedforward network with two inputs, one output and three nodes in the hidden layer.

pre-specified positive real number). During optimization, only a handful of α_k parameters becomes different from zero. The input patterns $\mathbf{x}_k$ corresponding to such α_k are denoted as **support vectors**.

The primary difference between ANNs and SVMs is that with SVMs the number of nodes in the hidden layer (*support vectors*) is not fixed beforehand. It is set during training. For further details we refer to Scholkopf and Smola (2002).

In the current work, the classifiers are trained off-line using a portion of available process data. Residuals from the observes designed previously are used as inputs. The ANN (or SVM) is expected to be able to distinguish nominal working regime from the faulty one with high accuracy on the entire data set.

5. EXPERIMENTAL FAULT DETECTION OF A HYDRAULIC PLANT

The discrete state estimation algorithm previously described will now be exploited to detect certain faults in a laboratory hydraulic system. The experimental hydraulic setup is shown in Figure 4. The centrifugal pump (**P**) draws the water from the lower tank (**TL**) into the upper tank (**TU**). The flow is adjusted by the electrical servo valve (**FV**). The flow is measured using the flow meter and the pulse flow transmitter (**FT**). The level in the upper tank is measured by the float level sensor and transmitter (**LT**).

The considered fault is a unintentional and undesirable reduction of the pump rotating speed, which reduces the amount of flow. The fault has been reproduced in the experimental setup by reducing the pump control signal from the nominal value (healthy condition) to about 70% of the nominal value (faulty condition). Three first order affine models have been derived via least square identification; two for the healthy and one for the faulty behaviour. Then, the suggested method for discrete state reconstruction can be applied as a fault diagnosis logic, by identifying whether the actual mode of operation is the faulty one, or is it one corresponding to the healthy working regime.

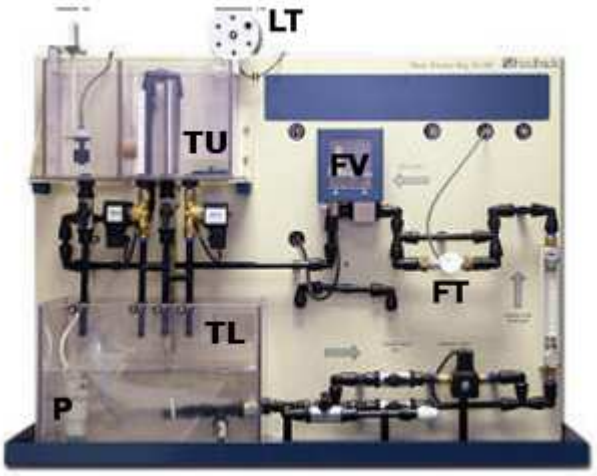


Fig. 4. "Feedback" Level/Flow Process Control System, PROCON 38-001

Two different data sets were obtained. In the first data set (FDS) a pump fault occurs at $t = 90s$ and is active during almost the entire recording. In the second data set (NDS) the pump fault does not occur at all. Before any processing the measurements were scaled into the $[0, 1]$ interval, with 0 and 1 denoting the minimal and maximal measured value of the involved physical quantity (either level or flow, or pump control signal). Let us denote the normalized flow measurements by $f(t)$, the normalized level measurement by $l(t)$ and the normalized pump control signal by $v(t)$. In the actual setup, the above are current signals in the standard range of 4-20 mA. For the purpose of model identification, the sampling time was selected as $T_s = 10ms$.

Two scalar ($n = m = 1$) affine models have been identified for the **nominal** working regime. The first model (NOMINAL-HIGH) roughly corresponds to the high value of the flow, while the second (NOMINAL-LOW) one roughly corresponds to the medium and low values of the flow. It has been established that a single model (FAULTY) is sufficient for describing the faulty working regime. Thus we have overall $q = 3$ models, two of which correspond to a healthy operating regime while the third one corresponds to a faulty condition.

The **NOMINAL-HIGH** model is

$$\dot{f}(t) = -64.43f(t) - 1.86l(t) + 14.82v(t) + 50.34, \quad (55)$$

the **NOMINAL-LOW** model is

$$\dot{f}(t) = -0.76f(t) + 0.15l(t) + 0.68v(t) - 0.03, \quad (56)$$

and the **FAULTY** model is

$$\dot{f}(t) = -1.38f(t) + 0.20l(t) + 0.76v(t) - 0.02. \quad (57)$$

The parallel stage of observers (4), (16)-(19) has been implemented with the gains $k_1 = 135$, $k_2 = 100$ $k_3 = 100$. As before, the residual signal were chosen as the integrals of $\|v_{21}\|$, $\|v_{22}\|$ and $\|v_{23}\|$ in a receding horizon time window of width $\delta = 1s$, according to the modified residual evaluation strategy (45)-(46).

Result of the simple thresholding classification procedure for the faulty data set is presented in Figure 5. The

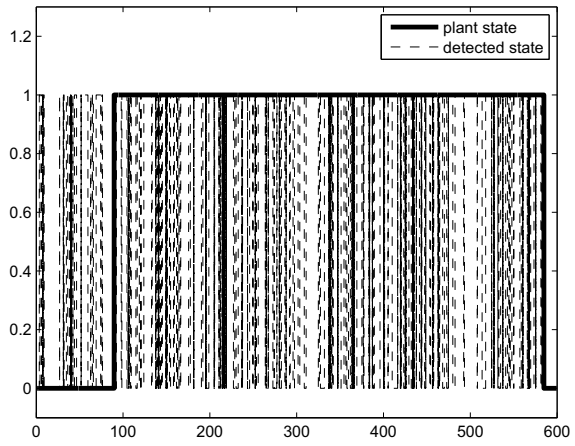


Fig. 5. Fault detection using simple tresholding techniques without postprocessing. The solid line corresponds to the actual state, while the dashed line corresponds to the estimated state.

classification accuracy is near 70%. Due to the modelling errors and measurement noise the performance is poor and frequent changes can be seen in the estimated discrete state.

The classification performance can be improved by utilization of soft computing techniques and postprocessing. The classifiers utilized in the current work are different ANNs and SVMs. The idea of the postprocessing is that it is highly unlikely that the process changes state from nominal to faulty and back rapidly. Therefore, the estimated state is allowed to change only when the classifier indicates the corresponding change for at least N consecutive time instants. With too small values of N it is not possible to prevent the frequent and unnecessary changes in the state estimation. On the other hand, too high values of N would result in long detection delays.

Several types of feedforward networks have been used, all with *tansig* activation function in both the hidden and the output layer, but with different number of neurons. It has been found that 20 nodes in the hidden layer are enough for obtaining an accurate classification. Fault detection performance obtained with different choices of N is illustrated in the Fig. 1. The detection accuracy without postprocessing is 93.17%. The network is also tested on a different data set in which the fault does not occur. Without postprocessing the network success rate is 89.83%. Result of the FDI procedure with $N = 10$ is presented in Fig. 6. SVMs show rather similar detection accuracy. Several different type of SVMs have been tested, the most representative results are reported in the Table 2 (N is the number of samples used in postprocessing). In the Table 2, N still denotes the number of samples used in postprocessing, poly n denotes a SVM with polynomial kernel of order n, and rbf x denotes a radial basis function kernel with width parameter $\sigma = x$. The satisfactory performance of the neural network based methods, and the positive effect of the post processing are apparent from the results displayed in the Tables

Table 1. Performance of the different fault detection setups based on feedforward neural networks.

N	FDS	NDS	delay
5	95.97 %	98.49 %	6
7	97.64 %	98.49 %	8
10	96.62 %	100 %	11

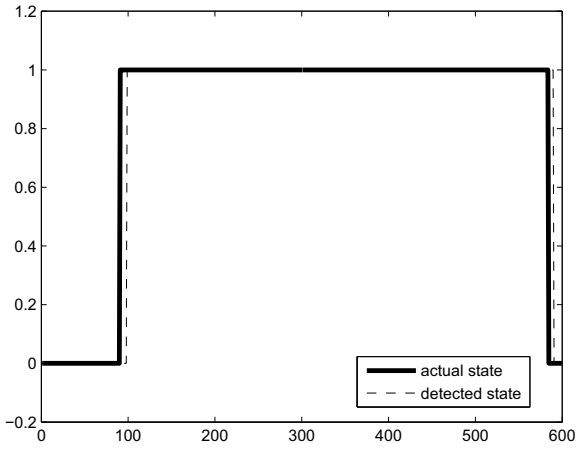


Fig. 6. Fault detection using feedforward neural network with 20 nodes in the hidden layer. $N = 10$ is used during postprocessing. The solid line corresponds to the actual state, while the dashed line corresponds to the estimated state.

Table 2. Performance of the different fault detection setups based on feedforward neural networks.

SVM type	N	FDS	NDS	delay
poly 5	0	86.5 %	77.5 %	-
poly 5	5	96.64 %	91.95 %	11
poly 5	7	96.8 %	100 %	13
poly 5	10	95.76 %	100 %	16
rbf 1	0	87.83 %	65.83 %	-
rbf 1	5	91.10 %	69.97 %	7
rbf 1	7	93.77 %	85.18 %	9
rbf 1	10	96.45 %	92.22 %	12
rbf 0.5	0	83.83 %	85.5 %	-
rbf 0.5	5	97.48 %	100 %	4
rbf 0.5	7	96.8 %	100 %	13
rbf 0.5	10	95.77 %	100 %	16

6. CONCLUSIONS

The present work addressed the problem of fault-detection using model-based approach with second-order sliding-mode observers for residual generation and soft computing for the residual evaluation. The faults have been modeled as abrupt changes in the system dynamics, and a switched affine dynamics has been considered. The FDI problem has been reduced to that of discrete state estimation.

The present study demonstrates that in practice it is not easy to achieve effective fault detection using simple thresholding techniques, mainly due to modeling errors and measurement noise. Several types of soft-computing based classifiers have been trained and tested. Such classi-

fiers are capable of learning complex relationships between the residual signals and the actual state, and are, therefore, more suitable for industrial applications.

Also, it has been perceived that a proper postprocessing procedure is necessary in order to prevent rapid changes of the detected state, even in the cases when soft-computing has been utilized. In fact, the postprocessing step enables us to construct an FDI system which is at the same time robust and agile: with swift detection capabilities and low false alarm rate.

Next investigations will be devoted to generalize the present approach to appropriate classes of nonlinear switched dynamics.

REFERENCES

V. Kecman (2001) Learning and soft-computing. MIT Press.

A. Levant (2003) Higher-order sliding modes, differentiation and output-feedback control. Int. J. Contr., 76(2003), 924–941

J. Moreno, M. Osorio (2009) A Lyapunov approach to second-order sliding mode controllers and observers. Proc $47-th$ Decision Contr Conf 2856–2861, Cancun, December 9–11.

A. Polyakov, A. Poznyak, Lyapunov function design for finite-time convergence analysis: "Twisting" controller for second-order sliding mode realization. Automatica 45(2), 444–448, (2009).

B. Scholköpf, A.J. Smola (2002) Learning with kernels. The MIT Press

S. Simani, C. Fantuzzi, and R.J. Patton, "Model-based fault diagnosis in dynamic systems using identification techniques", Springer-Verlag, 2002. ISBN 1852336854. Advances in Industrial Control Series. London, UK.

W. Wan, L. Li, D. Zhou and K. Liu , "Robust state estimation and fault diagnosis for uncertain hybrid nonlinear systems ", Nonlinear analysis: Hybrid systems, vol. 1, 2007.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Flow and Infiltration Estimation in Open Channel Hydraulic Systems

Siro Pillosu, Alessandro Pisano, Elio Usai

Dept. of Electrical and Electronic Engineering (DIEE)
University of Cagliari (e-mail: siro.pillosu,pisano,eusai@diee.unica.it).

Abstract: This paper addresses a problem of state and disturbance estimation for an open channel hydraulic system. More precisely, a cascade of n canal reaches, joined by gates, is considered, and, by using measurements of the water level in three points per reach, we design an observer capable of estimating both the infiltration and discharge in the middle point of each reach. To facilitate the observer design, the system dynamics is modeled by considering a linearized approximation of the underlying nonlinear dynamics around the subcritical uniform flow condition. The proposed solution is based on the unknown-input and proportional-integral observers theory. Simulation results are discussed to verify the effectiveness of the proposed schemes.

Keywords: Unknown-input observers, strong observability, open channel hydraulic system.

1. INTRODUCTION

Most open-channel hydraulic systems are currently manually operated by flow control gates. Medium-term research goals in this field are to operate those systems automatically in order to improve water distribution efficiency and safety. A key problem is to reconstruct all the informations needed for control or monitoring purposes (water levels, flow rates, and infiltrations), some of which are intrinsically impossible or difficult to measure, by reducing the number of required sensors in the field. The aim of this paper is the development of new estimation algorithms for a cascade chain of open channel hydraulic systems subject to unmeasurable disturbances. Open channel hydraulic systems are described by two nonlinear coupled partial differential equations (Saint-Venant Equations). The two main methods used to solve Saint-Venant equations are basically the finite-difference method (Strelkoff et al. (1970)) and the finite-element method (Colley et al. (1976)). Many variations and improvements have been proposed that have allowed getting better simulation results, but at the same time they increased model complexity and, correspondingly, the difficulty to embed them in model-based observers or controllers. For the considered open channel hydraulic system it has been shown how a three-point orthogonal collocation model (Villadsen et al. (1978)) can be used to design a model-based nonlinear controller with guaranteed properties of closed loop stability. It has even been shown how the behavior obtained solving the simplified equations of the collocation model is close enough to that obtained using high-accuracy solvers of the commercial software packages (Dulhoste et al. (2001); Besancon et al. (2001)). In (Besancon et al. (2001)) a three-point collocation-based nonlinear model of a single-reach irrigation canal was developed. In that model, the three points of interest where the system variables are evaluated are the canal start, middle and end points, respectively, and a constant uncertain infiltration is taken into account. An observer for the level variables and for the constant infiltration was designed by measuring the level in the middle of the reach and the upstream and downstream flows. Our developments take, as a starting point, the results presented in (Besancon et al. (2001)) for a single-reach canal. Here the main task is to **consider a cascade of n canal reaches** instead of a single reach, and to relax some of the standing modeling assumptions made in Besancon et al. (2001). More precisely we:

i.) dispense with the need of flow rate measurements by allowing only level measurements;

ii.) consider a time varying infiltration;

iii.) consider, in the case of absence of infiltration, a number of sensors less than those normally required.

Section 2 recalls the Saint-Venant equations that rigorously describe the dynamics of a canal reach. In Section 3 the three-point collocation based nonlinear model of a single reach presented in (Besancon et al. (2001)), is extended to a cascade of n canal reaches, and in Subsection 3.1 the linearization around the uniform flow condition of the obtained model is computed by linearizing the discharge balance equations through the gates. In Section 4, the two estimation problems addressed in the manuscript (called "Problem 1" and "Problem 2") are stated. Only level measurement are permitted in both. Problem 1 involves the simplifying assumption of absence of infiltration, and, as a counterpart, it considers certain level variables to be unavailable for measurements. The flow variables are wanted to be estimated along with the unmeasured level variables. Problem 2 deals with the more general non-zero infiltration case but requires the measurement of the level variables in all of the collocation points (three points per channel). The flow variables are wanted to be estimated again, along with the unknown infiltrations, after a **finite-time estimation transient**. Problem 1 and

Problem 2 give rise to an observation problem for Linear Time-Invariant System with Unknown Inputs (LTISUI). In Section 5 a method for state estimation and unknown input reconstruction in LTISUI is recalled. The approach is based on the assumption of "**Strong Observability**" (Molinari et al. (1976); Hautus et al. (1983); Bejarano et al. (2007)), a structural geometric restriction involving the matrices of the LTISUI mathematical model. Such restriction has different, although equivalent, formulations, the simplest of which establishes that a certain matrix pair should be observable. In the Section 6, a case study of a canal with rectangular section and three reaches in cascade is illustrated. The parameters of the linearized models previously developed are computed. In the successive Sections 7 and 8 the techniques described in the Section 5 are applied to solve the estimation Problem 1 and Problem 2, respectively. It is shown, in both cases, that the structural requirement of strong observability holds for the resulting models, and corresponding simulation results are illustrated and commented, which will confirm the expected performance of the suggested observers.

2. FORMULATION OF THE PROBLEM

Water flow dynamics in an open channel are governed by the Saint-Venant Equations

$$\frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = w \tag{1}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial (Q^2/S)}{\partial x} + gS\left(\frac{\partial H}{\partial x} - I + J\right) = k_q(w)\frac{Q}{S}w \tag{2}$$

where $x \in [0, L]$ is the spatial variable (L being the channel length), t being the time variable, and $S(x,t)$, $Q(x,t)$ and $H(x,t)$ being the wet section, water flow rate and relative water level, respectively, and the term $w = w(x,t)$ in the right-hand side of (1), (2) represents the infiltration. J represents the friction term, which has the expression $J = \frac{Q|Q|}{Di^2}$, $Di = kS\left(\frac{S}{P}\right)^{\frac{2}{3}}$, with k being Strickler friction coefficient and $P(x,t)$ being the transversal wet length, and I is the canal slope. Finally the term $k_q(w)$ is 0 if $w \geq 0$ and 1 if $w < 0$. In this paper we refer to the case of positive infiltration ($w \geq 0$) and to canals with rectangular section, so if B is the constant canal width one has that $S = BH$ and $P = B + 2H$.
Thus, model (1)-(2) can be rewritten in terms of the Q and H variables only, and, in particular, eq. (1) modifies as

$$\frac{\delta H}{\delta t} = -\frac{1}{B}\frac{\delta Q}{\delta x} + \frac{1}{B}w \tag{3}$$

If the slope of the canal is low, as it is the case, e.g., in irrigation channels, it can be assumed subcritical flow condition which makes it reasonable to complement (1) with the Dirichlet boundary conditions for the upstream and downstream value of the flow

$$Q(0,t) = Q_U(t); \quad Q(L,t) = Q_D(t); \tag{4}$$

Initial conditions are given by:

$$H(x,0) = H_0(x); \quad Q(x,0) = Q_0(x) \tag{5}$$

3. COLLOCATION-BASED FINITE-DIMENSIONAL MODEL

In (Dulhoste et al. (2001)) it was shown that Saint Venant equation can be approximated by ordinary differential equations of finite dimension using a collocation point Galerkin method (Fletcher et al. (1984)). It has been also shown that three collocation points located at the canal upstream, middle, and downstream points (say points A, M, and B, respectively) leads to a sufficiently accurate representation for observation and control purposes. Consider a cascade of n canal reaches connecting the two upstream and downstream reservoirs, separated by $n+1$ adjustable gates, and subject to infiltration losses, as represented in the Figure 1. By choosing three collocation points for each channel, and using the same notation as before to denote the resulting points A_i, M_i, B_i ($i = 1, 2, ..., n$) the equation (3) can be discretized as follows (Besancon et al. (2001)):

$$\begin{aligned}
\dot{H}_{Ai} &= \frac{1}{B_i L_i}[-4Q_{Mi} + 3Q_{Ai} + Q_{Bi}] + \frac{w_i}{B_i} \\
\dot{H}_{Mi} &= \frac{1}{B_i L_i}[Q_{Ai} - Q_{Bi}] + \frac{w_i}{B_i} \\
\dot{H}_{Bi} &= \frac{1}{B_i L_i}[4Q_{Mi} - Q_{Ai} - 3Q_{Bi}] + \frac{w_i}{B_i}
\end{aligned} \tag{6}$$

$$Q_{Ai} = \eta_i \Sigma_i \sqrt{2g(H_{Bi-1} - H_{Ai})} \tag{7}$$

$$\begin{aligned}
Q_{Bi} &= Q_{Ci} + Q_{Ai+1} = \\
&= Q_{Ci} + \eta_{i+1}\Sigma_{i+1}\sqrt{2g(H_{Bi} - H_{Ai+1})}
\end{aligned} \tag{8}$$

where H_{Ai}, H_{Mi} and H_{Bi} ($i = 1, 2, ..., n$) are the state variables, Q_{Ai}, Q_{Mi} and Q_{Bi} ($i = 1, 2, ..., n$) denote the flow at the collocation points, w_i ($i = 1, 2, ..., n$) is the infiltration in the $i-$th reach, η_j and Σ_j ($j = 1, 2, ..., n+1$) are the discharge coefficients and the opening sections of the i-th gate, and Q_{Ci} is the widthdrawal request from the users. H_{B0} and H_{An+1} represent the constant levels in the upstream and downstream reservoirs. Considering (7) and (8) into (6) one obtains a more compact expression for the system nonlinear dynamics:

$$\begin{aligned}
\dot{H}_{Ai} &= \frac{1}{B_i L_i}[f_A(H_{Bi-1}, H_{Bi}, H_{Ai}, H_{Ai+1}, \Sigma_i, \Sigma_{i+1}) - \\
&\qquad -4Q_{Mi} + Q_{Ci}] + \frac{w_i}{B_i} \\
\dot{H}_{Mi} &= \frac{1}{B_i L_i}[f_M(H_{Bi-1}, H_{Bi}, H_{Ai}, H_{Ai+1}, \Sigma_i, \Sigma_{i+1}) - \\
&\qquad -Q_{Ci}] + \frac{w_i}{B_i} \\
\dot{H}_{Bi} &= \frac{1}{B_i L_i}[f_B(H_{Bi-1}, H_{Bi}, H_{Ai}, H_{Ai+1}, \Sigma_i, \Sigma_{i+1}) + \\
&\qquad +4Q_{Mi} - 3Q_{Ci}] + \frac{w_i}{B_i}
\end{aligned} \tag{9}$$

with implicit definition of the nonlinear functions $f_A(\cdot)$, $f_M(\cdot)$ and $f_B(\cdot)$. The dynamic relationship between the middle point flow variable Q_{Mi} and the remaining system variables can be derived by generalizing the "one-reach canal" relationship (Dulhoste et al. (2001)) as follows:

$$\begin{aligned}
\dot{Q}_{Mi} &= \psi_{iq}(Q_{Ai}, Q_{Bi}, Q_{Mi}, H_{Ai}, H_{Mi}, H_{Bi}, w_i) \\
&= gB_i H_{Mi}\left(I + \frac{H_{Ai} - H_{Bi}}{L_i}\right) + \left(\frac{2(Q_{Ai} - Q_{Bi})}{BL_i}\right)\frac{Q_{Mi}}{H_{Mi}} +
\end{aligned} \tag{10}$$

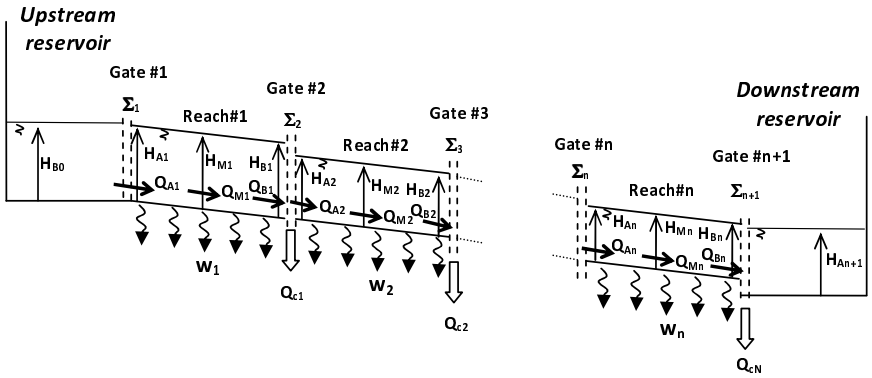


Fig. 1. Cascade of n canal reaches with infiltration losses.

$$+\left(\frac{H_{Bi}-H_{Ai}}{B_i L_i H_{Mi}^2}-\frac{g}{K^2 B_i H_{Mi}(\frac{B_i H_{Mi}}{B_i+2H_{Mi}})^{\frac{4}{3}}}\right)Q_{Mi}^2$$

3.1 Linearized Model

The nonlinear model (9) can be linearized in a vicinity of the uniform flow condition (Corriga et al. (1983)). Let $\overline{Q}_i$ $(i=1,2,...,n)$, denote the flow value in the i-th channel in the uniform flow condition. Let also $\overline{H}_i$ $(i=1,2,...,n)$ be the corresponding water levels, and $\overline{\Sigma}_j$ $(j=1,2,...,n+1)$ be the corresponding values for the gates opening sections. Define the corresponding variables $h_{Ai}, h_{Mi}, h_{Bi}, q_{Mi}$, and σ_i like the deviation of $H_{Ai}, H_{Mi}, H_{Bi}, Q_{Mi}$ and Σ_i from the uniform condition, then relation (7) can be linearized as follows in a vicinity of the uniform flow condition (Corriga et al. (1983))

$$Q_{Ai}=\overline{Q}_i+a_i\sigma_i(t)+b_i[h_{Bi-1}(t)-h_{Ai}(t)] \tag{11}$$

with the coefficients a_i and b_i as follows

$$a_i=\eta_i\sqrt{2g(\overline{H}_{i-1}-\overline{H}_i)};\quad b_i=\frac{\eta_i\overline{\Sigma}_i\sqrt{2g}}{\sqrt{2(\overline{H}_{i-1}-\overline{H}_i)}} \tag{12}$$

The corresponding linearized form for (8) can be derived from the next continuity equation

$$Q_{Bi}=Q_{Ai+1}+Q_{Ci},\qquad i=1,...n \tag{13}$$

which leads to

$$Q_{Bi}=Q_{Ci}+\overline{Q}_{i+1}+a_{i+1}\sigma_{i+1}+b_{i+1}[h_{Bi}-h_{Ai+1}] \tag{14}$$

with $i=1,2,...,n$ and the deviation variables h_{An+1} and h_{B0} customarily set both to zero as a consequence of the fact that the water level in the upstream and downstream reservoirs is supposed to keep constant

$$h_{An+1}=0;\quad h_{B0}=0 \tag{15}$$

Substituting (11)-(14) into (6)-(8), considering the next continuity condition

$$\overline{Q}_i=\overline{Q}_{i+1}+Q_{Ci};\qquad i=1,...n \tag{16}$$

and considering vectors

$$\begin{aligned}&h=[h_{A1}\ h_{M1}\ h_{B1}\ ...h_{An}\ h_{Mn}\ h_{Bn}]^T, h\in R^{3n}\\&\sigma=[\sigma_1\ \sigma_2...\sigma_{n+1}]^T,\quad \sigma\in R^{n+1}\end{aligned} \tag{17}$$

$$q_M=[q_{M1}\ q_{M2}...q_{Mn}]^T,\quad q\in R^n \tag{18}$$

$$w=[w_1\ w_2...w_n]^T\quad w\in R^n \tag{19}$$

$$\tag{20}$$

it is possible to rewrite the system (9) in the compact state space form

$$\dot{h}=Ah+M_\sigma\sigma+M_q q_M+M_w w \tag{21}$$

with implicitly defined constant matrices A, M_σ, M_q and M_w of appropriate dimension. Vector q_M depends (dynamically) on the other system variables according to the nonlinear differential equation

$$\dot{q}_M=\psi(h,q_M,\sigma,w)=[\psi_{1q}(\cdot),\ \psi_{2q}(\cdot),\ \dots\ \psi_{nq}(\cdot)]^T \tag{22}$$

and the functions $\psi_{iq}(\cdot)$ $(i=1,2,...,n)$ are defined in (10). Note that the nonlinear dynamics (22) **need not to be linearized** since vector q_M is going to be treated as an unknown input of the system, rather than as a part of the system state. Since the first and last equation in (9) are not affected by the level variables h_{Mi}, it is possible to consider a reduced-order version of system (21) where the state vector h is replaced by the reduced-order version

$$\tilde{h}=[h_{A1}\ h_{B1}\ h_{A2}\ ...h_{An}\ h_{Bn}]^T\quad \tilde{h}\in R^{2n} \tag{23}$$

The corresponding reduced-order state space model is given by

$$\dot{\tilde{h}}=\tilde{A}\tilde{h}+\tilde{M}_\sigma\sigma+\tilde{M}_q q_M+\tilde{M}_w w, \tag{24}$$

whose matrices $\tilde{A},\tilde{M}_\sigma,\tilde{M}_q,\tilde{M}_w$ can be trivially derived by removing selected rows and columns from the matrices A, M_σ, M_q, M_w of the full-order model (21)

4. FLOW AND INFILTRATION ESTIMATION PROBLEM STATEMENT

In this paper we make reference to the linearized dynamics (21) and we address two distinct state and disturbance estimation problems under the common constraint that **only level measurements are allowed**. In fact, flow sensors are expensive devices, and, furthermore, it is difficult to make precise flow measurements in the real scenarios. Level sensors are on the contrary cheap and accurate. Vector σ is supposed to be known, while vectors w and q_M are both unmeasurable. We cast the following problems:

Problem 1. By measuring only a portion of the reduced-order state vector $\tilde{h}$, and assuming no infiltrations ($w = 0$), estimate the flow vector q_M and reconstruct the unmeasured part of vector $\tilde{h}$.

Problem 2. By measuring the full vector h, reconstruct the infiltration vector w and the flow vector q_M in finite time.

Both Problems 1 and 2 will be solved by making use of unknown-input observers (UIO) under the requirement of "strong observability" (Molinari et al. (1976); Hautus et al. (1983)) for certain subsystems that shall be specified later on. The UIO design for general, strongly observable, linear time-invariant systems with unknown inputs is recalled in the next Section V. The successive sections VI and VII address the Problem 1 and 2, respectively, by exploiting the presented UIO design framework.

5. STRONG OBSERVABILITY AND UIO DESIGN FOR LINEAR SYSTEMS WITH UNKNOWN INPUTS

Consider the linear time invariant dynamics

$$\begin{aligned} \dot{x} &= Ax + Gu + F\xi \\ y &= Cx \end{aligned} \tag{25}$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^p$ are the state and output variables, $u(t) \in \mathbb{R}^h$ is a *known input* to the system, $\xi(t) \in \mathbb{R}^m$ is an *unknown input* term, and A, G, F, C are known constant matrices of appropriate dimension. Let us make the following assumptions:

A1. The matrix triplet (A, F, C) is strongly observable
A2. $rank\,(CF) = rank\,F = m$.

If conditions A1 and A2 are satisfied then it can be systematically found a state coordinates transformation together with an output coordinates change which decouple the unknown input ξ from a certain subsystem in the new coordinates. Such a transformation is outlined below. For the generic matrix $J \in \mathbb{R}^{n_r \times n_c}$ with $rankJ = r$, we define $J^{\perp} \in \mathbb{R}^{n_r - r \times n_r}$ as a matrix such that $J^{\perp}J = 0$ and $rankJ^{\perp} = n_r - r$. Matrix $J^{\perp}$ always exists and, furthermore, it is not unique [1]. Let $\Gamma^{+} = [\Gamma^T\Gamma]^{-1}\Gamma^T$ denote the left pseudo-inverse of Γ and it is such that $\Gamma^{+}\Gamma = I_{n_c}$, with I_{n_c} being the identity matrix of order n_c. Consider the following transformation matrices T and U:

$$T = \begin{bmatrix} F^{\perp} \\ (CF)^{+} C \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \tag{26}$$

$$U = \begin{bmatrix} (CF)^{\perp} \\ (CF)^{+} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \tag{27}$$

$$(CF)^{+} = \left[(CF)^T (CF) \right]^{-1} (CF)^T \tag{28}$$

and the transformed state and output vectors

$$\bar{x} = Tx, \qquad \bar{y} = Uy \tag{29}$$

Consider the following partitions of vectors $\bar{x}$ and $\bar{y}$

$$\bar{x} = \begin{bmatrix} T_1 x \\ T_2 x \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \quad \bar{x}_1 \in R^{n-m} \quad \bar{x}_2 \in R^m \tag{30}$$

[1] A Matlab instruction for computing $M_b = M^{\perp}$ for a generic matrix M is `Mb = null(M')'`

$$\bar{y} = \begin{bmatrix} U_1 y \\ U_2 y \end{bmatrix} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix}; \quad \bar{y}_1 \in R^{p-m} \quad \bar{y}_2 \in R^m \tag{31}$$

After simple algebraic manipulations the **transformed system** in the new coordinates takes the form:

$$\begin{aligned} \dot{\bar{x}}_1 &= \bar{A}_{11}\bar{x}_1 + \bar{A}_{12}\bar{x}_2 + F^{\perp}Gu \\ \dot{\bar{x}}_2 &= \bar{A}_{21}\bar{x}_1 + \bar{A}_{22}\bar{x}_2 + (CF)^{+} CGu + \xi \\ \bar{y}_1 &= \bar{C}_1\bar{x}_1 \\ \bar{y}_2 &= \bar{x}_2 \end{aligned} \tag{32}$$

with the matrices $\bar{A}_{11}$, ..., $\bar{A}_{22}$ and $\bar{C}_1$ such that

$$\begin{bmatrix} \bar{A}_{11} & \bar{A}_{11} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} = TAT^{-1}; \quad \bar{C}_1 = (CF)^{\perp} CT_1 \tag{33}$$

It turns out that the triple (A, C, F) is strongly observable if, and only if, **the pair $(\bar{A}_{11}, \bar{C}_1)$ is observable** (Molinari et al. (1976); Hautus et al. (1983)). In light of the Assumption A1, this property, that can be also understood in terms of a simplified algebraic test to check the strong detectability of a matrix triple, opens the way to design stable observers for the state of the transformed dynamics (32). The peculiarity of the transformed system (32) is that $\bar{x}_2$ constitutes a part of the transformed output vector $\bar{y}$. Hence, **the observation of the state of systems (32) can be accomplished by estimating $\bar{x}_1$ only**, whose dynamics is not affected by the unknown input vector. The observability of the $(\bar{A}_{11}, \bar{C}_1)$ permits the implementation of the following Luenberger observer for the $\bar{x}_1$ subsystem of (32):

$$\dot{\hat{\bar{x}}}_1 = \bar{A}_{11}\hat{\bar{x}}_1 + \bar{A}_{12}\bar{y}_2 + F^{\perp}Gu + L(\bar{y}_1 - \bar{C}_1\hat{\bar{x}}_1) \tag{34}$$

which gives rise to the error dynamics

$$\dot{e}_1 = (A - LC)e_1, \qquad e_1 = \hat{\bar{x}}_1 - \bar{x}_1 \tag{35}$$

whose eigenvalues can be arbitrarily located by a proper selection of the matrix L. Therefore, with properly chosen L we have that $\hat{\bar{x}}_1 \to \bar{x}_1$ as $t \to \infty$, which implies that the overall system state can be reconstructed by the following relationships

$$\hat{x} = T^{-1}[\hat{\bar{x}}_1 \;\; \bar{y}_2]^T \tag{36}$$

Note that the convergence of $\hat{\bar{x}}_1$ to $\bar{x}_1$ is exponential and can be made as fast as desired.

5.1 Reconstruction of the unknown inputs

An estimator can be designed which gives an exponentially converging estimate of the unknown input vector ξ. Consider the following estimator dynamics

$$\dot{\hat{\bar{x}}}_2 = \bar{A}_{21}\hat{\bar{x}}_1 + \bar{A}_{22}\bar{y}_2 + (CF)^{+} CGu + v(t) \tag{37}$$

with the estimator injection input $v(t)$ yet to be specified. Let us assume that a constant Ξ_d can be found such that

$$|\dot{\xi}(t)| \leq \Xi_d \tag{38}$$

Define

$$s = \hat{\bar{x}}_2 - \bar{y}_2 = \hat{\bar{x}}_2 - \bar{x}_2 \tag{39}$$

By (37) and (32), the dynamics of the sliding variable σ takes the following form

$$\dot{s} = f(t) - v(t), \quad f(t) = \bar{A}_{21}\bar{e}_1(t) + \xi(t) \tag{40}$$

Considering (35) , the time derivative of the uncertain term $f(t)$ can be evaluated as

$$\dot{f}(t) = \bar{A}_{21}(A - LC)e_1(t) + \dot{\xi}(t) \tag{41}$$

where $e_1(t)$ is exponentially vanishing. Then, considering (38), by taking any $\overline{\Psi} > \Xi_d$, the next condition

$$|\dot{f}(t)| \leq \overline{\Psi}, \qquad t > T^*, \qquad T^* < \infty \tag{42}$$

will be established starting from a finite time instant $t = T^*$ on. As shown in (Levant et al. (1993)), if the injection input $v(t)$ is designed according to

$$v\,(t) = \lambda\,|\sigma|^{1/2}\,sign\sigma + v_1; \quad \dot{v}_1\,(t) = \alpha sign\;\sigma, \tag{43}$$

$$\alpha > \overline{\Psi}, \qquad \lambda > \frac{1-\theta}{1+\theta}\sqrt{\frac{\overline{\Psi}-\alpha}{\overline{\Psi}+\alpha}}, \quad \theta \in (0,1) \tag{44}$$

It steers to zero in finite time both σ and its time derivative $\dot{\sigma}$. Therefore, condition

$$v(t) = \xi(t) + \bar{A}_{21}(A - LC)e_1(t) \tag{45}$$

holds after a finite transient time. It readily follows from the constraction property of $e_1(t)$ that the second term in the right hand side of (45) is exponentially vanishing, which implies that $|v(t) - \xi(t)| \to 0$ as $t \to \infty$ and, furthermore, the convergence takes place exponentially. Therefore, under the condition (38), the estimator (37), (39), (43)-(44) allows one to reconstruct the unknown input vector ξ acting on the original system (25).

6. CASE STUDY

We shall consider a test canal with rectangular section and three reaches in cascade with width of 2 m and length respectively 4, 5 and 2 km; the supposed discharge coefficient is $\eta = 0.6$; the roughness coefficient $K_s = 50\frac{m^{\frac{1}{3}}}{s}$; and the constant slope is $I = 0,001$;. The water level in upstream reservoir is $H_{B0} = 3m$; and in downstream reservoir is $H_{A4} = 1m$. We have the following withdrawals ($\frac{m^3}{s}$): $Q_{C1} = 2,\ Q_{C2} = 2,\ Q_{C3} = 1$. The opening section of the $4-th$ gate is kept constant $\Sigma_4 = 0.538m^2$. The uniform flow condition is characterized by the following values: flow rates ($\frac{m^3}{s}$): $\overline{Q}_1 = 6.017,\ \overline{Q}_2 = 4.007,\ \overline{Q}_3 = 1.966$; levels ($m$): $\overline{H}_1 = 2,40,\ \overline{H}_2 = 1.72,\ \overline{H}_3 = 0.99$; opening sections ($m^2$): $\overline{\Sigma}_1 = 2.923,\ \overline{\Sigma}_2 = 1.829,\ \overline{\Sigma}_3 = 0.866$.
The opening sections of the gates 1, 2 and 3 are adjusted according to

$$\begin{aligned} \Sigma_1 &= \overline{\Sigma}_1 + 0.8sin[(2\pi/1000)t] \\ \Sigma_2 &= \overline{\Sigma}_2 + 0.5sin[(2\pi/1000)t] \\ \Sigma_3 &= \overline{\Sigma}_3 + 0.3sin[(2\pi/1000)t] \end{aligned} \tag{46}$$

and the infiltration variables are set as $w_1 = w_2 = w_3 = 0.1e^{-0.001t}$.

7. FLOW ESTIMATION WITH PARTIAL LEVEL MEASUREMENTS AND NO INFILTRATION (PROBLEM 1)

Consider the reduced-order linearized dynamics (24) by assuming no infiltration (i.e., $w = 0$) according to the statement of Problem 1 (see Section 4):

$$\dot{\tilde{h}} = \tilde{A}\tilde{h} + \tilde{M}_\sigma\sigma + \tilde{M}_q q_M \tag{47}$$

Only five elements of vector $\tilde{h}$ are supposed to be measured, according to the next output equation

$$\tilde{y} = \tilde{C}\tilde{h}; \quad \tilde{C} = \begin{bmatrix} 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&1&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&1 \end{bmatrix} \tag{48}$$

It is worth noting that system (47)-(48) is a special case of the general dynamics (25) with the modified notation $\tilde{h} = x$, $\sigma = u$, $q_M = \xi$. It is easy to check that the matrix triplet $(\tilde{A}, \tilde{M}_q, \tilde{C})$ is strongly observable, hence the design method previously described can be applied to reconstruct, both, the unknown vector q_M and the **unmeasured** level variable h_{A2}. By computing the matrices of the transformed system dynamics, it can be readily verified that $(\bar{A}_{11}, \bar{C}_1)$ is an observable pair. Therefore, it can be implemented the suggested scheme (34),(37), (43)-(44), with the observer gain matrix

$$\tilde{L} = 10^{-3}\begin{bmatrix} -1.7 & 2.1 \\ -0.6 & -3.0 \\ -3.3 & -0.9 \end{bmatrix} \tag{49}$$

that has been computed in order to assign to the observation error matrix $(\bar{A}_{11} - \tilde{L}\bar{C}_1)$ the desired spectrum of eigenvalues [-0.05,-0.05,-0.005]. The gain parameters α and λ of the unknown input reconstruction algorithm are set as $\alpha = 1.5\sqrt{5}, \lambda = 5$. The performance of the observer are tested by means of simulations made in the Matlab-Simulink environment. The system and the observers are integrated by fixed step Runge-Kutta method, with the integration step $T_s = 10^{-4}s$. The system's initial conditions are: $\tilde{h}(0) = [0.1, 0.1, 0.1, 0.1, 0.1, 0.1]$. All the observer's initial conditions are set to zero. For simulation purposes the actual Q_M profiles are generated by solving the corresponding system of nonlinear differential equations (22), with the initial conditions $Q_M(0) = [6.017, 4.007, 1.966]$. The next Figures 2 show the actual and estimated profiles of the unknown flow variable q_{M1} during the TEST 1, of duration 500 seconds. The left and right plot show the transient and long term behaviour. After a transient of about twenty seconds, the estimated flow converges towards the actual one. The estimation performance for the flow variables q_{M2} and q_{M3} is pretty equivalent and it is not shown for brevity. The reconstruction of the

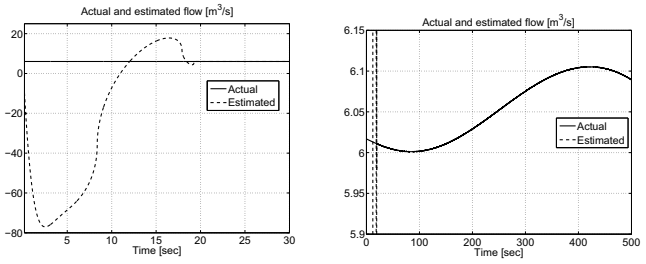


Fig. 2. Actual and estimated flow variable q_{M1} in the TEST 1

unmeasured level variable h_{A2} is shown in the Figure 3. The left and right plot show the transient and long term behaviour, respectively.

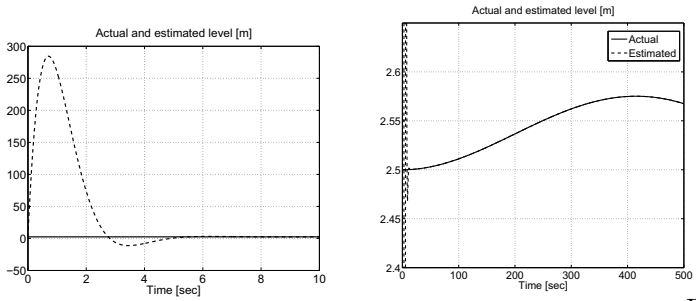


Fig. 3. Actual and estimated level variable h_{A2} in the TEST 1

8. FLOW AND INFILTRATION ESTIMATION WITH FULL LEVEL MEASUREMENTS (PROBLEM 2)

We consider here the full-order model (21), and we assume the availability for measurement of the entire state vector $h \in \Re^{3n} \equiv \Re^9$, i.e. the considered output is $y = h$. The problem here is to reconstruct after a finite observation transient the unknown input terms, namely the flow q_M in the middle of the channels and the infiltration vector w. System (21) along with the considered output equation $y = h$ can be rewritten as

$$\begin{aligned} \dot{h} &= Ah + M_\sigma \sigma + [M_q \;\; M_w] \cdot [q_M \;\; w]^T \\ y &= Ch \end{aligned} \tag{50}$$

which belongs to the class of dynamics (25) with the modified notation $h = x$, $\sigma = u$, $[q_M^T \quad w^T]^T = \xi$, $M_\sigma = G$ and $[M_q \quad M_w] = F$, and with the output transformation matrix $C = I$ (I being the identity matrix of order $3n = 9$). It is easy to check that the matrix triplet $(\tilde{A}, [M_q \quad M_w], I)$ is strongly observable. Since the state vector is supposed to be fully available, a simplified version of the design methodology previously described can be applied to reconstruct the unknown vectors q_M and w. By computing the matrices of the transformed system dynamics, it can be readily verified that $(\bar{A}_{11}, \bar{C}_1)$ is an observable pair. Since the state vector is already available, only the observer (39)-(44) for reconstructing the unknown input is necessary. The gain parameters α and λ of the observation algorithm are set as $\alpha = 1.5\sqrt{5}, \lambda = 5$. The performance of the observer is tested by means of simulations made in the same way of the previous case. For simulation purposes the actual Q_M profiles are generated by solving the corresponding system of nonlinear differential equations (22), with the initial conditions $Q_M(0) = [6.017, 4.007, 1.966]$ The next Figures 4 shows the actual and estimated profiles of the unknown flow variable q_{M2} during the TEST 2, of duration 100 seconds. The left and right plot show the transient and long term behaviour. After a transient of about half a second, the estimated flow converges towards the actual one. The estimation performance for the flow variables q_{M1} and q_{M3} is pretty equivalent and it is not shown for brevity. The reconstruction of of the unknown infiltration variable w_3 is shown in the Figure 5. The left and right plot show the transient and long term behaviour, respectively.

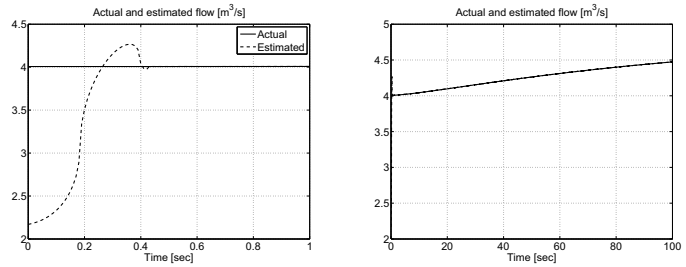


Fig. 4. Actual and estimated flow variable q_{M2} in the TEST 2

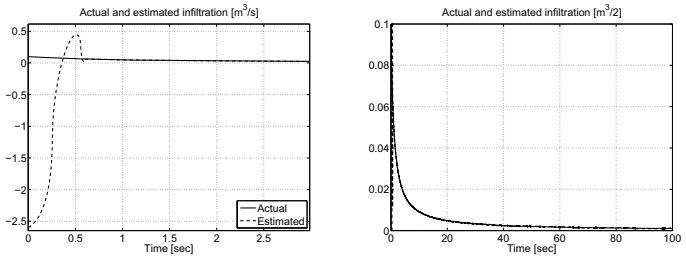


Fig. 5. Actual and estimated infiltration variable w_3 in the TEST 2

REFERENCES

Strelkoff, T. Numerical solution of Saint-Venant equation,Journal of Hydraulical Engineering. Division ASCE, Vol.96, No.HY1, pp.223-252. 1970.

Colley, R.L. and Moin, S.A. Finite element solution of Saint-Venant equations,Journal of Hydraulical Engineering. Division ASCE, Vol.102, No.HY6, pp.759-775. 1976.

Villadsen J.V., and Michelsen M.L. Solution of differential equations models by polynomial approximation. Prentice Hall. Englewood Cliffs, N.J. 1978.

Dulhoste, J.-F., Besancon, G., and Georges, D. 2001. Nonlinear control of water flow dynamics by inputoutput linearization based on a collocation method model. European Control Conf.

Besancon, G.,Dulhoste, J.-F. and Georges, D. 2001. Nonlinear observer design for water level control in irrigation canals. Conf. on Decision and Control, Dec. 2001.

B.P. Molinari. "A strong contollability and observability in linear multivariable control" IEEE Transaction on Automatic Control, 21(5), pp. 761-764, 1976.

M. L. J. Hautus. "Strong detectability and observers" Linear Algebra and its Applications, 50, pp. 353-368, 1983.

F.J. Bejarano, L. Fridman and A. Poznyak. "Exact state estimation for linear systems with unknown inputs based on hierarchical super-twisting algorithm". Int. J. Robust Nonlinear Control, 17(18), pp. 1734-1753, 2007.

Fletcher C.A.J. Computational Galerkin methods. Springer Series in Computational Physics. Springer-Verlag, 1984.

G. Corriga, S. Sanna, G. Usai, Sub-optimal constant volume control for open-channel networks, Applied Mathematical Modelling (1983).

A. Levant, "Sliding order and sliding accuracy in sliding mode control", Int. J. Contr., 58(1993), 1247-1263.

C. Edwards, S. K. Spurgeon, and R. J. Patton,"Sliding mode observers for fault detection and isolation," *Automatica*, vol. 36, pp. 541–553, 2000.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Unknown Input Observer with sliding mode disturbance estimator for the Diffusion PDE

Alessandro Pisano, Stefano Scodina, Elio Usai

Dept. of Electrical and Electronic Engineering (DIEE)
University of Cagliari (Italy)
Email:{pisano,stefano.scodina,eusai}@diee.unica.it

Abstract: In this note an Unknown Input Observer and a disturbance estimator are proposed for a class of distributed parameter systems. The 1D diffusion equation subject to an uncertain exogenous input is dealt with, and point-wise measurements are considered. The observer/estimator design is carried out by making reference to a finite dimensional modal decomposition of the solution. A combined state and output transformation is applied to the resulting finite dimensional approximation, yielding a special form for the transformed system that allows the implementation of a linear observer for reconstructing the system state and a sliding mode observer for reconstructing the unknown input. Numerical simulations show the applicability of the suggested approach to the considered class of PDEs.

1. INTRODUCTION

There are many kind of systems whose dynamical behavior are described by Partial Differential Equations (PDE), (Curtain, 1995), (Schiesser, 2009). In the last decade, this field has broadened considerably as more realistic models have been introduced and investigated in different areas such as thermodynamics, elastic structures, fluid dynamics and biological systems, to name a few (Imanuilov, 2005), (Mondaini, 2008). In spite of the fact that optimization and control of systems governed by PDE is a very active field of research, no much have been developed for observer design. A main result of this field is described in (Krstic, 2005).

In this paper we consider a problem of state and disturbance estimation for a perturbed version of the diffusion PDE with collocated measurement sensors. It is assumed that the system model is corrupted by an in-domain uncertain, distributed, disturbance. Related investigation were made in (Demetriou, 2005) where an Unknown Input Observer (UIO) was proposed for a class of PDEs in abstract form with a concrete example developed for the perturbed diffusion-convection equation. (Demetriou, 2005) basically extends to Distributed Parameter Systems (DPS) the finite dimensional results of UIO design (Chen, 1996), (Edwards, 2000). The key point of the design method in (Demetriou, 2005) was that of selecting the sensor type and location in such a way that the resulting measurement operator fulfills certain operator equations. After deriving the finite-dimensional modal discretization of the involved PDE, here we follow a similar idea but we develop our design conditions for the approximate finite dimensional model directly. We exploit known results involving the concept of strong observability (Hautus, 1983), (Molinari, 1976) to derive in closed form the expression of a linear, reduced order, UIO. Hence the sensor type and location are to be chosen in such a way that the resulting approximate finite dimensional system is Strongly Observable. Furthermore we suggest a sliding mode estimator to reconstruct the unmeasurable external disturbance affecting the original infinite dimensional dynamics.

The paper is structured as follows. Next Section II contains the problem formulation. In Section III the modal decompostion is developed for the considered class of PDE and the resulting finite dimensional model is developed. In Section 4 and 5 the state observer and the disturbance estimator design are illustrated. Section VI discusses some numerical simulation example and Section 7 presents some final conclusions.

2. FORMULATION OF THE PROBLEM

Consider a physical phenomenon represented by the space and time varying scalar field $z(x,t)$ evolving in a Hilbert space $L_2(0,l)$, where $0 \leq x \leq l$ is the mono-dimensional (1D) spatial variable and $t > 0$ is the time variable. Let the scalar field behavior be governed by a perturbed diffusion (PDE).

$$z_t(x,t) = \theta z_{xx}(x,t) + f(x)w(t) \tag{1}$$

where θ is a positive coefficient called *diffusivity*, $z_t(x,t)$ denotes the partial time derivative and $z_{xx}(x,t)$ denotes the second order spatial derivative. The vector field $f(x)w(t)$ represent an **uncertain** source term where $f(x) \in L_2[0,l]$ is a **known** function, and $w(t)$ **uncertain**. The initial condition (ICs) is:

$$z(x,0) = z_0(x), \;\; z_0(x) \in L_2[0,l] \tag{2}$$

It is well-known, given the initial condition (2), that the Hilbert space-valued heat equation (1) has a unique mild solution $z(x,t)$ (Pazy, 1983). By default, only mild solutions are under study in the present work.

We consider two types of boundaries conditions (BCs), namely, homogenous Neumann BCs

$$\texttt{Neumann-type} \quad z_x(0,t) = z_x(l,t) = 0 \tag{3}$$

or homogenous Dirichlet BCs

$$\texttt{Dirichlet-type} \quad z(0,t) = z(l,t) = 0. \tag{4}$$

The available measurements are the p-dimensional vector $y = [y_1\ y_2 ... y_p]$

In particular we consider point-wise measurements along the spatial domain hence

$$y_k(t) = z(x_s^k, t), \quad k = 1, ..., p \tag{5}$$

where x_s^k is the location of the k-th measurement sensor. The aim of this work is to provide the approximate reconstruction of the state $z(x,t)$ and of the unknown disturbance signal $w(t)$.

3. MODAL REPRESENTATION

By expanding the solution of the equation (1) in an infinite series in terms of eigenfunctions (modal expansion) it is possible to express the solution as

$$z(x,t) = \sum_{n=1}^{\infty} q_n(t)\phi_n(x) \tag{6}$$

where $\phi_n(x)$ are the eigenfunctions corresponding to the the boundary conditions (4) or (3) and $q_n(t)$ are appropriate functions to be determined. Substituting the modal expansion for the solution $z(x,t)$ into the system we obtain an infinite-dimensional system of ODE $q_n(t)$

$$\begin{aligned} \dot{q}_n &= \lambda_n q_n + f_q^n w(t), \quad n = 1, ..., \infty. \\ y_k &= \sum_{n=1}^{\infty} q_n(t) c_k^n, \quad k = 1, ..., p. \end{aligned} \tag{7}$$

where λ_n are the eigenvalues and:

$$\begin{aligned} q_n(0) &= \frac{\int_0^l z_0(x)\phi_n(x)dx}{\int_0^l \phi_n^2(x)dx}, \quad f_q^n = \frac{\int_0^l f(x)\phi_n(x)dx}{\int_0^l \phi_n^2(x)dx} \\ c_k^n &= \int_0^l s_k(x)\phi_n(x)dx = \phi_n(x_s^k) \end{aligned} \tag{8}$$

We consider a finite number N of modes, yielding the finite dimensional approximation of (8):

$$\begin{bmatrix} \dot{q}_1 \\ \vdots \\ \dot{q}_N \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_N \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_N \end{bmatrix} + \begin{bmatrix} f_q^1 \\ \vdots \\ f_q^N \end{bmatrix} w(t) \tag{9}$$

with the output equation

$$\begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} c_1^1 & \cdots & c_1^N \\ \cdots & \cdots & \cdots \\ c_p^1 & \cdots & c_p^N \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_N \end{bmatrix} \tag{10}$$

Finally we can rewrite (9) and (10) in compact form:

$$\begin{aligned} \dot{q}(t) &= \Delta q(t) + F w(t) \\ y(t) &= C q(t) \end{aligned} \tag{11}$$

where $q = [q_1\ q_2\ ...\ q_N]$ and

$$\Delta = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_N \end{bmatrix}, \ C = \begin{bmatrix} c_1^1 & \cdots & c_1^N \\ \cdots & \cdots & \cdots \\ c_p^1 & \cdots & c_p^N \end{bmatrix}, \ F = \begin{bmatrix} f_q^1 \\ \vdots \\ f_q^N \end{bmatrix} \tag{12}$$

4. STRONG OBSERVABILITY AND UIO DESIGN FOR LINEAR SYSTEMS WITH UNKNOWN INPUT

Consider the linear finite-dimensional system (11) obtained from the modal decomposition

$$\begin{aligned} \dot{q} &= \Delta q + F w(t) \\ y &= C q \end{aligned} \tag{13}$$

where $q \in \mathbb{R}^N$ and $y \in \mathbb{R}^p$ are the state and output variables, $w(t) \in \mathbb{R}^m$ is an *unknown input* term, and Δ, F, C are known constant matrices. Let us make the following assumptions:

A1. The matrix triplet (Δ, F, C) is strongly observable
A2. $rank\,(CF) = rank\ F = m$.

The notion of strong observability has been introduced more than thirty years ago (Molinari, 1976), (Hautus, 1983) and (Kratz, 2005) in the framework of the unknown-input observers theory. Recently it has been exploited to design robust observers based on the high-order sliding mode approach (Bejarano, 2007). It has been shown in (Molinari, 1976) that the following statements S_1 and S_2 are equivalent

S_1. The triple (Δ, F, C) is strongly observable.
S_2. The triple (Δ, F, C) has no invariant zeros.

If conditions A1 and A2 are satisfied then it can be systematically found a state coordinates transformation together with an output coordinates change which decouple the unknown input $w(t)$ from a certain subsystem in the new coordinates. Such a transformation is outlined below.
For the generic matrix $J \in \mathbb{R}^{n_r \times n_c}$ with $rank J = r$, we define $J^{\perp} \in \mathbb{R}^{n_r - r \times n_r}$ as a matrix such that $J^{\perp} J = 0$ and $rank J^{\perp} = n_r - r$. Matrix $J^{\perp}$ always exists and, furthermore, it is not unique [1]. Let $\Gamma^+ = [\Gamma^T \Gamma]^{-1} \Gamma^T$ denote the left pseudo-inverse of Γ and let $\Gamma^{\dagger} = \Gamma^T [\Gamma \Gamma^T]^{-1}$ denote the right pseudo-inverse of Γ. It is such that $\Gamma^+ \Gamma = I_{n_c}$ and $\Gamma \Gamma^{\dagger} = I_{n_c}$, with I_{n_c} being the identity matrix of order n_c. Consider the following transformation matrices T and U:

$$T = \begin{bmatrix} F^{\perp} \\ (CF)^+ C \end{bmatrix} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \tag{14}$$

[1] A Matlab instruction for computing $M_b = M^{\perp}$ for a generic matrix M is $\texttt{Mb} = \texttt{null(M')'}$

$$U = \begin{bmatrix} (CF)^{\perp} \\ (CF)^{+} \end{bmatrix} = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \tag{15}$$

and the transformed state and output vectors

$$\bar{q} = Tq, \qquad \bar{y} = Uy \tag{16}$$

Consider the following partitions of vectors $\bar{q}$ and $\bar{y}$

$$\bar{q} = \begin{bmatrix} T_1 q \\ T_2 q \end{bmatrix} = \begin{bmatrix} \bar{q}_1 \\ \bar{q}_2 \end{bmatrix}, \quad \bar{q}_1 \in R^{n-m} \quad \bar{q}_2 \in R^m \tag{17}$$

$$\bar{y} = \begin{bmatrix} U_1 y \\ U_2 y \end{bmatrix} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix}, \quad \bar{y}_1 \in R^{p-m} \quad \bar{y}_2 \in R^m \tag{18}$$

The inverse transformation matrices T^{-1} and U^{-1} take the following form

$$T^{-1} = \left[\left(I - F\left(CF\right)^{+} C \right) \left(F^{\perp}\right)^{\dagger} F \right] = \left[\tilde{T}_1 F \right] \tag{19}$$

$$U^{-1} = \left[\left[I - \left(CF\right)\left(CF\right)^{+} \right] \left[\left(CF\right)^{\perp} \right]^{+} \left(CF\right) \right] \tag{20}$$

The partitioned transformed vectors take the form

$$\begin{aligned} &\bar{q}_1 = F^{\perp} q, \quad \bar{q}_2 = \left(CF\right)^{+} Cq, \\ &\bar{y}_1 = \left(CF\right)^{\perp} y, \quad \bar{y}_2 = \left(CF\right)^{+} y \end{aligned} \tag{21}$$

After simple algebraic manipulations the **transformed system** in the new coordinates is given by:

$$\begin{aligned} \dot{\bar{q}}_1 &= \bar{\Delta}_{11}\bar{q}_1 + \bar{\Delta}_{12}\bar{q}_2 \\ \dot{\bar{q}}_2 &= \bar{\Delta}_{21}\bar{q}_1 + \bar{\Delta}_{22}\bar{q}_2 + w(t) \\ \bar{y}_1 &= \bar{C}_1\bar{q}_1 \\ \bar{y}_2 &= \bar{q}_2 \end{aligned} \tag{22}$$

with the matrices $\bar{\Delta}_{11}$, ..., $\bar{\Delta}_{22}$ such that

$$\begin{bmatrix} \bar{\Delta}_{11} & \bar{\Delta}_{12} \\ \bar{\Delta}_{21} & \bar{\Delta}_{22} \end{bmatrix} = T\Delta T^{-1} \tag{23}$$

and

$$\bar{C}_1 = \left(CF\right)^{\perp} C\tilde{T}_1 \tag{24}$$

It turns out that the triple (Δ, C, F) is strongly observable if, and only if, **the pair $(\bar{\Delta}_{11}, \bar{C}_1)$ is observable** (Molinari, 1976), (Hautus, 1983). In light of the Assumption A1, this property, that can be also understood in terms of a simplified algebraic test to check the strong detectability of a matrix triple, opens the way to design stable observers for the state of the transformed dynamics (22). The peculiarity of the transformed system (22) is that $\bar{q}_2$ constitutes a part of the transformed output vector $\bar{y}$. Hence, **the observation of the state of systems (22) can be accomplished by estimating $\bar{q}_1$ only**, whose dynamics is not affected by the unknown input vector. The observability of the $(\bar{\Delta}_{11}, \bar{C}_1)$ permits the implementation of the following Luenberger observer for the $\bar{q}_1$ subsystem of (22):

$$\dot{\hat{\bar{q}}}_1 = \bar{\Delta}_{11}\hat{\bar{q}}_1 + \bar{\Delta}_{12}\bar{y}_2 + L(\bar{y}_1 - \bar{C}_1\hat{\bar{q}}_1) \tag{25}$$

which gives rise to the error dynamics

$$\dot{e}_1 = (\bar{\Delta}_{11} - L\bar{C}_1)e_1, \qquad e_1 = \hat{\bar{q}}_1 - \bar{q}_1 \tag{26}$$

whose eigenvalues can be arbitrarily located by a proper selection of the matrix L. Therefore, with properly chosen L we have that

$$\hat{\bar{q}}_1 \to \bar{q}_1 \quad as \quad t \to \infty \tag{27}$$

which implies that the overall system state can be reconstructed by the following relationships

$$\hat{q} = T^{-1} \begin{bmatrix} \hat{\bar{q}}_1 \\ \bar{y}_2 \end{bmatrix} \tag{28}$$

Note that the convergence of $\hat{\bar{q}}_1$ to $\bar{q}_1$ is exponential and can be made as fast as desired.

5. RECONSTRUCTION OF THE UNKNOWN INPUT

A disturbance estimator can be designed which gives an exponentially converging estimate of the unknown input. Consider the following estimator dynamics

$$\dot{\hat{\bar{q}}}_2 = \bar{\Delta}_{21}\hat{\bar{q}}_1 + \bar{\Delta}_{22}\bar{y}_2 + v(t) \tag{29}$$

with the estimator injection input $v(t)$ yet to be specified. The next assumption is fulfilled:
A3. It can be found a constant w_d such that: $|\dot{w}(t)| \leq w_d$.

Define

$$\sigma(t) = \hat{\bar{q}}_2 - \bar{q}_2 \tag{30}$$

The time derivative of $\sigma(t)$ is

$$\dot{\sigma} = \dot{\hat{\bar{q}}}_2 - \dot{\bar{q}}_2 = \bar{\Delta}_{21}e_1(t) + v(t) - w(t) \tag{31}$$

Define the output injection $v(t)$ as follows

$$\begin{aligned} v(t) &= k_1 \left|\sigma\right|^{\frac{1}{2}} sign\,\sigma - k_2\sigma + \alpha(t) \\ \dot{\alpha}(t) &= -k_3\, sign\,\sigma - k_4\sigma \end{aligned} \tag{32}$$

Algorithm (32) implements the Second Order Sliding Mode Controller (2-SMC) called "Super-Twisting". The main reason why we used a (2-SMC) algorithm is that guarantees a continuous estimate of the output injection $v(t)$ without the aid of filters. Considering (31) into (32) yields:

$$\dot{\sigma} = \bar{\Delta}_{21}e_1(t) - w(t) - k_1 \left|\sigma\right|^{\frac{1}{2}} sign\,\sigma - k_2\sigma + \alpha(t) \tag{33}$$

To simplify the notation define

$$\Gamma(t) = \bar{\Delta}_{21}e_1(t) + \alpha(t) - w(t) \tag{34}$$

The derivative of $\Gamma(t)$ is

$$\dot{\Gamma} = \bar{\Delta}_{21}\dot{e}_1 + \dot{\alpha} - \dot{w} = \psi - k_3\, sign\,\sigma - k_4\sigma \tag{35}$$

where

$$\psi = \bar{\Delta}_{21}\dot{e}_1 - \dot{w} = \bar{\Delta}_{21}(\bar{\Delta}_{11} - LC_1)e_1 - \dot{w} \quad (36)$$

The error dynamics in $\sigma - \Gamma$ coordinates is:

$$\begin{aligned} \dot{\sigma} &= \Gamma - k_1 |\sigma|^{\frac{1}{2}} \, sign\,\sigma - k_2\sigma \\ \dot{\Gamma} &= \psi - k_3 \, sign\,\sigma - k_4\sigma \end{aligned} \quad (37)$$

Let the tuning parameters be chosen according to the next inequalities

$$k_1, k_3 > 0;\ k_2, k_4 \geq 0;\ min\left\{\frac{k_1}{2}, \frac{k_1 k_3}{1+k_1}, k_3\right\} > M \quad (38)$$

where M is any constant such that

$$M \geq w_d + \rho^2, \quad \rho \neq 0. \quad (39)$$

Considering (37)-(39), the condition $e_1(t) \to 0$, derived from (27), guarantees that the next condition (40) holds starting from a finite time instant $\bar{T}$.

$$|\psi| \leq M, \quad t \geq \bar{T}. \quad (40)$$

The stability of (37) can be demonstrated by means of the Lyapunov function (Moreno, 2008)

$$V(\sigma, \Gamma) = 2k_3|\sigma| + k_4\sigma^2 + \frac{1}{2}\Gamma^2 + \frac{1}{2}s^2(\sigma, \Gamma) \quad (41)$$

where

$$s(\sigma, \Gamma) = \Gamma - k_1|\sigma|^{\frac{1}{2}} \, sign\,\sigma \quad (42)$$

Differentiating the Lyapunov function (41) along the trajectory of the system (37) gives

$$\dot{V} \leq -|\sigma|^{-\frac{1}{2}} W(\sigma, s) \quad (43)$$

where

$$\begin{aligned} W(\sigma, s) &= [k_2 s^2|\sigma|^{\frac{1}{2}} + (\frac{k_1}{2} - M)s^2 + k_1 k_2] + \\ &+ (k_2 k_3 - M k_2)|\sigma|^{\frac{3}{2}} + (k_1 k_3 - M(1-k_1))|\sigma| + \\ &+ k_2 k_4 |\sigma|^{\frac{5}{2}} \geq \gamma V(\sigma, \Gamma) \end{aligned} \quad (44)$$

for some $\gamma > 0$ and for all $\sigma, \Gamma, s \in \mathbb{R}$. By taking advantage of (27) it can be easily shown (see (Moreno, 2008)) that $|v(t) - w(t)| \to 0$ as $t \to \infty$. Then, under the conditions (38), the estimator (29), (37) allows one to reconstruct the unknown input $w(t)$ acting on the original system (13).

6. NUMERICAL EXAMPLE

Consider the perturbed heat equation.

$$z_t = 0.5 z_{xx} + f(x)w(t) \quad (45)$$

with homogeneous Dirichlet BCs (4) where $f(x) = 2 + 6\sin(4\pi x)$ and $w(t) = 2(10 + sin(4t))$. The initial conditions are set to $z_0(x) = sin(\pi x)$ and the location of the two sensors are: $x_c^1 = 0.1$ and $x_c^2 = 0.7$. The corresponding eigenvalues and eigenfunctions are

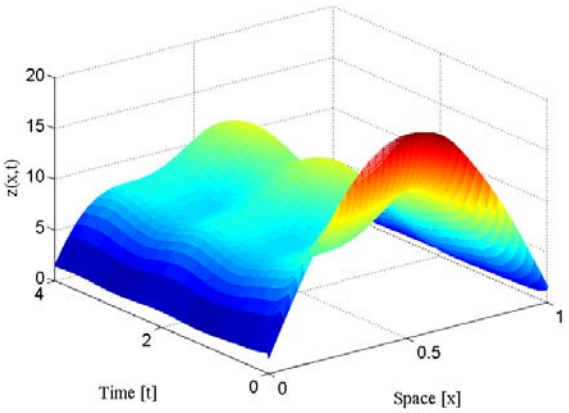


Fig. 1. TEST 1: actual state $z(x, t)$.

$$\lambda_n = -0.5\frac{n^2}{\pi^2}, \quad \phi_n(x) = sin(n\pi x), \quad n = 1, ..., \infty. \quad (46)$$

To generate the measurements and the "true" states accurately, the equation (45) has been simulated using $N = 50$ modes. Figure 1 shows the actual state evolution.

In TEST 1 the UIO and the disturbance estimator are implemented with $N = 5$ modes. The next matrices are obtained for the original system (16):

$$\Delta = \begin{bmatrix} -4.934 & 0 & 0 & 0 & 0 \\ 0 & -19.739 & 0 & 0 & 0 \\ 0 & 0 & -44.413 & 0 & 0 \\ 0 & 0 & 0 & -78.956 & 0 \\ 0 & 0 & 0 & 0 & -123.370 \end{bmatrix}$$

$$F = \begin{bmatrix} 2.546 \\ 0 \\ 0.888 \\ 6 \\ 0.509 \end{bmatrix}, \ C = \begin{bmatrix} 0.338 & 0.637 & 0.860 & 0.982 & 0.987 \\ 0.827 & -0.929 & 0.218 & 0.684 & -0.987 \end{bmatrix} \quad (47)$$

The transformation matrices (14), (15) are

$$\begin{aligned} U &= \begin{bmatrix} -0.737 & 1 \\ 0.081 & 0.059 \end{bmatrix} \\ T &= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -0.333 & 0 & 1 & 0 & 0 \\ -2.356 & 0 & 0 & 1 & 0 \\ -0.200 & 0 & 0 & 0 & 1 \\ 0.076 & -0.003 & 0.082 & 0.120 & 0.021 \end{bmatrix} \end{aligned} \quad (48)$$

and the transformed system (22) is characterized by the following matrices

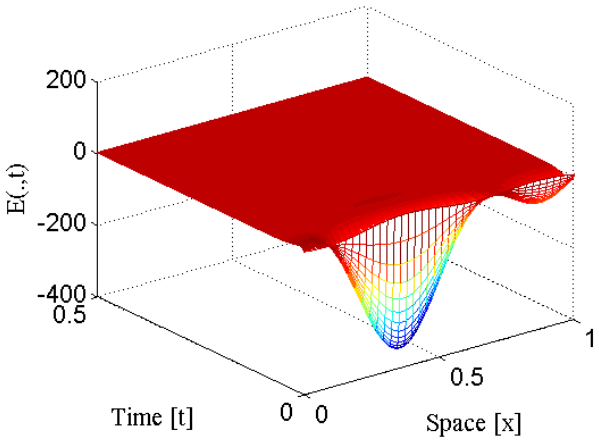


Fig. 2. Observation error $E(x,t)$ for N=5 (TEST 1).

$$\bar{\Delta}_{11} = \begin{bmatrix} -19.739 & 0 & 0 & 0 \\ -0.132 & -41.638 & 4.039 & 0.703 \\ -1.750 & 36.772 & -25.424 & 9.318 \\ -0.237 & 4.994 & 7.270 & -122.104 \end{bmatrix}$$

$$\bar{\Delta}_{12} = \begin{bmatrix} 0 \\ -33.510 \\ -444.132 \\ -60.318 \end{bmatrix} \tag{49}$$

$$\bar{\Delta}_{21} = [-0.168 \ 1.498 \ -1.982 \ -1.276]$$

$$\bar{\Delta}_{22} = -62.507$$

$$C_1 = [-1.400 \ -0.417 \ -0.040 \ -1.716]$$

It can be checked that $\bar{\Delta}_{11}$ and C_1 is an observable pair which confirms that the conditions A1 and A2 hold, system (13) and (22) is strongly observable and hence we can implement the Luenberger observer (25) with the following L matrix

$$L = \begin{bmatrix} 1834 \\ -1955 \\ -10639 \\ -836 \end{bmatrix} \tag{50}$$

which assign all eigenvalues of the error matrix $[\bar{\Delta}_{11} - LC_1]$ the same value: -80. The variable $\hat{\tilde{\mathbf{q}}}_1$ generated by the Luenberger observer is used to implement the disturbance estimator (29) and the estimator control signal $v(t)$ is obtained by setting the super-twisting gains (37) as follows: $k_1 = 44, k_2 = 0, k_3 = 10, k_4 = 0$.
Fig.2 shows the spatio-temporal profile of the state estimation error $E(x,t) = z(x,t) - \hat{z}(x,t)$ where

$$\hat{z}(x,t) = \sum_{n=1}^{5} \hat{q}_n(t)\phi_n(x)$$

while Fig.3 depicts the corresponding L_2 error norm. Fig.4 show the actual and estimated profile of the disturbance $w(t)$ which confirms the good performance of the suggested estimator. In TEST 2 the observer is built considering $N = 20$ modes. The actual and estimated profiles of the disturbance are shown in Fig.5. The final test (TEST 3) considers homogeneous Neumann BCs (3) instead of (4) and an observer implemented with $N = 5$ modes. Fig.6, Fig.7 and Fig.8 show the corresponding results.

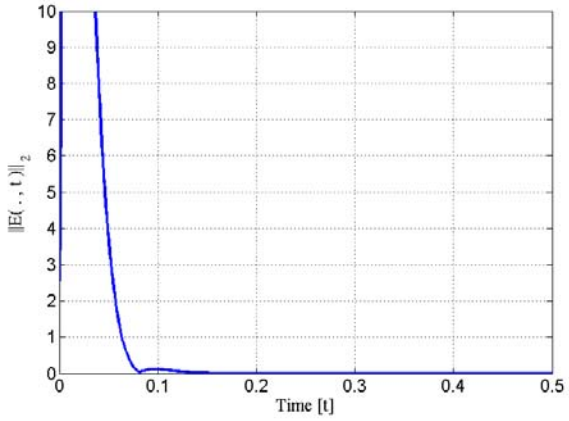


Fig. 3. The L_2 norm of the observation error $||E(\cdot,t)||_2$ for N=5 (TEST 1).

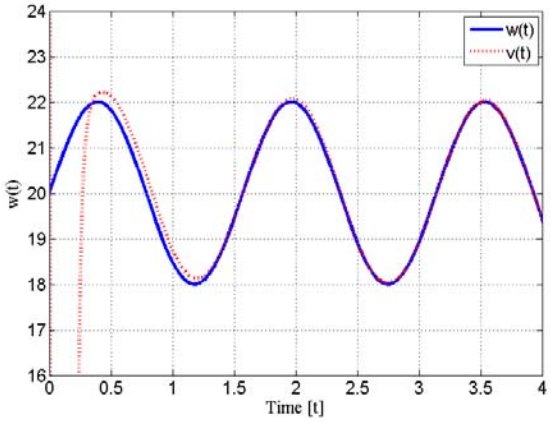


Fig. 4. Unknown disturbance reconstruction for N=5 (TEST 1).

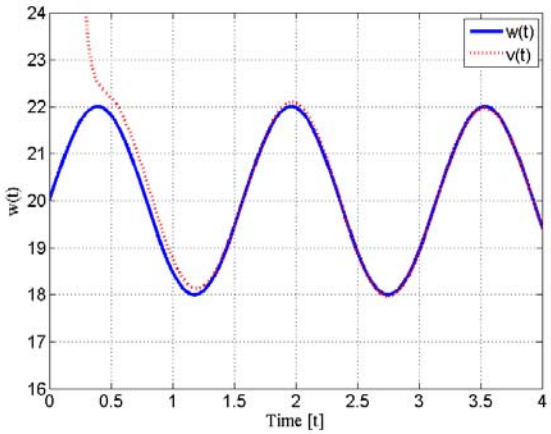


Fig. 5. Unknown disturbance reconstruction for N=20 (TEST 2).

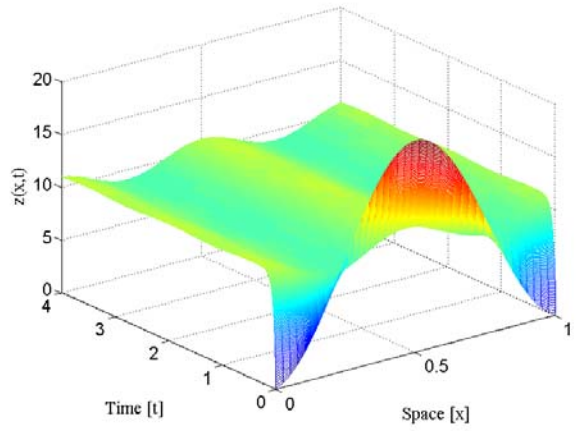


Fig. 6. Actual state in the TEST 3 (Neumann BCs).

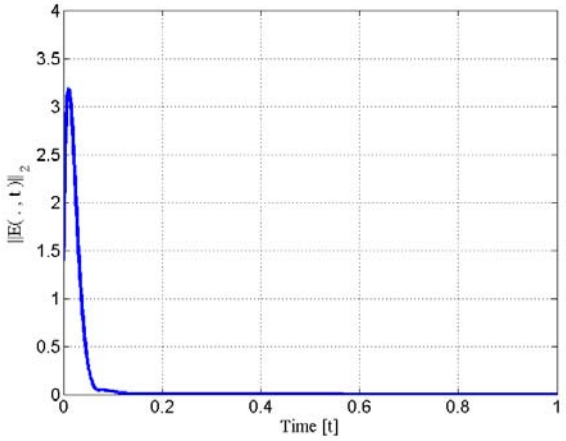


Fig. 7. The L_2 norm of the observation error $||E(\cdot,t)||_2$ in TEST 3.

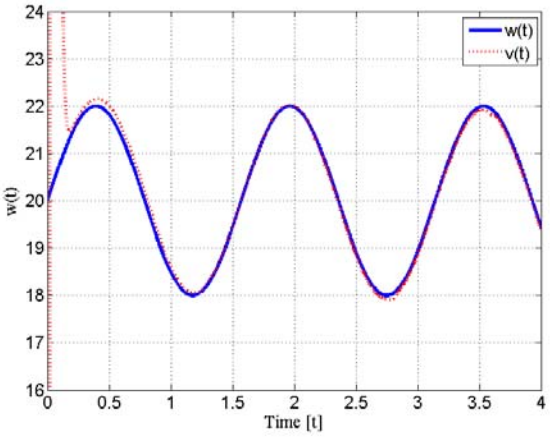


Fig. 8. Unknown disturbance reconstruction in TEST 3.

7. CONCLUSIONS AND FUTURE WORK

In this work an approach to approximate state observation and unknown input reconstruction for a class of distributed parameter system is proposed. By means of a simulation example, it is checked that both the modal and finite-difference approximations of the diffusion equation solution fulfills the property of strong observability when two point-wise measurements are located in the solution spatial domain. Next investigations will be devoted to better analyze the properties of the suggested scheme and to extend the present analysis to more general classes of PDEs.

REFERENCES

F.J. Bejarano, L. Fridman and A. Poznyak, "Exact state estimation for linear systems with unknown inputs based on hierarchical super-twisting algorithm" Int. J. Robust Nonlinear Control, 17(18), pp. 1734-1753, 2007.

J.Chen, R. Patton and H. Y. Zhang, "Design of unknown input observers and robust fault detection filters", Int. J. Control, 63, pp. 85-105, 1996.

R. F. Curtain, H. J. Zwart, "An Introduction to Infinite-Dimensional Linear Systems Theory", Texts in Applied Mathematics, Vol. 21, Springer-Verlag, Berlin, 1995.

M.A. Demetriou and I.G. Rosen, "Unknown Input Observers for a class of distributed parameter systems" Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on, Dec 2005.

C. Edwards, S. K. Spurgeon, and R. J. Patton ,"Sliding mode observers for fault detection and isolation", *Automatica*, vol. 36, pp. 541–553, 2000.

M. L. J. Hautus, "Strong detectability and observers" Linear Algebra and its Applications, 50, pp. 353368, 1983.

G. T. R. Imanuilov, O. Leugering and Z. Bing-Yu, "Control Theory of Partial Differential Equations", C. . H. T. . F. Group, Ed. 1st, 2005.

W. Kratz, "Characterization of Strong Observability and Construction of an Observer", Linear Algebra and its Appl. pages 31-40, 1995.

M. Krstic, A. Smyshlyaev, "Backstepping observers for a class of parabolic pdes", Systems & Control letters, vol. 54, p. 613 U 625, 2005.

B.P. Molinari, "A strong contollability and observability in linear multivariable control" IEEE Transaction on Automatic Control, 21(5), pp. 761764, 1976.

R. P. Mondaini and P. M. Pardalos, "Mathematical Modelling of Biosystems", Springer, Ed. 1st, 2008.

J. A. Moreno, M. Osorio, "A Lyapunov approach to second-order sliding mode controllers and observers", Proceedings of the 47th IEEE Conference on Decision and Control Cancun, Mexico, Dec. 9-11, 2008.

A. Pazy, "Semigroups of linear operators and applications to Partial Differential Equations" New York: (Springer-Verlag), 1983.

W. E. Schiesser and G. W. Griffiths, "A Compendium of Partial Differential Equation Models", C. U. Press, Ed. 1st, 2009.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

An efficient algorithm for fault tolerant sensor network design

F. Rouissi. G. Hoblos, N. Langlois

IRSEEM EA 4353, Institut de Recherche en Systèmes Electroniques Embarquées
ESIGELEC, Ecole Supérieure d'Ingénieurs Généralistes
Avenue Galilée, BP 10024, 76801 Saint Etienne du Rouvray, France
((rouissi, hoblos, langlois)@esigelec.fr).

Abstract: This paper deals with the problem of fault tolerant sensor network design. New criterion based algorithm is proposed to provide a fault tolerant architecture. The proposed criterion is based on the number of admissible losses situations which guarantees some robustness for operating safety. An oriented graph is used to compute this criterion with respect to some given properties. Some relations with strong and weak redundancy degrees are given for computing them together in order to propose multi-criteria based algorithm design. The proposed algorithm will be illustrated with an academic example.

Keywords: Fault tolerance, sensor network design, observability

1. INTRODUCTION

Interest for more reliable and safety systems is growing as control systems become increasingly complex and encounter various unexpected component failures. Indeed, fault-tolerant control takes into account, at the early design stage, unexpected system failures. For this reason, the ability of a system (sensors, actuators, components) to tolerate a fault must be taken from designing. Such a problem is called sensor or actuator network design or selection.

The sensor network design problem for steady state has been addressed in the past using different approaches. Ali and Narasimhan proposed to maximize reliability, which is based on sensor failure probability, observability of variables as well as redundancy. They introduced the concept of reliability of estimation of a variable (Ali and Narasimhan, 1995), where the concept was lately extended to redundant networks (Ali and Narasimhan, 1993). Their algorithm uses graph theory to build networks with a specific number of sensors and maximum system reliability.
A minimum cost model for the design of reliable sensor networks was presented in (Bagajewicz and Sànchez, 2000) where the connection between the models based on reliability goals and the minimum cost model subject to reliability constraints were explored.

Optimization formulation incorporating various constraints on the sensor network design problem as well utilizing quantitative information such as fault occurrence and sensor failure probabilities, and constraints for sensor location were presented in (Bhushan and Rengaswamy, 2000a) and (Bhushan and Rengaswamy, 2000b). In (Bhushan, Narasimhan and Rengaswamy, 2008) a formulation for incorporating robustness to the uncertain fault occurrence and sensor failure probability data is presented.

The problem of sensor location to ensure observability of dynamic systems has been addressed for linear time invariant systems, and extended to bilinear (Ragot, Maquin and Bloch, 1992). Optimal sensor location for parameters estimation has been considered (Firdaus and Udwadia, 1994).

Turbatte *et al* (Turbatte et al., 1991) have proposed a concept of system reliability that gives the probability for all variables to be observable when sensors are likely to fail, and have introduced redundancy degrees as robustness measures for the observability property. Further contributions have been given in (Luong et al., 1994), where two notions have been defined, namely the "principal redundancy of degree k" and the "weak redundancy degree".

In (Chamseddine, Noura and Raharijaona, 2007), the work of Staroswiecki (Staroswiecki, Hoblos and Aïtouche, 2004) is extended to non linear large scale complex systems. The complex system is decomposed into interconnected subsystems. The decomposition enables the use of reduced order observers for subsystems. This simplifies observers design and reduces the calculation requirement. For each subsystem, the minimum set of sensors allowing its observability is determined while taking into consideration its interconnection with other subsystems. A method to design a minimal cost sensor network ensuring the observability of complex systems consisted of interconnected subsystems is proposed (Chamsedine et al., 2008).

Recently, Sircoulomb *et al.* proposed in (Sircoulomb et al., 2007) a sensor redundancy elaboration method in systems with physical and costs constraints. The method allows making redundant a sensor relatively to estimation quality degradation.

This paper is organized as follows. Section 2 introduces the definition of the new criterion proposed and presents the observability graph used for iterative computation. Some

relations with strong and weak redundancy degrees are also introduced. In section 3, two proposed sensor network design algorithms are described for simple and multi criterion respectively. Finally, an illustrative example is provided in section 4.

2. FAULT TOLERANCE ANALYSIS AND NEW CRITERION FORMULATION

Consider the discrete deterministic *LTI* system:

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) \\ z(k) = Hx(k) \end{cases} \quad (0)$$

where $x \in R^n$ is the state vector, $u \in R^m$ is the control input, $y \in R^p$ is the measurement vector, and $z \in R^q$ is the vector which is to be estimated ($q < n$). *A, B, C* and *H* are matrices of suitable dimensions. Obviously, the classical observability problem corresponds to the particular setting $H = I$.

If *I* is the whole sensor set, the sensor subsets of I can be organized in a subset lattice, where two levels are given in figure 1 (Staroswiecki, Hoblos and Aïtouche, December, 1999).

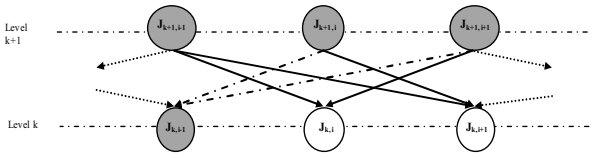


Fig. 1. Lattice of sensors subsets

- a node is a subset of *I*,
- an edge is an oriented link between a node at level *k* (subsets containing *k* sensors) and a node at level *k+1*.
- A grey node is a node which keeps the functional state *z* observable. The observability test of each subset *J*, included in *I*, for estimating *z*, is:

$$rank\begin{bmatrix} H \\ OBS(J) \end{bmatrix} = rank[OBS(J)] \quad (0)$$

where *OBS(J)* is the observability matrix constituted with the sensor subset *J*.

2.1 Assumption

A sensor loss situation is said to be admissible if it leads to an observable sensor subset. This supposes that the sensor subset before sensor losing is grey.

2.2 Criterion formulation

Based on assumption 1, for each sensor subset *J* in the graph, an indicator $I_{i,J}$ can be associated to any belonged sensor *i*. This indicator gives the number of admissible losses situations concerning this sensor. Obviously, this indicator depends on the level and the sensor subset where the sensor is localized. Indeed, for a level, a sensor can have indictors with different values w.r.t. to sensor subset in which the sensor is belonged.

The robustness of the observability property can be measured using the proposed sensors indicators. The more the sensor indicator is high the more the observability robustness is better.

2.3 Example

Let *I* be a sensor set of *3* sensors *a*, *b* and *c*. The corresponding graph is given in figure 2.

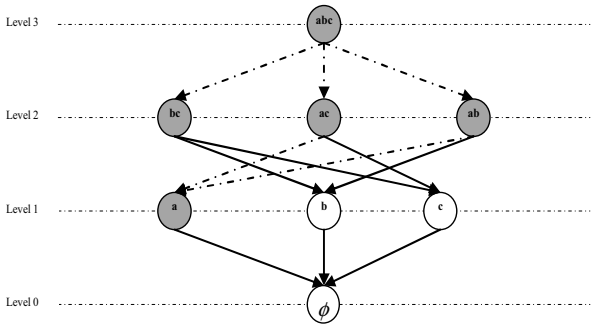


Fig. 2. Graph of 3 sensors set

There are only two sensors subsets {b} and {c} where the system is not observable. Then, all indicators are s. t.:
- In level 1: $I_{a/\{a\}}=0$, $I_{b/\{b\}}=0$ and $I_{c/\{c\}}=0$
- In level 2: $I_{b/\{a,b\}}=1$, $I_{a/\{a,b\}}=0$, $I_{c/\{a,c\}}=1$, $I_{a/\{a,c\}}=0$, $I_{b/\{b,c\}}= I_{c/\{b,c\}}=0$
- In level 3: $I_{a/\{a,b,c\}}=1$, $I_{b/\{a,b,c\}} = I_{c/\{a,b,c\}}=2$

2.4 Properties

Property 1:

The sensors indicators associated with white or terminal nodes are equal to zero. A terminal node is a node where all successors are white.

Property 2:

A node is grey if there is at least a sensor indicator different of zero.

2.5 Relations with strong and weak redundancy degrees

First of all, definitions of strong and weak redundancy degrees are reminded:

Weak Redundancy Degree of J: Weak redundancy degree of *J* is defined as the maximal number of sensors of *J* which may be lost while continuing to estimate z.

Strong Redundancy Degree of J: Strong redundancy degree of *J* is defined as the maximal number of indifferentiate sensors of *J* that may be lost while continuing to estimate z.

Property 3:

For a node *J*, if $\min_j(I_{j,J}) = l,\ \ l \in \{0,1\}$, then the strong redundancy degree of *J* is equal to *l*. Obviously, if $\min_j(I_{j,J}) \geq 2$, then the strong redundancy degree of *J* is ≥ 1.

Property 4:

For a node *J*, if $\max_j(I_{j,J}) = l,\ \ l \in \{0,1,2\}$, then the weak redundancy degree of *J* is equal to *l*.

3. SENSOR NETWORK DESIGN ALGORITHMS

In this paragraph, the objective consists in designing a sensor network that ensures system's observability by guarantying a certain desired threshold l_d of sensors indicators defined previously. The problem can be reduced to choose a subset *J* s.t.:

$$\min_j(I_{j,J}) = l_d \qquad (3)$$

This desired threshold guarantees that first sensors losses does not lead to the loss of observability property.

The proposed algorithm is based on the observability graph by:

- testing system's observability for a number of sensor sets where observability test is simple for linear systems and for relatively small scale nonlinear systems,
- and computing of the new proposed criterion.

The next three properties will be necessary to simplify computing.

Property 5:

If a node *J* is grey, for all predecessor nodes of *J*, all sensors indicators, except those belonged to *J*, are incremented of 1.

Property 6:

If a node J is white, for all predecessor nodes of *J*, all sensors indicators keep their values.

Then, the algorithm can be done as follows in three main steps:

3.1 Algorithm 1

Step *INI*

- All sensor indicators of all nodes are equal to zero. Particularly, sensor indicators in level 1 are always equel to zero (*property 1*).

Step level (*k:* $1 \rightarrow p$)

- For each node, if at least one sensor indicator is different of zero then the node is grey, else observability test is necessary for marking (*property 2*).
- If a node is grey all sensors indicators for all predecessor nodes are incremented (*property 5*).
- If a node is white all sensors indicators for all predecessor nodes keep their values (*property 6*).

Stop condition

- Desired sensors indicators l_d reached
- or top level reached

This algorithm can be extended from simple criterion concerning sensors indicators to couple criteria by combining with weak and strong redundancy degrees. This leads to the next algorithm.

3.2 Algorithm 2

Step *INI*

- All sensor indicators of all nodes are equal to zero (*property 1*)
- Strong and weak redundancy degree are equal to zero (*property 3*)

Step level (*k:* $1 \rightarrow p$)

- For each node, if at least one sensor indicator is different of zero then the node is grey, else observability test is necessary for marking. (*property 2*)
- If a node is grey all sensors indicators for all predecessor nodes are incremented (*property 5*).
- If a node is white all sensors indicators for all predecessor nodes keep their values (*property 6*).
- Compute strong and weak redundancy degrees of each node by applying properties 3 and 4. Other properties can be necessary to compute these redundancy degrees (Staroswiecki, Hoblos and Aïtouche, 2004).

Stop condition

- Desired criteria, concerning sensors indicators and strong and/or weak redundancy degrees, reached
- or top level reached

4. APPLICATION EXAMPLE

To illustrate the sensor network design algorithm, a *LTI* system with 7 states and 4 possible sensors is used (Staroswiecki, Hoblos and Aïtouche, December, 1999), (Hoblos, Staroswiecki and Aïtouche, 2000). The Jordan I given by:

$$A=\begin{pmatrix}1&0&0&0&0&0&0\\0&0.5&0&0&0&0&0\\0&0&3&0&0&0&0\\0&0&0&4&0&0&0\\0&0&0&0&2&0&0\\0&0&0&0&0&1.5&0\\0&0&0&0&0&0&2.5\end{pmatrix},\ C=\begin{pmatrix}1&1&1&1&0&0&0\\0&0&0&0&1&1&1\\1&1&0&0&0&0&1\\0&1&1&1&1&1&0\end{pmatrix}$$

$$H=\begin{pmatrix}1&0&0&0&0&0&0\\0&1&0&0&0&0&0\\0&0&1&0&0&0&0\\0&0&0&1&0&0&0\end{pmatrix}$$

From the definition of matrix H, only a part of the states, namely $z=[x_1, x_2, x_3, x_4]$ has to be estimated.

The objective is to find a sensor subset J s.t.: $\min_j(I_{j,J})=1$.

Let us give in details the execution of the proposed algorithm.

Step *INI*
All sensor indicators of all nodes are equal to zero (Table 1).

Table 1

Level 1	Level 2	Level 3	Level 4
	$cd\begin{vmatrix}I_c=0\\I_d=0\end{vmatrix}$		
$a\vert I_a=0$	$bd\begin{vmatrix}I_b=0\\I_d=0\end{vmatrix}$	$bcd\begin{vmatrix}I_b=0\\I_c=0\\I_d=0\end{vmatrix}$	
$b\vert I_b=0$	$bc\begin{vmatrix}I_b=0\\I_c=0\end{vmatrix}$	$acd\begin{vmatrix}I_a=0\\I_c=0\\I_d=0\end{vmatrix}$	$abcd\begin{vmatrix}I_a=0\\I_b=0\\I_c=0\\I_d=0\end{vmatrix}$
$c\vert I_c=0$	$ad\begin{vmatrix}I_a=0\\I_d=0\end{vmatrix}$	$abd\begin{vmatrix}I_a=0\\I_b=0\\I_d=0\end{vmatrix}$	
$d\vert I_d=0$	$ab\begin{vmatrix}I_a=0\\I_b=0\end{vmatrix}$	$abc\begin{vmatrix}I_a=0\\I_b=0\\I_c=0\end{vmatrix}$	
	$ac\begin{vmatrix}I_a=0\\I_c=0\end{vmatrix}$		

Step level (1)

All sensors indicators (Table 1) are equal to zero so observability test is necessary for coloring (figure 3), (property 2). Then, only node {a} is grey.
Predecessors of {a} are {a,d}, {a,b}, {a,c}, {a,c,d}, {a,b,d}, {a,b,c} and {a,b,c,d}, then, $I_{d/\{a,d\}}$, $I_{b/\{a,b\}}$, $I_{c/\{a,c\}}$, $I_{c/\{a,c,d\}}$, $I_{d/\{a,c,d\}}$, $I_{b/\{a,b,d\}}$, $I_{d/\{a,b,d\}}$, $I_{b/\{a,b,c}$, $I_{c/\{a,b,c\}\}}$ and $I_{a/\{a,b,c,d\}}$ are incremented.

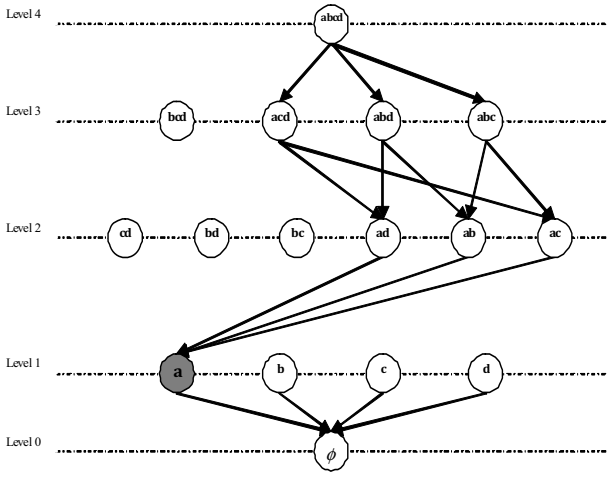


Fig.3. node {a}

Table 2 contains the computation results for level 1.

Table 2

Level 1	Level 2	Level 3	Level 4
	$cd\begin{vmatrix}I_c=0\\I_d=0\end{vmatrix}$		
$a\vert I_a=0$	$bd\begin{vmatrix}I_b=0\\I_d=0\end{vmatrix}$	$bcd\begin{vmatrix}I_b=0\\I_c=0\\I_d=0\end{vmatrix}$	
$b\vert I_b=0$	$bc\begin{vmatrix}I_b=0\\I_c=0\end{vmatrix}$	$acd\begin{vmatrix}I_a=0\\I_c=1\\I_d=1\end{vmatrix}$	$abcd\begin{vmatrix}I_a=0\\I_b=1\\I_c=1\\I_d=1\end{vmatrix}$
$c\vert I_c=0$	$ad\begin{vmatrix}I_a=0\\I_d=1\end{vmatrix}$	$abd\begin{vmatrix}I_a=0\\I_b=1\\I_d=1\end{vmatrix}$	
$d\vert I_d=0$	$ab\begin{vmatrix}I_a=0\\I_b=1\end{vmatrix}$	$abc\begin{vmatrix}I_a=0\\I_b=1\\I_c=1\end{vmatrix}$	
	$ac\begin{vmatrix}I_a=0\\I_c=1\end{vmatrix}$		

Step level (2)
{a,d}, {a,b} and {a,c} are grey (*property 2*).
For each node where all sensor indicators are equal to zero observability test is necessary for coloring. Then, only {c,d} is grey. For {c,d}, predecessors are {b,c,d}, {a,c,d} and {a,b,c,d}, (figure 4).

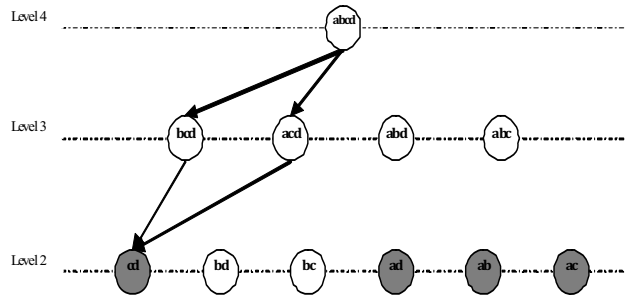


Fig.4. node {c,d}

And then all corresponding sensors indicators are incremented.

For {a,d}, predecessors are {a,c,d}, {a,b,d} and {a,b,c,d},

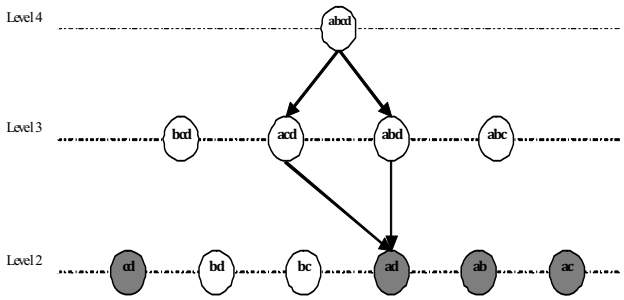


Fig.5. node {a,d}

All corresponding sensors indicators are incremented.

For {a,b}, predecessors are {a,b,d}, {a,b,c} and {a,b,c,d},

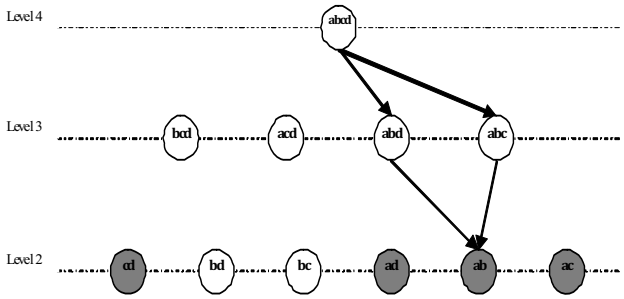


Fig.6. node {a,b}

$I_{d/\{a,b,d\}}$, $I_{c/\{a,b,c\}}$, $I_{c/\{a,b,c,d\}}$ and $I_{d/\{a,b,c,d\}}$ are incremented.

For {a,c}, predecessors are {a,c,d}, {a,b,c} and {a,b,c,d},

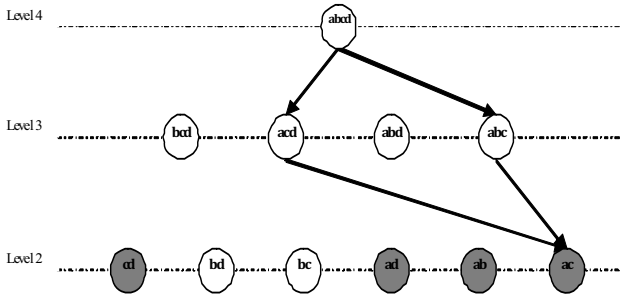


Fig.7. node {a,c}

$I_{d/\{a,c,d\}}$, $I_{b/\{a,b,c\}}$, $I_{b/\{a,b,c,d\}}$ and $I_{d/\{a,b,c,d\}}$ are incremented.

Table 3 present the sensors indicators once all nodes in level 3 were tested.

Table 3

Level 1	Level 2	Level 3	Level 4
	$cd\begin{vmatrix} I_c = 0 \\ I_d = 0 \end{vmatrix}$		
$a\vert I_a = 0$	$bd\begin{vmatrix} I_b = 0 \\ I_d = 0 \end{vmatrix}$	$bcd\begin{vmatrix} I_b = 2 \\ I_c = 0 \\ I_d = 0 \end{vmatrix}$	
$b\vert I_b = 0$	$bc\begin{vmatrix} I_b = 0 \\ I_c = 0 \end{vmatrix}$	$acd\begin{vmatrix} I_a = 1 \\ I_c = 2 \\ I_d = 2 \end{vmatrix}$	$abcd\begin{vmatrix} I_a = 1 \\ I_b = 4 \\ I_c = 3 \\ I_d = 3 \end{vmatrix}$
$c\vert I_c = 0$	$ad\begin{vmatrix} I_a = 0 \\ I_d = 1 \end{vmatrix}$	$abd\begin{vmatrix} I_a = 0 \\ I_b = 2 \\ I_d = 2 \end{vmatrix}$	
$d\vert I_d = 0$	$ab\begin{vmatrix} I_a = 0 \\ I_b = 1 \end{vmatrix}$	$abc\begin{vmatrix} I_a = 0 \\ I_b = 2 \\ I_c = 2 \end{vmatrix}$	
	$ac\begin{vmatrix} I_a = 0 \\ I_c = 1 \end{vmatrix}$		

Step level (3)

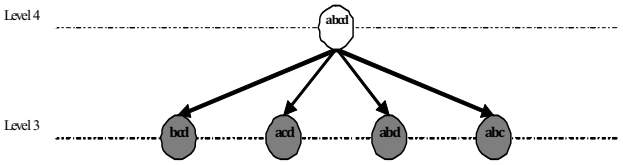


Fig.8. Step level (3)

In level 3, all nodes are grey (*property 2*). $I_{a/\{a,b,c,d\}}$, $I_{b/\{a,b,c,d\}}$, $I_{c/\{a,b,c,d\}}$ and $I_{d/\{a,b,c,d\}}$ are incremented.

Final resulting indicators are presented in table 4.

Table 4

Level 1	Level 2	Level 3	Level 4
	$cd\begin{vmatrix} I_c = 0 \\ I_d = 0 \end{vmatrix}$		
$a\vert I_a = 0$	$bd\begin{vmatrix} I_b = 0 \\ I_d = 0 \end{vmatrix}$	$bcd\begin{vmatrix} I_b = 2 \\ I_c = 0 \\ I_d = 0 \end{vmatrix}$	
$b\vert I_b = 0$	$bc\begin{vmatrix} I_b = 0 \\ I_c = 0 \end{vmatrix}$	$acd\begin{vmatrix} I_a = 1 \\ I_c = 2 \\ I_d = 2 \end{vmatrix}$	$abcd\begin{vmatrix} I_a = 2 \\ I_b = 5 \\ I_c = 4 \\ I_d = 4 \end{vmatrix}$
$c\vert I_c = 0$	$ad\begin{vmatrix} I_a = 0 \\ I_d = 1 \end{vmatrix}$	$abd\begin{vmatrix} I_a = 0 \\ I_b = 2 \\ I_d = 2 \end{vmatrix}$	
$d\vert I_d = 0$	$ab\begin{vmatrix} I_a = 0 \\ I_b = 1 \end{vmatrix}$	$abc\begin{vmatrix} I_a = 0 \\ I_b = 2 \\ I_c = 2 \end{vmatrix}$	
	$ac\begin{vmatrix} I_a = 0 \\ I_c = 1 \end{vmatrix}$		

The algorithm execution can be stopped at level 3 where the desired threshold $l_d = 1$ is reached.

4. CONCLUSION

A new criterion based algorithm for fault tolerant sensor network design has been developed in this paper. The objective is to guarantee the observability properties and to optimize the sensor network design. The sensor indicator is

used as a selective criterion to choose the appropriate subset which gives an acceptable estimation situation of the system. It was showed that this criterion can be combined with strong and weak redundancy degrees for multi-criteria sensor network design.

In the field of sustainable energy, these results will be explored for designing of a benchmark for the monitoring and supervision of systems with multi source of renewable energy.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Haute Normandie Region and the project CHAMP (Low-Carbon Hybrid Advanced Motive Power) within the framework of Interreg IVA FRANCE (CHANNEL–ENGLAND TERRITORIAL COOPERATION that have financially supported this study.

REFERENCES

Ali, Y. and Narasimhan, S. (1993) 'Sensor network design for maximizing reliability of linear processes', *American Institute of Chemical Engineers Journal*, vol. 39, p. 820.

Ali, Y. and Narasimhan, S. (1995) 'Redundant sensor network for linear processes', *American Institute of Chemical Engineers Journal*, vol. 41, p. 2237.

Bagajewicz, M. and Sànchez, M. (2000) 'Cost-optimal design of reliable sensor networks', *Computers and Chemical Engineering*, vol. 23, pp. 1757-1762.

Bhushan, M., Narasimhan, S. and Rengaswamy, R. (2008) 'Robust sensor network design for fault diagnosis', *Computer and Chemical Engineering*, vol. 32, pp. 1067-1084.

Bhushan, M. and Rengaswamy, R. (2000a) 'Design of sensor location based on various fault diagnostic observability and reliability criteria', *Computers & Chemical Engineering*, vol. 24, p. 735.

Bhushan, M. and Rengaswamy, R. (2000b) 'Design of sensor network based on the SDG of the process for efficient fault diagnosis', *Industrial & Engineering Chemistry Research*, vol. 41, pp. 1826-1839.

Chamseddine, A., Noura, H. and Raharijaona (2007) 'Optimal sensor network design for observability of complex systems', *American control conference ACC'07*, July, pp. 1705-1619.

Chamsedine, A., Noura, H., Ouladsine, M. and Raharijaona, T. (2008) 'Observability of complex systems: Minimal cost Sensor network design', *Proceedings of the 17th world Congress, The international federation of Automatic Control*, pp. 13287-13292.

Firdaus, E. and Udwadia, E. (1994) 'Methodology for optimum sensor location for parameter identification in dynamic systems', *Journal of Enginering Mechanics*, vol. 120, no. 2, pp. 368-390.

Hoblos, G., Staroswiecki, M. and Aïtouche, A. (2000) 'Fault Tolerant Sensor Network Design Using Redundancy Degrees', *4th IFAC Symposium on Intelligent Components and Instruments for Control Applications*, pp. 93-98.

Hoblos, G., Staroswiecki, M. and Aïtouche, A. (2000) 'Optimal design of fault tolerant sensor networks', *Proc of the IEEE International Conf. on control application CCA'2000*, pp. 467-472.

Luong, M., Maquin, D., Huynh, C.T. and Ragot, J. (1994) 'Observability, Redundancy, Reliability and Integrated design of Measurment System', *SICICA'94*.

Ragot, J., Maquin, D. and Bloch, G. (1992) 'Sensor positioning for processes described by linear linear processes.', *Diagn. Sûreté Fonct.*, vol. 2.

Sircoulomb, V., Hoblos, G., Chafouk, H. and Ragot, J. (2007) 'Analyse et synthèse de redondance de capteurs en vue d'améliorer l'estimation d'état d'un système', *4 ème Colloque Interdisciplinaire en instrumentation, C2I.*

Staroswiecki, M., Hoblos, G. and Aïtouche, A. (2004) 'Sensor network design for fault tolerant estimation.', *International Journal of Adaptive Control and Signal Processing*, vol. 18, pp. 55-72.

Staroswiecki, M., Hoblos, G. and Aïtouche, A. (December, 1999) 'Fault Tolerance Analysis of sensor Systems', *38th IEEE CDC'99 (conf. on Decision & Control)*, pp. 3581-3586.

Turbatte, H.C., Maquin, D., Cordier, C. and Huynh, C.T. (1991) 'Analytical Redundancy and reliability of measurement systems', *IFAC Safeprocess'91*, vol. 1, pp. 49-54.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Multi-Scale PCA based fault diagnosis for rotating electrical machines

Francesco Ferracuti * **Andrea Giantomassi** *
Gianluca Ippoliti * **Sauro Longhi** *

* *Dipartimento di Ingegneria Informatica Gestionale e dell'Automazione*
Universitá Politecnica delle Marche, via Brecce Bianche, 60131 Ancona
Email: frank_f@hotmail.it, {a.giantomassi, gianluca.ippoliti, sauro.longhi}@univpm.it.

Abstract: In high competition for reducing costs, the maintenance of the production plants has a key role. Machine fault detection and diagnosis can decrease the costs of maintenance by minimizing the loss of production due to machine breakdown. In this paper, Multi-Scale Principal Component Analysis (MSPCA) is used for fault detection and diagnosis. MSPCA simultaneously extracts both, cross correlation across the sensors (PCA approach) and auto-correlation within a sensor (Wavelet approach). The advantage of MSPCA is also demonstrated on a laboratory-scale experimental system, by means of vibration measurements.

Keywords: Fault detection, Fault diagnosis, Vibration measurements, Data acquisition, Electrical machines, Signal analisys.

1. INTRODUCTION

In the field of operation efficiency, the monitoring activity of rotating electrical machines for fault detection and diagnosis is depth investigated (Bagheri et al., 2007; Taniguchi et al., 1999; Benloucif and Balaska, 2006; Verucchi and Acosta, 2008; Tavner, 2008). Vibration analysis is widely accepted as a tool to detect faults of a machine as it is non-destructive, reliable and it permits continuous monitoring without stopping the machine (Liu et al., 2009; Wan et al., 2002; Bellini et al., 2008; Wu et al., 2009; Gani and Salami, 2002; Ciandrini et al., 2010). In particular analyzing the vibration power spectrum it is possible to detect different fault that can be arise in rotating machines. Most common defects in these machines are unbalance and misalignment. Unbalance generates a radial component at the rotation frequency. Misalignment generates a radial component at double of rotating frequency and an axial component at rotation frequency. Moreover components of the spectrum over the rotation frequency are due to bearings, events that occur many times per round, signal distortion, mechanical non linearities (i.e. backlash). Often, a fault arising in a rotating machine increases the vibration amplitude associated with the fault. For instance, if a fault occurs in gears, the vibration amplitude of a whole family of sidebands increases in a specific region of its frequency spectrum, while a ball-bearing fault is characterized by an increment in the amplitude of a family of harmonics.

In traditional Machine Vibration Signature Analysis (MVSA), the Fourier transform is used to determine the vibration spectrum (Lachouri et al., 2008), and the signature at different frequencies are identified and compared with initial measurement to detects faults in the machine. The short coming of this approach is that the Fourier analysis is limited to stationary signals while vibrations are not stationary by nature.

The Multi-Scale Principal Component Analysis (MSPCA) proposed by Bakshi (Bakshi, 1998) deals with processes that operate at different scales, and have contributions from:

- events occurring at different localizations in time and frequency;
- stochastic processes whose energy or power spectrum changes with time and/or frequency;
- variables measured at different sampling rate or containing missing data.

Thus, a MSPCA formulation, in which contributions from events occurring at different scales are captured by Principal Component Analysis (PCA) models at the corresponding scale, seems to be appropriated for extracting information from vibration data. Moreover wavelets, with their time-frequency localization and multi-resolution property, can be used as a useful framework for multi-scale representation of data (Manish et al., 2002).

The primary motivation for jointly using PCA and Wavelet Transform comes from the idea that, in PCA, the correlation among sensors is used to transform the multivariate space into a subspace which preserves maximum variance of the original space. However, PCA fails to make use of correlation within the sensor along the time line. In other words, it does not utilize the information pertaining to the frequency or scale characteristics of the individual sensors. Wavelets, on the other hand, capture correlation within a sensor whereas PCA correlates across sensors (Manish et al., 2002). Thus, wavelets and PCA based analysis of multivariate data represent two extremes, one, making use of only the signal trend, and the other, using

only correlation. The Multi Scale PCA (MSPCA) is a way to combine these two techniques, to extract maximum information from multivariate sensor data.

In the present paper vibration analysis of rotating electrical machines is considered. MSPCA is applied for fault detection and diagnosis: once a fault is detect a multiscale fault identification is performed, and the recursive version of MSPCA is used because it is able to detect and identify faults on-line. The paper is organized as follows. In Section 2 and 3 Principal component analysis and Wavelet Transform are briefly introduced respectively. In section 4 the fault detection and diagnosis procedure is discussed. The test bench used is described in Section 5. Results on experimental tests are reported in Section 6. The paper ends with comments on the performance of the proposed diagnostic methods.

2. PRINCIPAL COMPONENT ANALYSIS

By projecting data into a lower-dimensional space that accurately characterizes the state of the process, dimensionality reduction techniques can greatly simplify and improve process monitoring procedures. PCA is a dimensionality reduction technique. It produces a lower-dimensional representation in a way that preserves the correlation structure between the process variables, and it can capture the variability in the data.

In PCA, the correlation among sensors is used to transform the multivariate space into a subspace which preserves maximum variance of the original space in minimum number of dimensions. In other words, PCA rotates the original coordinate system along the direction of maximum variance.

Consider a data matrix $\boldsymbol{X} \in \mathbb{R}^{n \times m}$ consisting of n sample rows and m variable columns that are normalized to zero mean and unit variance. The matrix $\boldsymbol{X}$ can be decomposed into a score matrix $\boldsymbol{T}$ and a loading matrix $\boldsymbol{P}$ whose columns are the right singular vectors of $\boldsymbol{X}$ as follows:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^T + \tilde{\boldsymbol{X}} = \boldsymbol{T}\boldsymbol{P}^T + \tilde{\boldsymbol{T}}\tilde{\boldsymbol{P}}^T \tag{1}$$

where $\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{T}}\tilde{\boldsymbol{P}}^T$ is the residual matrix (Manish et al., 2002). Once a PCA model is built and a new data sample $\boldsymbol{x}$ is to be tested for fault detection, it is first scaled and then decomposed as follows:

$$\boldsymbol{x} = \hat{\boldsymbol{x}} + \tilde{\boldsymbol{x}} \tag{2}$$

where,

$$\hat{\boldsymbol{x}} = \boldsymbol{P}\boldsymbol{P}^T\boldsymbol{x} \in S_p \tag{3}$$

is the projection on the Principal Component Subspace (PCS), S_p, and

$$\tilde{\boldsymbol{x}} = (\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^{\boldsymbol{T}})\boldsymbol{x} \in S_r \tag{4}$$

is the projection on the Residual Subspace (RS), S_r (Manish et al., 2002).

For fault detection in the new sample x, a deviation in x from the normal correlation could change the projections onto the subspaces, either S_p or S_r. Consequently, the magnitude of either $\tilde{\boldsymbol{x}}$ or $\hat{\boldsymbol{x}}$ could increase over the values obtained with normal data.

The Square Prediction Error (SPE), also known as Q, is a statistic that measures lack of fit of a model to data. The SPE statistic indicates the difference, or residual, between a sample and its projection into the k components retained in the model. The exact description of the distribution of SPE is given in (Jackson, 2003):

$$SPE \equiv \|\bar{\boldsymbol{x}}\|^2 = \left\|(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^{\boldsymbol{T}})\boldsymbol{x}\right\|^2. \tag{5}$$

The process is considered normal if:

$$SPE \leq \delta^2 \tag{6}$$

where δ^2 is a confidence limit for SPE. A confidence limit expression for SPE, when x follows a normal distribution, is developed in (Jackson and Mudholkar, 1979; Manish et al., 2002; Lachouri et al., 2008; Ciandrini et al., 2010). In this work, the SPE test is used as the main criterion for fault detection.

3. WAVELET TRANSFORM

The Wavelet Transform (WT) is defined as the integral of the signal $\boldsymbol{x}(t)$ multiplied by scaled, shifted version of basic wavelet function $\boldsymbol{\psi}(t)$, a real valued function who's Fourier transform satisfies the admissibility criteria (Li et al., 1999). Then the wavelet transformation $C(\cdot,\cdot)$ of a signal $s(t)$ is defined as below:

$$\begin{array}{l} c(a,b) = \int_{\mathbb{R}} s(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)dt \\ a \in \mathbb{R}^+ - \{0\} \\ b \in \mathbb{R} \end{array} \tag{7}$$

where a is the so-called scaling parameter, b is the time localization parameter. Both a and b can be continuous or discrete variables. Multiplying each coefficient by an appropriately scaled and shifted wavelet yields the constituent wavelets of the original signal. For signals of finite energy, continuous wavelets synthesis provides the reconstruction formula:

$$s(t) = \frac{1}{K_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}^+} c(a,b)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)\frac{da}{a^2}db \tag{8}$$

where

$$K_\psi = \int_{-\infty}^{+\infty} \frac{\mid \hat{\psi}(\xi) \mid}{\mid \xi \mid}\mathrm{d}\xi \tag{9}$$

denotes a (Wavelet specific) normalization parameter in which $\hat{\psi}$ is the Fourier transform of ψ.

Mother wavelets must satisfy the following properties:

$$\int_{-\infty}^{+\infty} |\psi(t)|dt < \infty,\ \int_{-\infty}^{+\infty} |\psi(t)|^2 dt = 1,\ \int_{-\infty}^{+\infty} \psi(t)dt = 0. \tag{10}$$

To avoid intractable computations when operating at every scale of the Continuous WT (CWT), scales and positions can be chosen on a power of two, i.e. dyadic scales and positions. The Discrete WT (DWT) analysis is more efficient and just as accurate. In this scheme a and b are given by:

$$(j,k) \in \boldsymbol{Z}^2 : a = 2^j,\ b = 2^j k,\ \boldsymbol{Z} := \{0, \pm 1. \pm 2, \cdots\}. \tag{11}$$

Then defining:

$$(j,k) \in \boldsymbol{Z}^2 : \psi_{j,k} = 2^{-j/2}\psi\left(2^{-j}t - k\right), \tag{12}$$

the discrete wavelet analysis can be described mathematically as:

$$\begin{gathered} c(a,b) = c(i,j) = \textstyle\sum_{n\in\boldsymbol{Z}} s(n)\psi_{j,k}(n), \\ a = 2^j,\ b = 2^j k, \\ j \in \boldsymbol{Z},\ k \in \boldsymbol{Z}. \end{gathered} \tag{13}$$

The inverse transform, also called discrete synthesis, is defined as:

$$s(n) = \sum_{j\in\boldsymbol{Z}} \sum_{k\in\boldsymbol{Z}} c(j,k)\psi_{j,k}(n). \tag{14}$$

The detail level j, and the approximation at level j are defined as:

$$\begin{gathered} D_j(t) = \textstyle\sum_{k\in\boldsymbol{Z}} c(j,k)\psi_{j,k}(t) \\ A_{j-1} = \textstyle\sum_{j>J} D_j \end{gathered} \tag{15}$$

and the following equations hold:

$$\begin{gathered} A_{j-1} = A_j + D_j, \\ s = A_j + \textstyle\sum_{j\le J} D_j. \end{gathered} \tag{16}$$

4. MSPCA FORMULATION

Wavelet Transformation and Principal Component Analysis can be combined to extract both correlation within the sensors and cross correlation among sensors, in this way it is possible to extract maximum information from multivariate sensor data. MSPCA can be used as a tool for fault detection and diagnosis by means of statistical indexes. In particular, faults are detected by use the SPE (5) and the diagnosis is conducted by the SPE contribution method. The contribution is the technique of computing the SPE of the sensors separately. In this way it is possible to detect which sensor is most affected by fault (Manish et al., 2002).

Process monitoring by MSPCA involves computing independent principal components loadings and detection limits for the scores and residuals at each scale from data representing normal operations. For new data, a statistically significant change is indicated if the residuals based on wavelet coefficients computed from the most recent measurements violate the detection limits at any scale. Since the wavelet coefficients are sensitive only to changes, if a variable goes outside the region of normal operation and stays there, the wavelet coefficients will be statistically significant only when the change first occur. The change is first detected at the finest scale that covers the frequencies present in the feature representing abnormal operation. If the change persists, it is detected by wavelet coefficients at coarser scales present in the abnormal feature. In this condition, the most recent scaling function coefficients are the last ones to violate the detection limit, and continue to do so for as long as the process remains in an abnormal state. Similarly, when a process returns from abnormal operation, the change will be detected by wavelet coefficients, but the scaling function coefficients will continue to indicate abnormal operation for several samples due to the coarser scale of representation. Thus, process monitoring based only on PCA of the wavelet and scaling function coefficients will not allow quick and continuous detection of faults, and may create false alarms. The last MSPCA step of reconstructing the signal to the time domain, and computing the scores and residuals for the reconstructed signal improves the speed of detection abnormal operation and eliminates false alarm (Bakshi, 1998).

The following algorithm is derived from (Ciandrini et al., 2010), and it is improved to include the diagnosis by means of SPE contribution:

1. The Wavelet analysis is used, to refine the data, with a level of detail, L;
2. Normalize mean and standard deviation of detail and approximation matrices and apply PCA to the approximation matrix A_L, of order L, and to the L detail matrices D_l, where $l = 1, ...L$;
3. The PCA transformation matrix P and the covariance matrix S are computed for each approximation and detail matrices;
4. The SPE index (5) is computed, for each wavelet matrix;
5. The SPE threshold δ^2 is computed, for each detail matrix and for the approximation matrix of order L, using the following equation (Jackson and Mudholkar, 1979):

$$\begin{gathered} \delta^2 = \theta_1 \left[\frac{h_0\ C_\alpha\ \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2(h_0^2 - h_0)}{\theta_1^2}\right]^{1/h_0} \\ h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \\ \theta_i = \textstyle\sum_{j=p+1}^{n} \lambda_j^i \end{gathered} \tag{17}$$

   where p is the PCA subspace dimension, and λ_j^i is the j-th eigenvalue of the covariance matrix S_l of the detail matrix D_l, where $l = 1, .., L$ is the level of detail index; C_α is the normal cumulative distribution, with a significance level α;
6. The inverse Wavelet is applied, to reconstruct the variables behavior;
7. The PCA is applied to the reconstructed and normalized signals;
8. The SPE index (5) is computed for the transformation and covariance matrix of the reconstructed signal;

9. The previous steps are repeated for each new dataset, except for the threshold computation, computing the PCA and SPE index using the P and S matrices, obtained with the first dataset;
10. **If** the SPE is over the threshold δ^2, the fault is detected and the SPE contribution diagnosis is performed, **else** the next data set is analyzed (return to 1);
11. Compute the SPE contribution for each sensor for all details and approximation matrices and diagnose the type of fault.

5. DEVELOPED EXPERIMENTAL SETUP

The monitoring and diagnosis system have been prototyped for plants whose critical components, that need to be monitored, are electrical rotating machines. Signal-based fault detection and prediction procedures have been developed and tested on a laboratory-scale experimental system, in which vibration and absorbed current are acquired. In rotating machines vibrations arise along two main directions: axial and radial. Thus two accelerometers are used. Integrated Circuit-Piezoelectric (ICP®) accelerometers produced by PCB (model $333B50$) have been used (PCB Piezotronics, 2009). These sensors incorporate built-in, signal-conditioning electronics. The Table 1 contains the complete performance characteristics of the sensor. The NI CompactRIO (NI cRIO) 9004, product by National Instruments, is used for data acquisition (National Instruments, 2009). The module equipped the NI cRIO 9004 is NI 9233 used to acquire signal from accelerometers, that is characterized by a 24-bit (delta-sigma) resolution analog-to-digital converter (ADC) with a sampling frequency up to $50kS/s$. An oversampling frequency is used for the ADC converter and then a digital band pass filter is applied. The filter band has been designed on the basis of an accurate study of the considered machines. The acceleration is acquired, the velocity is computed, and the following indices are obtained: Root Mean Square (RMS), standard deviation, skewness and kurtosis. These indices are computed over a samples window. The frequency spectrum is calculated for vibration acceleration and velocity and it is averaged over a proper set of observations. These operations allow a memory saving acquisition but both drift faults and abrupt ones can be detect with sufficient accuracy. The motor current is measured by a LEM current sensor LTSR 6-NP (LEM U.S.A., 2009), and the signal is processed in the same way of accelerometers. A scalable system has been developed where data are stored and many procedures have been tested and prototyped. The scalable system allows to monitor many machines with simultaneous data acquisition and fault detection analysis.

6. RESULTS

This approach has been implemented, using a Haar wavelet kernel, a level of detail $L = 3$, and the dimension of Principal Components subspace, p, is chosen by the Kaiser's rule (Jolliffe, 2002). For the training stage of PCA a faultless dataset is selected and an outlier elimination is performed in order to ensure robustness to analysis procedure. In (Ho et al., 2002) the training procedure with different outlier elimination techniques is presented.

Table 1. Integrated Circuit-Piezoelectric (ICP®) accelerometers.

Sensitivity	$1000[mV/g]$
Measurement Range	$\pm 5V$
Frequency Range	$0,5 - 3000Hz$
Resonant Frequency	$\geq 20kHz$
Non-Linearity	$\leq 1\%$
Temperature Range	-18 to $+66C$
Excitation Voltage	$18 - 30VDC$
Constant Current Excitation	$2 - 20mA$
Output Impedance	$\leq 500\Omega$
Sensing Element	Ceramic
Size(HxLxW)	$11.4mm$ x $17.3mm$ x $11.4mm$
Weight	$7.5gm$
Electrical Connector	$10 - 32$ Coaxial Jack

The hipotesys, under which this procedure can be implemented, is the Gaussian distribution of signals. Here, the outlier elimination is performed using a Huber estimator for matrix T with parameter $k = 4$ (Ho et al., 2002). The fault detection and diagnosis algorithm is performed online recursively, by means of a moving window using, at each step, at least 2^L samples needed by the DWT.

The diagnosis procedure is implemented using two accelerometer and a current sensor, to demonstrate the effectiveness of MSPCA approach, three different faults are induced. A sample is shown in Figure (1), where the sensors signal for the backlash fault are shown.

Considered faults are: unbalance, backlash and misalignment of the rotor shaft respectively. The objective is early detection and identification of abnormal situations. Incoming data samples from the moving window are then fed into the MSPCA model and the SPE is computed at each scale, and it is compared with the threshold. In this work the fault detection is performed using the SPE of reconstructed signal only. If it is below the limit, the sample is assumed to be normal. Else the sample is considered faulty and the SPE contribution is performed for the approximation matrix A_3.

In Figure 2 the test on the unbalance fault is performed, the SPE of the reconstructed signal is shown in Figure 2(a), while the Approximation matrix A_3 is shown in Figure 2(b). Once the fault is detected the SPE contribution weights are calculated for the approximation matrix and they are shown in Figure 2(c). In the same way it is possible to detect the backlash fault and isolate the respectively signature by means of contribution of the approximation matrix A_3. In Figure 3(a) is shown the reconstructed SPE and in Figure 3(b) the approximation matrix SPE. The SPE contribution weights are shown in Figure 3(c), this represent the signature of backlash fault.

In the case of misalignment, the MSPCA performance are shown in Figure 4. In particular the SPE of reconstructed signal is shown in Figure 4(a), while the approximation matrix SPE is shown in Figure 4(b). The SPE contribution weights at sample 50 are highlighted in Figure 4(c).

7. CONCLUSIONS

In the field of operation efficiency, fault detection and diagnosis procedures help to decrease the cost of maintenance. In this area vibration analysis is reliable, non destructive and allows continuous monitoring. MSPCA is a fault de-

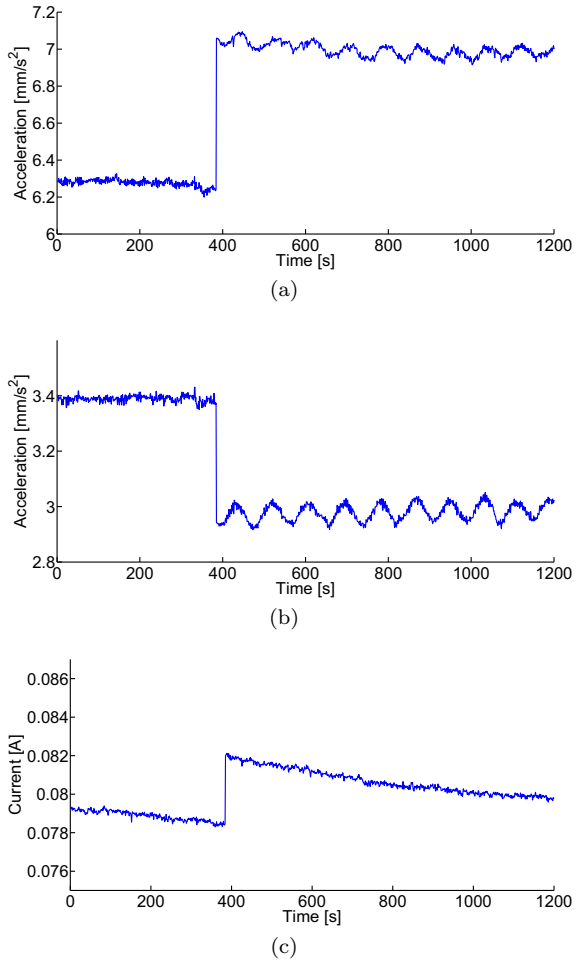


Fig. 1. Backlash fault occur at time 384. (a) Radial accelerometer signal; (b) axial accelerometer signal; (c) current signal.

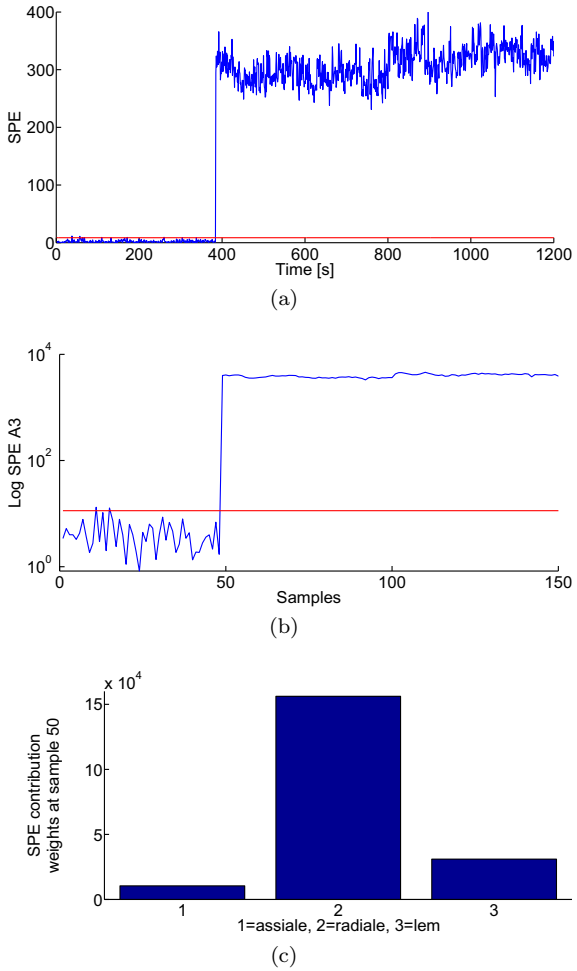


Fig. 2. Unbalance fault at time 384. (a) SPE of reconstructed signal; (b) SPE of approximation matrix A_3, in samples; (c) SPE contribution weights of approximation matrix A_3, at sample 50.

tection and diagnosis procedure that deal with stochastic processes whose power spectrum changes in time and/or frequency. The MSPCA combine two techniques that allow extraction of correlation among sensors (PCA approach) and correlation within sensor (Wavelet approach).

The fault detection is performed by means of statistical index: the SPE of the reconstructed signal. Diagnosis is performed using SPE contribution weights on the approximation matrix, which represent the signature of the isolated fault. Results show that, this approach produce a single signature for each fault tested: unbalance, misalignment and backlash. This approach is effective, the SPE is very sensitive to faults, and the procedure keep robustness thanks to outlier elimination. In conclusion this approach is suitable for monitoring critical machines in plants, and helps for early detection and identification of faults.

An open problem is the generalization of the proposed approach, in other words to extend the procedure to different plants without a redesign of the components. Future researches intend to analyze this aspect by the introduction of intelligent components for tuning and adapting each step of the proposed algorithm.

REFERENCES

Bagheri, F., Khaloozaded, H., and Abbaszadeh, K. (2007). Stator fault detection in induction machines by parameter estimation, using adaptive kalman filter. In *MED Conf. on Control & Automation*, 1–6. Athens.

Bakshi, B.R. (1998). Multiscale pca with application to multivariate statistical process monitoring. *AIChE Journal*, 44, 1596–1610.

Bellini, A., Immovilli, F., Rubini, R., and Tassoni, C. (2008). Diagnosis of bearing faults of induction machines by vibration or current signals: A critical comparison. In *IEEE Industry Applications Society Annual Meeting, IAS '08*, 1–8. Edmonton Alta.

Benloucif, M. and Balaska, H. (2006). Robust fault detection for an induction machine. In *World Automation Cong. WAC '06*, 1 – 6. Budapest.

Ciandrini, C., Gallieri, M., Giantomassi, A., Ippoliti, G., and Longhi, S. (2010). Fault detection and prognosis methods for a monitoring system of rotating electrical machines. In *Proc. of the IEEE Int. Sym. on Ind. El.*, 2085 – 2090. Bari.

Gani, A. and Salami, M. (2002). Vibration faults simulation system (vfss): a system for teaching and training on fault detection and diagnosis. In *Student Conf.*

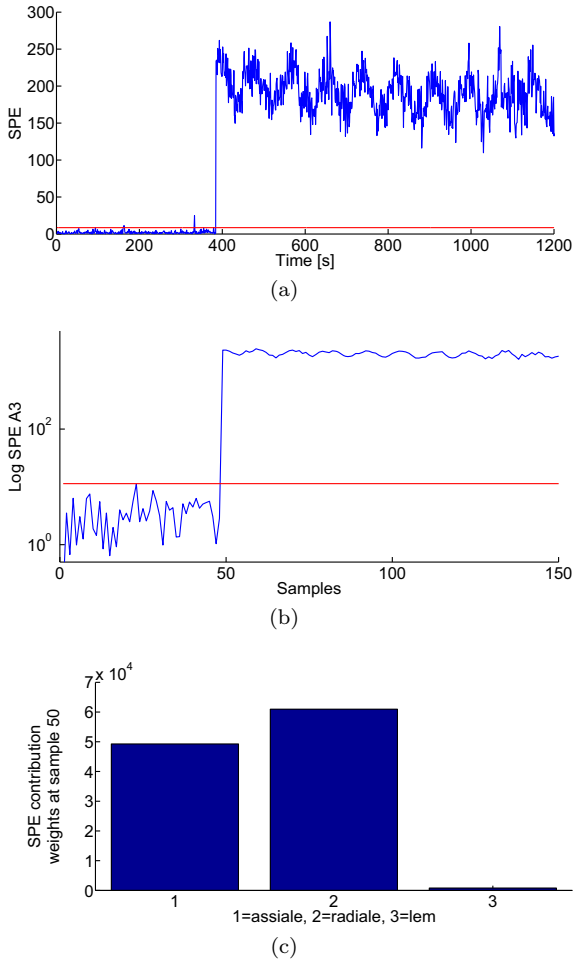


Fig. 3. Backlash fault induced at time 384. (a) SPE of reconstructed signal; (b) SPE of approximation matrix A_3, in samples; (c) SPE contribution weights of approximation matrix A_3, at sample 50.

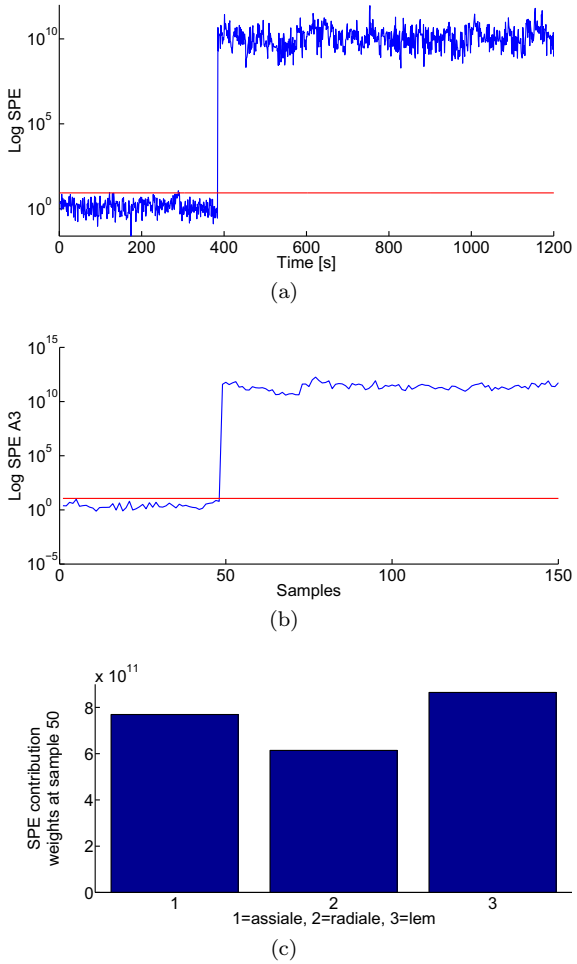


Fig. 4. Misalignment fault induced at time 384. (a) SPE of reconstructed signal; (b) SPE of approximation matrix A_3, in samples; (c) SPE contribution weights of approximation matrix A_3, at sample 50.

on Research and Devel. Proceed., 15–18. Shah Alam Malaysia.

Ho, K.A., Tvarlapat, K.J., Piovoso, M.J., and Hajare, R. (2002). A method of robust multivariate outlier replacement. *Computer and Chemical Engineering*, 26, 17–39.

Jackson, J.E. (2003). *A User's Guide to Principal Components*. Wiley-Interscience, New York.

Jackson, J. and Mudholkar, G. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, 21, 341–349.

Jolliffe, I.T. (2002). *Principal component analysis*. Springer, Berlin.

Lachouri, A., Baiche, K., Djeghader, R., Doghmane, N., and Ouhtati, S. (2008). Analyze and fault diagnosis by multi-scale pca. In *Int. Conf. on Information and Communication Technologies*, 1–6. Damascus Syria.

LEM U.S.A., Inc., h. (2009).

Li, X., Dong, S., and Yuan, Z. (1999). Discrete wavelet transform for tool breakage monitoring. *Int. Journal of machine tool manufacture*, 99, 1944–1955.

Liu, H., Li, Z., and Zhaowei (2009). Time frequency distribution for vibration signal analysis with application to turbo-generator fault diagnosis. In *Chinese Control and Decision Conf. CCDC '09*, 5492 – 5495. Guilin.

Manish, M., Yuea, H.H., Qin, S.J., and Lingb, C. (2002). Multivariate process monitoring and fault diagnosis by multi-scale pca. *Computers & Chemical Engineering*, 26, 1281–1293.

National Instruments, Inc., h. (2009).

PCB Piezotronics, Inc., h. (2009).

Taniguchi, S., Akhmetov, D., and Dote, Y. (1999). Fault detection of rotating machine parts using novel fuzzy neural network. In *Proc of IEEE Int. Conf. on Systems, Man, and Cybernetics*, 365–369. Tokyo.

Tavner, P. (2008). Review of condition monitoring of rotating electrical machines. *Electric Power Applications IET*, 2, 215–247.

Verucchi, C.J. and Acosta, G.G. (2008). Fault detection and diagnosis techniques in induction electrical machines. *IEEE Trans. Latin America*, 5, 41–49.

Wan, S., Li, H., and Li, Y. (2002). Adaptive radial basis function network and its application in turbine-generator vibration fault diagnosis. In *Proc. on Int. Conf. on Power System Technology*, 1607–1610. Kunming.

Wu, Z., Li, F., Yan, S., and Wang, B. (2009). Motor fault diagnosis based on the vibration signal testing and analysis. In *Third Int. Sym. on Int. Inf. Technology Application*, 433–436. Nanchang.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Reconfiguration of over-actuated consecutive-k-out-of-n: F systems based on Bayesian Network Reliability model

P. Weber, C. Simon, D. Theilliol

CRAN, UMR 7039 Nancy-Université, CNRS
Boulevard des Aiguillettes B.P. 70239 F-54506 Vandœuvre lès Nancy
e-mail: { Philippe.Weber, Christophe.Simon, Didier.Theilliol }@cran.uhp-nancy.fr

Abstract: This paper presents a new approach based on reliability to reconfigure redundant actuators when failures occur. The aim is to preserve the health of the actuators and the availability of the system both in the nominal situation and in the presence of some unavailable actuators. This paper proposes a solution to control an over-actuated system that is structured as a typical consecutive-k-out-of-n: F system. In a degraded situation, this work proposes a reconfigured control allocation based on the on-line estimation of actuators reliability. The analytic solution proposed in the literature to compute the reliability of consecutive-k-out-of-n: F system is not appropriated to compute the reliability on-line. The graphical aspect of Bayesian Network is very interesting because it formalizes the model by coupling a generic model structure with simple parameter matrices and the inference computes the reliability of actuators according to on-line observations (evidences). It is applied on circular and linear typical consecutive-k-out-of-n: F system to estimate its reliability and provide the parameters to distribute the control efforts among the redundant set of actuators.

Keywords: Probabilistic model, Bayesian Network, Reconfiguration, Fault Tolerant System

1. INTRODUCTION

In order to respect the growing of economic demand for high plant availability and reliability, fault-tolerant control (FTC) is introduced. The aim of FTC is to keep plant available by the ability to achieve the objectives that have been assigned in the faulty behavior and accept reduced performances when critical faults occur (Blanke *et al*. 2006 and Noura *et al*. 2009).

In most safety critical systems, the actuators redundancy is often used. Particular cases of *k-out-of-n* (*koon*) systems are consecutive-*koon* systems proposed by (Kontoleon 1980) which have been used to model various engineering systems such as microwave stations of telecom networks, oil pipelines, and vacuum systems in an electronic accelerator… All these systems are over actuated systems based on redundancy of actuators to increase the system reliability.

The management of complex industrial systems contributes to higher competitiveness and higher performances. Thus, the relevance of dependability analyses increased due to their role in improving availability, performance, products quality, on-time delivery, and environment requirements (Alsyouf, 2007 and Kutucuoglu, *et al*. 2001). Nowadays, one of the major problems in the dependability field is addressing the system modeling in relation to the increase of its complexity. This modeling task underlines issues concerning the quantification of the model parameters and the representation, propagation and quantification of the uncertainty (Zio, 2009). To model the reliability of complex industrial systems, it is observed a growing interest focused on Bayesian Network (BN) in the recent literature (Weber *et al.* 2010). This modeling method is not the solution to all problems, but it seems to be very relevant in the context of complex systems (Langseth, 2008). BN are particularly interesting because they are able to compute the reliability taking into account observations (evidences) about the state of some components. For instance, it is possible to estimate the reliability of the system and all its components knowing that a part of them are out of order.

Our proposal consists mainly in formalizing a standard structure model as the consecutive-*koon*: F system with a BN model and demonstrate the usability of this model in a reconfigurable control problem of over-actuated system. For this purpose, the paper is organized as follow. Section 2 recalls the reliability computation of consecutive-*koon*: F system**.** In section 3, a solution for a reliable reconfigurable control of over-actuated systems is presented based on the on-line actuators reliability indicators. Section 4 introduces BN and explains the generic structure and parameter to model reliability of consecutive-*koon*: F system. Section 5 presents numerical applications of the BN modeling capabilities. Finally the conclusion is given by integrating also future research directions.

2. CONSECUTIVE-K-OUT-OF-N: F SYSTEM

The consecutive-*koon*: F system has attracted considerable attentions since it was first proposed by (Kontoleon 1980). Consecutive-*koon* systems can be classified according to the linear or circular arrangement of its components and the functioning or malfunctioning principle. A consecutive-*koon*: F system consists of a set of *n* ordered components that composed a chain such that the system is failed if at least *k*

consecutive components are failed. A consecutive-*koon*: G system is a chain of *n* components such that the system works if at least *k* consecutive components work. Thus, four types of *koon* can be enumerated: linear consecutive-*koon*: F, linear consecutive-*koon*: G, circular consecutive-*koon*: F and circular consecutive-*koon*: G. An illustration of these specific structures can be found in telecommunication systems with *n* relay stations that can be modeled as a linear consecutive-*koon* system, if the signal transmitted from each station is strong enough to reach the next *k* stations. An oil pipeline system for transporting oil from point to point with *n* spaced pump stations is another example of linear consecutive-*koon* system. A closed recurring water supply system with *n* water pumps in a thermo-electric plant is a good example of a circular system. The system ensures its mission if each pump can be powerful enough to pump water and steam to the next *k* consecutive pumps (Yam, 2003).

The size of the system given by the number of components is one factor of its complexity. But, the distribution of component reliability and the independence between components are the major complexity factors. The assumption about the distribution is required before studying the reliability of the system based on the reliability of its components. The different assumptions are: independent and identically distributed (iid), independent and non-identically distributed (inid) and the same cases with a dependence between components.

The exact system reliability for linear consecutive-*koon* in iid case has been provided by (Lambiris and Papastavridis 1985):

$$R(p,n,k)=\sum_{i=0}^{n}\binom{n-ik}{i}(-1)^i\left(pq^k\right)^i - q^k\sum_{i=0}^{n}\binom{n-ik-k}{i}(-1)^i\left(pq^k\right)^i \quad (1)$$

Lambiris also provides the exact formula for the circular case:

$$R_c(p,n,k)=\sum_{i=0}^{n}\binom{n-ik}{i}(-1)^i\left(pq^k\right)^i + k\sum_{i=0}^{n}\binom{n-ik-k-1}{i}(-1)^{i+1}\left(pq^k\right)^{i+1} - q^n \quad (2)$$

if $n \geq k$

Notation:

- n number of components in the system
- k minimum number of consecutive failed components which cause system failure
- p probability that a component functions q=1-p
- $R(p,n,k)$ reliability of a linear consecutive-*koon*: F system
- $R_c(p,n,k)$ reliability of a circular consecutive-*koon*: F system

3. ON-LINE CONTROL RE-ALLOCATION BASED ON RELIABILITY

Let us consider the model of a consecutive-*koon*: F system, with *n* actuators, given by:

$$\begin{aligned}\dot{x}(t) &= Ax(t)+B_u u(t)\\ y(t) &= Cx(t)\end{aligned} \quad (3)$$

where $A\in R^{m\times m}$, $B_u\in R^{m\times n}$ and $C\in R^{p\times m}$ are respectively, the state, the control and the output matrices. $x\in R^m$ is the system state, $u\in R^n$ is the control input, $y\in R^p$ is the system output, and (A,B_u) is stabilizable.

Control allocation is generally used for over-actuated systems, where the number of operable control is greater than the controlled variables. It is the case of a *koon*: F system that $rank(B_u)=q<n$. This implies that B_u can be factorized as: $B_u=B_v B$, where $B_v\in R^{m\times q}$ and $B\in R^{q\times m}$. Thus an alternative description of (3) can be given by:

$$\begin{aligned}\dot{x}(t) &= Ax(t)+B_v u(t)\\ v(t) &= Bu(t);\ y(t)=Cx(t)\end{aligned} \quad (4)$$

where $v\in R^q$ is the virtual control input, called as the total control efforts produced by the actuators and defined by the controller. For simplicity and for this study, the case $q = p$ is considered (i.e. when the number of virtual controls equals the number of variables to be controlled). The control allocation problem can be expressed as a constrained linear mapping problem based on the relation:

$$v(t)=Bu(t) \quad (5)$$

with $u_{\min}\leq u\leq u_{\max}$ the physical actuators saturation. Optimized based control allocation methods aim to find an optimal solution. If there is no exact solution, the optimal control is the feasible one such that $Bu(t)$ approximates $v(t)$ as well as possible. The optimal control input can be obtained by a two step optimization, namely sequential quadratic programming:

$$\psi=\arg\min_{u_{\min}\leq u\leq u_{\max}}\left\|Bu-v\right\|_2 \text{ and } u=\arg\min_{u\in\psi}\left\|W_u u\right\|_2 \quad (6)$$

where ψ is the set of feasible solutions subject to the objective criterion (6). The weighting matrix $W_u\in R^{n\times n}\succ 0$ is used to give a specific priority level to the actuators. Moreover, the optimal control input $u=(u_1,u_2,...u_n)$, solution of the control allocation problem (6) is defined according to the values of the weighting matrix W_u.

In order to improve the safety of the system and preserve the actuators, a specific choice of the weighting matrix W_u is proposed based on the actuators reliability (Khelassi *et al*. 2010). The weighting matrix W_u is considered as a key to manage the redundant actuators and contribute to a reliable controller which improves the system reliability.

To perform the solution of the control allocation problem, and keep the set of the actuators available as long as possible, the desired efforts $v(t)$ defined by the controller are distributed relating to the actuators contribution in the system reliability, as follows:

$$W_u = \begin{pmatrix} R_1 & & 0 \\ & R_2 & \\ & \ddots & \\ 0 & & R_n \end{pmatrix} \succ 0 \qquad (7)$$

where R_i is the contribution of the actuator $i \in \{1,...n\}$. As a direct consequence the actuators are utilized in the control allocation proportionally to their functioning probability.

Let us consider the actuator *i* defined as a random variable w_i with two states {*Up*, *Down*}, and the system defined as a random variable *S* with two states {*Up*, *Down*}, the functioning probability of w_i is defined by:

$$R_i = \mathrm{P}(w_i = Up | S = Up) \qquad (8)$$

In order to integrate the actuators degradation in the reconfigured control allocation strategy, W_u is re-estimated and changed on-line according to the estimation of actuators reliabilities. Therefore if an actuator w_j is unavailable {*Down*}, the system is still working (because it is an over actuated system), and the functioning probabilities of each actuator are defined by:

$$\mathrm{P}(w_i = Up | w_j = Down, ..., S = Up) \qquad (9)$$

4. BAYESIAN NETWORK MODEL

BN appears to be a solution to model complex systems because they performs the factorization of variables joint distribution based on the conditional dependencies. The main objective of BN is to compute the distribution probabilities in a set of variables according to the observation of some variables and the prior knowledge of the others. The principles of this modeling tool are explained in (Jensen, 1996; Pearl *et al*. 1988).

Recall of BN characteristics

A BN is a directed acyclic graph (DAG) in which the nodes represent the system variables and the arcs symbolize the dependencies or the cause-effect relationships among the variables. A BN is defined by a set of nodes and a set of directed arcs. A probability is associated to each state of the node. This probability is defined, *a priori* for a root node and computed by inference for the others.

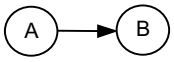


Figure 1. Basic example of a BN

The computation of nodes' probabilities is based on the probabilities of the parents' states and the conditional probability table (CPT). For instance, let's consider two nodes *A* and *B* with two states (S_{*1} and S_{*2}) each. The relation between nodes *A* and *B* is defines by the structure of the BN given on Figure 1. The *a priori* probabilities of node *A* are defined in Table 1:

A	
S_{A1}	S_{A2}
P(A=S_{A1})	P(A=S_{A2})

Table 1. *a priori* probabilities of node A

A CPT is associated to node *B*. This CPT defines the conditional probabilities P(B|A) attached to node *B* with a parent *A*, to define the probability distributions over the states of *B* given the states of *A* (Table 2).

	B	
A	S_{B1}	S_{B2}
S_{A1}	P(B=S_{B1}|A=S_{A1})	P(B=S_{B2}|A=S_{A1})
S_{A2}	P(B=S_{B1}|A=S_{A2})	P(B=S_{B2}|A=S_{A2})

Table 2. CPT of node B given node A.

Thus, the BN inference computes the marginal distribution for instance P(B=S_{B1}) by the following relation:

$$\begin{aligned} &\mathrm{P}(B = S_{B1}) = \\ &\mathrm{P}(B = S_{B1} | A = S_{A1}).\mathrm{P}(A = S_{A1}) \\ &+ \mathrm{P}(B = S_{B1} | A = S_{A2}).\mathrm{P}(A = S_{A2}) \end{aligned} \qquad (10)$$

The added value of a BN is linked to the computation of the probabilities attached to a node state given the state of one or several variables. BN are a powerful modeling and analysis tools for complex systems because it provides a lot of modeling advantages. A general inference mechanism (that permits the propagation as well as the diagnostic) is used to collect and to incorporate new information (evidences) gathered in a study. The Bayes´ theorem is the heart of this mechanism and allows updating a set of events´ probabilities according to the observed facts and the BN structure. It makes the strength of this knowledge management tool.

Generic consecutive-koon: F models

As mentioned in section 2, consecutive-*koon* systems are complex. Thanks to BN, their complexity is translated only in the graphical structure of the model. Moreover previous works based on formulas (1) and (2) are dedicated to iid components without dependences. A BN model is interesting because it provide an easy solution to model a non-identically

distributed (inid) component. The consecutive-*koon*: F model is completely defined from the combination of failed components. This combination respects a logical description as the union of minimal cutsets and is easily used to build the BN model.

Component $i \in \{1,...n\}$ is defined as a random variable w_j with two states $\{Up, Down\}$ and local aggregation of k consecutive components is defined as a random variable c_j with $j \in \{1,..., n-k+1\}$ for linear structure and $j \in \{1,..., n\}$ for circular structures. Variable c_j is defined by two states $\{Not\ Occurred, Occurred\}$. The system reliability is defined from the global system states modeled by variable S and its states $\{Up, Down\}$. Figure 2 and Figure 3 present the generic BN model structures for linear and circular consecutive-*koon*: F systems.

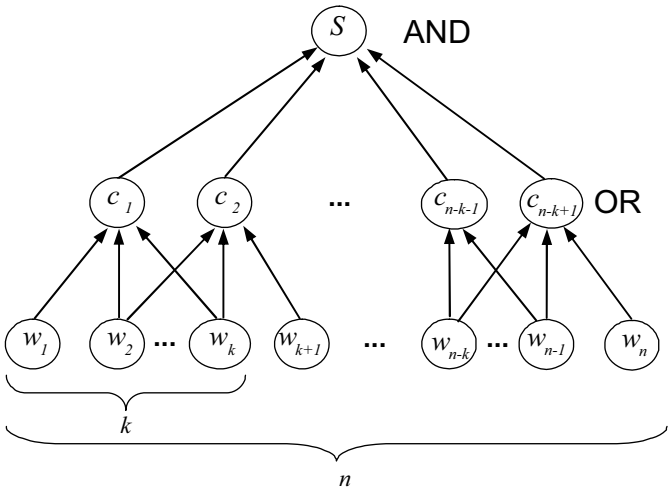


Figure 2. Generic BN model of linear consecutive-koon: F system

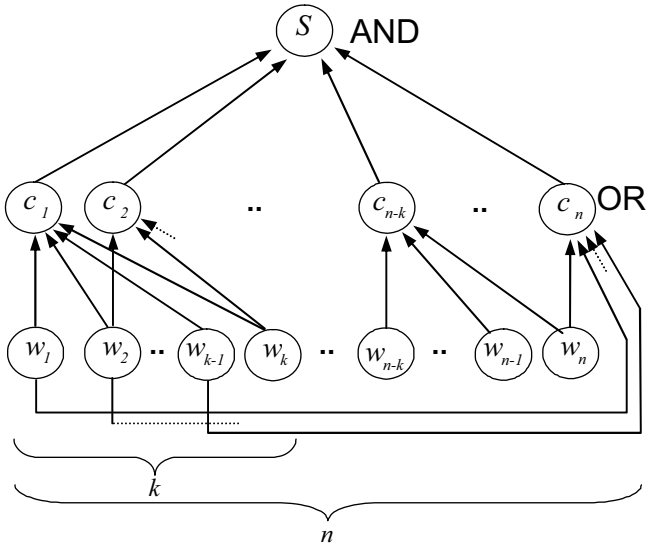


Figure 3.Generic BN model of circular consecutive-koon: F system

			c_j	
w_{i+k}	…	w_i	*Not Occurred*	*Occurred*
Up	*Up*	*Up*	1	0
	...	…	1	0
		Down	1	0
...	...	...	1	0
Down		*Up*	1	0
	...	...	1	0
	Down	*Down*	0	1

Table 3. CPT of node c_j.

			S	
c_{n-k} or c_{n-k+1}	…	c_1	*Up*	*Down*
Not Occurred	*Not Occurred*	*Not Occurred*	1	0
	...	…	0	1
		Occurred	0	1
...	...	...	0	1
Occurred		*Not Occurred*	0	1
	...	...	0	1
	Occurred	*Occurred*	0	1

Table 4. CPT of node S.

The CPT of c_j (Table 3) are defined from logical aggregation as OR gates to compute the occurrence probability of the minimal cutsets i.e. the local failure of k consecutive components and the system reliability is defined as a AND gate to compute the union of the minimal cutsets (Table 4). Therefore the parameters in the CPT are equal to 1 or 0 because there is no uncertainty on the combination of components' events leading to the failure of the system.

Finally the eq. (1) or (2) cannot be used to compute the probability that a component is available according to an on-line observation on the system (9). The BN is well appropriate to compute the probability that a component w_i is $\{Up\}$ according to the *a priori* knowledge on the component and the observations of some unavailable components.

$$\mathrm{P}(w_i = Up | w_j = Down, ..., S = Up) \qquad (11)$$

5. EXAMPLE

This section presents several numerical applications of consecutive-*2*-out-of-*5*: F system to linear and circular structure. The reliability of such systems is studied and a diagnosis with inspections scenario is realized to compute the on-line functioning probabilities of each actuator with the BN model.

Modeling structure and parameterization:
Figure 4 shows the application of the generic BN model of section 4 for a linear consecutive 2oo5: F system. Figure 5

shows the application of the generic BN model for Circular 2oo5: F system. The reader can see the spatial organization that shows the closeness of both linear and circular consecutive-*koon* BN models.

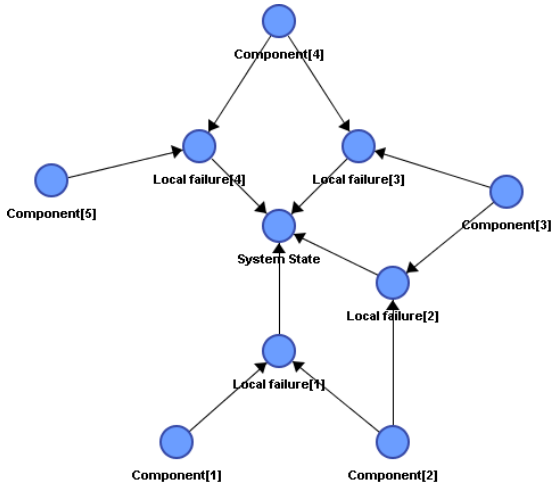


Figure 4. Linear consecutive-2-out-of-5: BN model

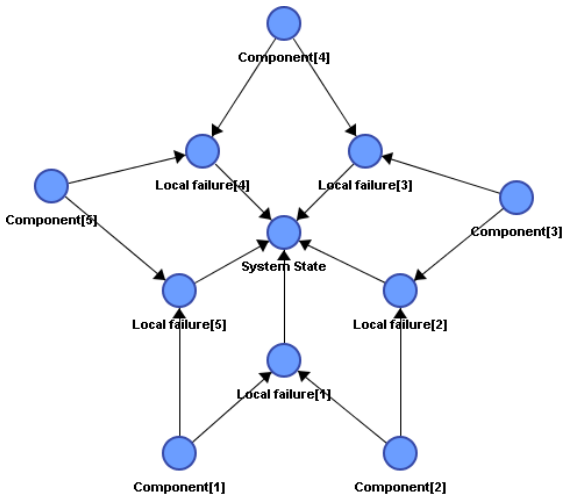


Figure 5. Circular consecutive-2-out-of-5: BN model

Reliability estimations:

For a first numerical application, let's consider *iid* failure rates of components $\lambda_i = 10^{-3} h^{-1}$ and the time of mission $T = 1000h$. The probability of components to be in state *Up* is: $P(w_i = Up) = \exp(-\lambda_i \times t)\big|_{t=T} = 0{,}3679$. The probability distribution of components and the probability distribution of the linear consecutive 2oo5: F system computed by the BN of Figure 4 is given on Figure 6. The circular consecutive 2oo5: F system reliability computed by the BN of Figure 5, based on its components reliability, is given on Figure 7.

The reader can verify that the probability of the systems to be in state *Up* corresponds to the reliability of the systems computed from relations (1) and (2), respectively for the linear and circular 2oo5: F system.

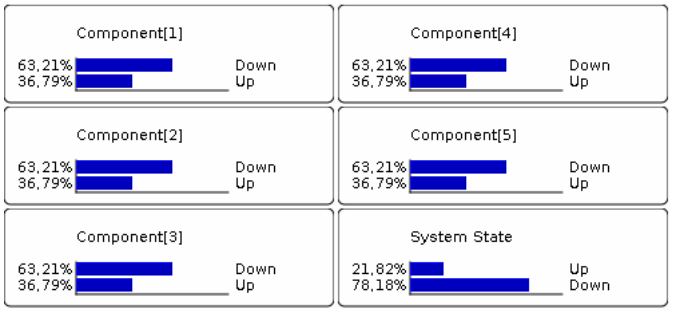


Figure 6. Linear consecutive-2oo5:F probability distribution

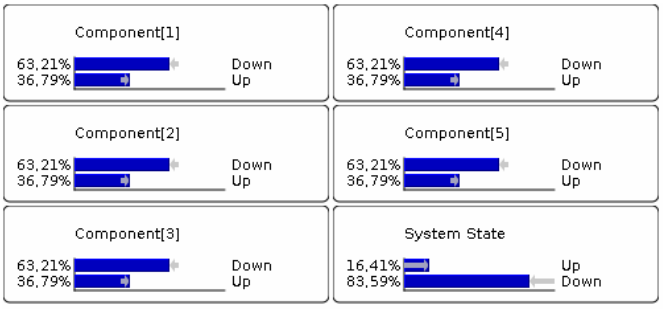


Figure 7. Circular consecutive-2oo5:F probability distribution

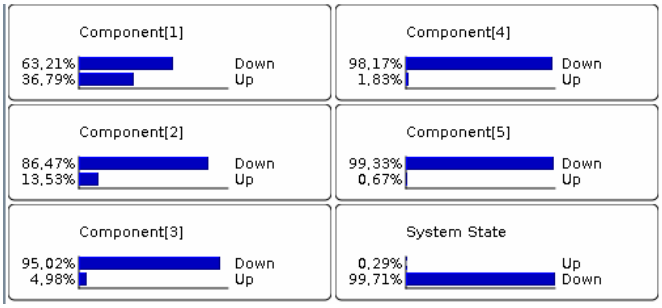


Figure 8. Circular consecutive-2oo5:F probability distribution

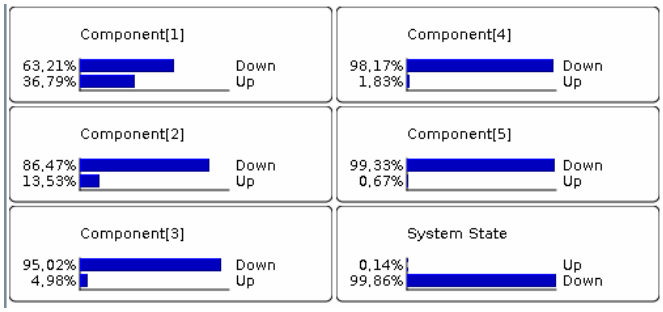


Figure 9. Circular consecutive-2-out-of-5 probability distribution

λ_1	λ_2	λ_3	λ_4	λ_5
10^{-3}	2.10^{-3}	3.10^{-3}	4.10^{-3}	5.10^{-3}
$P(w_1)$	$P(w_2)$	$P(w_3)$	$P(w_{41})$	$P(w_5)$
0,3976	0,1353	0,0498	0,0183	0,0067

Table 5. Failure rates and reliability of components.

For the second numerical application, let's consider an *inid* distribution of components failure rate and their

corresponding probability to be in state *Up* at the mission time. The numerical values are given in Table 5.
The reliability of the linear consecutive-2oo5: F and of the circular consecutive-2oo5: F systems are given on Figure 8 and 9. These two tests show the exactness of the BN models.

Diagnosis and weighting matrix computation
As mentioned in section 4, an advantage of BN is to naturally compute the probability of some variables given the value of several others. This ability of BN can significantly help managing the diagnostic of system. Moreover in this application to the on-line control re-allocation strategy, the BN model is used to compute the parameter of the weighting matrix W_u . For instance, let's consider the *inid* distribution (Table 5) of components probability given on Table 6 (step 0) for the circular consecutive-2oo5: F system at the time $T = 1000h$. In the (Step 1), the BN model computes the probability distribution of each component given *S* is *Up*. It can be interpreted as the importance contribution of each component to the functioning of the system. Then the actuators with the most important probability $P(w_i = Up)$ are more requested according to the equation (7). Considering the probabilities (Step 1), the component five is probably *Down* because $P(w_5 = Up) = 0.1199$ thus an inspection can be launched to verify its state. Two cases can occurred according to component five states. If the observation of the component five is *Down* then its probability $P(w_5 = Up) = 0$. This probability is used as evidence and the BN computes all other probabilities accordingly (step 2a). The probability of component one and four to be *Up* is equal to one, the next inspection should be launch on the component three. But, let's consider that component five is *Up*. The probability distribution of each component is different from (step 2a) as given on Table 6 (step 2b).

	Step 0	Step 1	Step 2a	Step 2b
$P(w_1 = Up)$	0,3976	0,9550	1	0,7911
$P(w_2 = Up)$	0,1353	0,7092	0,7586	0,3468
$P(w_3 = Up)$	0,0498	0,3549	0,2792	0,9102
$P(w_4 = Up)$	0,0183	0,8929	1	0,1064
$P(w_5 = Up)$	0,0067	0,1199	**0**	**1**
$P(S=Up)$	0,0014	**1**	**1**	**1**

Table 6. Diagnostic and inspection scenarios

6. CONCLUSION

In this paper, we have shown how a generic model of BN can easily handle the reliability analysis of complex system based on linear and circular consecutive-*koon* systems. In addition, we show how BN can be used to manage the diagnosis and inspection steps to identify the failed components in a complex system.

Moreover the computation of the functioning probabilities of each actuator is proposed to be done with BN inference. These distributions of probabilities are proposed to be used to define the weighing matrix W_u that is used to give a specific priority level to the actuators in the re-allocation problem of control. This method provides a control re-allocation that is based on on-line reliability estimation.

ACKNOWLEDGMENT

This work was suppported by the SAFE project 2010-2011 (Ageing Management in Fault-tolerant control system design project) from GIS 3SGS – France (https://www.gis-3sgs.fr/).

REFERENCES

Alsyouf, I. (2007) *The role of maintenance in improving companies' productivity and profitability.* International Journal of Production Economics, 105, 70–78.

Blanke M., M. Kinnaert, J. Lunze, and M. Staroswiecki. Diagnosis and fault tolerant control. *Control Systems Series, Springer-Verlag London*, 2006.

Jensen F.V. (1996). An Introduction to Bayesian Networks *Editions UCL Press*. London, UK.

Khelassi A., P. Weber, and D. Theilliol. Reconfigurable Control Design for Over-actuated Systems based on Reliability Indicators. Conference on Control and Fault-Tolerant Systems (SysTol'10), October 6-8, Nice, France, 2010.

Kontoleon J.M., *Reliability determination ofa r-successive-out-of-n:F system,* IEEE Trans. Reliability 29 (1980), 290-294.

Kutucuoglu, K., Hamali, J., Irani, Z. and Sharp, J., 2001. A framework for managing maintenance using performance measurement systems. International Journal of Operations and Production Management **21** 1/2, pp. 173–194

Lambiris, M. and Papastavridis, S., 1985, Exact reliability formulas for linear and circular consecutive-k-out-of-n: F systems, IEEE Trans. on Reliability, **34** (2), 124-126.

Langseth H. (2008). Bayesian Networks in Reliability: The Good, the Bad and the Ugly. Advances in Mathematical Modeling for Reliability. IOS Press. Amsterdam, Netherland.

Noura H., D. Theilliol, J.C. Ponsart, and A. Chamssedine. Fault tolerant control systems: Design and practical application. *Springer Dordrecht Heidelberg London*, 2009.

Pearl J. (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers Inc. San Francisco, USA.

Yam RCM., M.J. Zuo, Y.L. Zhang, A method for evaluation of *reliability* indices for repairable circular consecutive-*k*-out-of-*n*: *F* systems, Reliability Engrg. System Safety 79 (2003) 1–9.

Weber P., Medina-Oliva G., Simon C., Iung B., Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas, Engineering Applications of Artificial Intelligence, 2010.

Zio E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering and System Safety*. Volume 94, 125-141.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fault detection in flat systems by constraint satisfaction and input monitoring

Ramatou Seydou, Tarek Raïssi, Ali Zolghadri, David Henry

Bordeaux I University, IMS- lab, Automatic control group,
351 cours de la libération, 33405 Talence cedex, France

Abstract: This paper describes an application of a set-membership technique to robust fault detection for a class of nonlinear systems, the so-called flat systems. The proposed consistency test is built based on a comparison of an estimated feasible set and the expected value of the input vector. This strategy consists in eliminating models of the plant that are not consistent with the set of observations provided by the system sensors. The set-membership estimator design for the input vector takes into account the model uncertainties and disturbances, which makes the consistency test robust against such perturbations. The robustness of the proposed strategy is illustrated by simulations using several sensor/actuator faults scenarios.

Keywords: Flat systems, Constraint Satisfaction Problem (CSP), set-based observer, fault detection

1. INTRODUCTION

The issue of model-based Fault Detection and Isolation (FDI) in dynamic systems has been an active research area during the last three decades (see Ding [2008] for a recent survey). This paper considers observer-based fault detection for flat systems (Fliess et al [1992]). A system is called differentially flat, or just flat, if there exists a set of independent variables (to be called flat outputs of the system) such that both the system state and input vectors are functions of these flat outputs and a finite number of their successive derivatives. Flatness property offers an easy way to parameterize the dynamical behaviour of a system using flat outputs. In recent years the relevance of flat systems in control problems has been studied (see Agrawal and Sira-Ramirez [2004], Louembet et al. [2010], Rouchon [2008]). Most of the literature about flatness deals with control problems and few works are related to fault diagnosis. The main goal of this paper is to develop consistency tests for monitoring flat systems. The approach is based on "model invalidation". To achieve robust fault detection and to take into account uncertainties, a set-membership observer is used. The approach consists basically in formulating the state/input estimation into a Constraint Satisfaction Problem (CSP) (Norvig and Russell [2010]). CSPs consist of variables with constraints relating them. Many important real-world problems can be described as CSPs. The structure of a CSP can be represented by its constraint graph where the state and input vector constitutes the variable set and a mapping, relating the state and input to the flat outputs and their derivatives is taken as the constraints. Branch and prune algorithms (Goldsztejn [2006], Benhamou and Granvilliers [2006], Neumaier [2004]), based on consistency, are used to compute an outer approximation of the solution set of the CSP via interval analysis.

The set-membership observer is used to build consistency check tests based on a comparison of the feasible domain of the input resulting from a fault-free model simulation and the actual value. A fault occurrence will be indicated by an empty intersection. With respect to classical observer-based scheme, one advantage of the developed approach is the possibility to build consistency checks directly based on the input signal. In fact, the detection of incipient faults from output residuals may become difficult, especially when the permissible time window for detection is narrow. Here, the proposed methodology consists in estimating the input vector using a CSP-based observer and generating interval residual quantities that could be used to establish consistency checks in order to detect sensor or actuator faults.

The paper is structured as follows. Section 2 recalls some basic definitions for flatness, interval analysis, and CSP notions. In section 3, the CSP-based observer technique is detailed and illustrated through an example (section 4) in order to detect sensor and actuator faults. Finally some concluding remarks are given.

2. PRELIMINARIES

2.1 Flatness

1. Consider the following nonlinear system:

$$\begin{cases} \dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}, u) \\ y = h(\mathbf{x}) \end{cases} \quad (1)$$

System (1) is said to be flat with a flat output y if and only if one can describe the system states and inputs $(\mathbf{x}, u)$ only from the flat output and a finite number of its derivatives, i.e.:

$$\mathbf{x} = \boldsymbol{\theta}\left(y, \dot{y}, \dots, y^{(p)}\right) \text{ and } u = s(y, \dot{y}, \dots, y^{(p+1)}) \quad (2)$$

where $\boldsymbol{\theta}$ and s are respectively a smooth vector field and a map (Rouchon [2008]).

2. The controlled system $\dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}) + \boldsymbol{g}(\mathbf{x})u$ is said to be flat if there exists an output $y = h(\mathbf{x})$ such that the resulting SISO system

$\dot{\mathbf{x}} = \boldsymbol{f}(\mathbf{x}) + \boldsymbol{g}(\mathbf{x})u,\ y = h(\mathbf{x})$ (3)
has relative degree n. In that case, y is called a flat output defined by the output function $h(\mathbf{x})$.

2.2 Interval tools

A real interval $[a] = [\underline{a}, \overline{a}]$ is a connected and closed subset of $\mathbb{R}$. The set of all real intervals of $\mathbb{R}$ is denoted by $\mathbb{IR}$. Real arithmetic operations are extended to intervals (see Moore [1966], Hansen [2004]). Consider an operation $o \in \{+; -; *; /\}$ and $[a]$, $[b]$ two intervals, then
$[a] o [b] = \{\boldsymbol{x}\, o\, y \mid \boldsymbol{x} \in [a], y \in [b]\}$.
The width of an interval $[a]$ is defined by $w[a] = \overline{a} - \underline{a}$ and its midpoint by $mid[a] = (\underline{a} + \overline{a})/2$.

Inclusion functions

Let $\boldsymbol{f}: \mathbb{R}^n \to \mathbb{R}^m$; the range of the function $\boldsymbol{f}$ over an interval vector **[x]** is given by:
$\boldsymbol{f}([\mathbf{x}]) = \{\boldsymbol{f}(\mathbf{x}) \mid \mathbf{x} \in [\mathbf{x}]\}$
An interval function $[\boldsymbol{f}]: \mathbb{R}^n \to \mathbb{R}^m$ is an inclusion function for $\boldsymbol{f}$ if:
$\forall [\mathbf{x}] \in \mathbb{IR}^n, \boldsymbol{f}([\mathbf{x}]) \subseteq [\boldsymbol{f}]([\mathbf{x}])$
An inclusion function of $\boldsymbol{f}$ could be obtained by replacing each occurrence of a real variable by its corresponding interval and by replacing each standard function by its interval evaluation. Such a function is called the natural inclusion function. In practice, the inclusion function is not unique and depends on the formal expression of $\boldsymbol{f}$. When the manipulated intervals are not large, the centered form could give better results than the natural one.

2.3 Constraint Satisfaction Problems (CSPs)

A constraint satisfaction problem (or CSP) is defined by a set of variables, $X_1, X_2, \dots, X_n$, and a set of constraints, $C_1, C_2, \dots, C_m$. Each variable X_i has a nonempty domain D_i of possible values (Norvig and Russell [(2010]). Each constraint C_i involves some subset of the variables and specifies the allowable combinations of values for that subset. A state of the problem is defined by an assignment of values to some or all of the variables, $\{X_i = v_i, X_j = v_j, \dots\}$. An assignment that does not violate any constraint is called a consistent or legal assignment. A complete assignment is one in which every variable is mentioned, and a solution to a CSP is a complete assignment that satisfies all the constraints. Some CSPs also require a solution that maximizes an objective function.

Constraint propagation is a way to solve CSPs and the aim of propagation techniques is to contract as much as possible the domains for the variables without losing any solution. The Waltz filtering algorithm (Waltz [1972, 1975]) popularized the technique of constraint propagation and it was initially proposed as a way to reduce the combinatory associated with line labeling of three-dimensional scenes. The Waltz filtering is more addressed to the computer science and artificial intelligence domains (Van Hentenryck [1989], Kumar [1992]) but it has also proved its efficiency in solving some of control problems. When interval uncertainties are considered, consistency methods combining interval and constraint satisfaction techniques can be used to deal with problems such as parameter/ state estimation and further the fault detection problems.

Example:

Consider the three following constraints:
$(C_1): y = x^2$
$(C_2): xy = 1$
$(C_3): y = -2x + 1$
To each variable, x and y, we associate the domain $]\infty; +\infty[$. A constraint propagation consists in projecting all constraints until equilibrium:

$$(C_1) \to y \in]-\infty; +\infty[^2 = [0; +\infty[$$
$$(C_2) \to x \in 1/[0; +\infty[=]0; +\infty[$$
$$(C_3) \to y \in [0; +\infty[\cap ((-2) *]0; +\infty[+ 1)$$
$$= [0, +\infty[\cap]-\infty; 1[= [0, 1[$$
$$x \in]0; +\infty[\cap \left(-\frac{[0; 1[}{2} + \frac{1}{2}\right)$$
$$= [0, \frac{1}{2}[$$
$$(C_1) \to y \in [0, 1[\cap [0; \frac{1}{2}[^2 = [0, \frac{1}{4}[$$
$$(C_2) \to x \in [0; \frac{1}{2}[\cap \frac{1}{[0, \frac{1}{4}[} = \emptyset$$
$$y \in [0; \frac{1}{4}[\cap \frac{1}{\emptyset} = \emptyset$$

Thus, it has been proved that no solution exists for this CSP.

In the case of flat systems, the CSP variables correspond to the states and input related to the flat output by a specific and unique map. This map defines the constraints. In order to retrieve the state and input vectors satisfying the constraints, the set inversion technique is applied through the flatness equations (Jaulin et al. [2009]).

3. CSP-BASED FAULT DETECTION

3.1 Fault detection procedure

The fault detection strategy is based on the constraint satisfaction methodology discussed in the previous section. The idea is to build a simple procedure for "model invalidation". A CSP-based observer is designed in order to estimate a set containing the input u feasible values from the faulty real measurements. Interval residuals are then applied to determine the gap between the estimated set and the expected input. The lower, respectively upper, bound of the residual corresponds to the difference between the input u estimated set lower, respectively upper, bound and the fault free model input u value. The residual is then defined by:
$r = [\underline{u}_{est} - u, \overline{u}_{est} - u]$ (4)
The consistency test is based on the comparison of the input expected value (fault-free model) and the estimated domain $[\underline{u}_{est}, \overline{u}_{est}]$. Then, if u does not belong to the latter, the fault-free model is not compatible with the measurements and we can conclude that a fault has occurred. This is equivalent to checking if:
$r = [\underline{u}_{est} - u, \overline{u}_{est} - u] \not\ni 0.$ (5)

3.2 CSP-based observer

The proposed observer is built based on the flatness property. The main steps are detailed in the following:
a. Equation (2) can be rewritten as (Jaulin et al. [2009])
$\mathbf{z} = \left(y,\ y^{(1)}, \dots, y^{(p)}\right)^T = \boldsymbol{\varphi}(\mathbf{w}) = \varphi[(x, u)^T]$ (6)
The function $\boldsymbol{\varphi}$ can be obtained by successive derivatives of the flat output with respect to time. The goal is to estimate the input u based on the expression (6) at the sampling times t_j.

b. Denote respectively by $\boldsymbol{U_j}$, $\mathbf{Z_j}$, the domains of u and $\mathbf{z}$ at t_j. Note that if no prior information about the domain of u is available, we can select $\boldsymbol{U_j} =]-\infty, +\infty[$. Thus, the input estimation method consists in computing all the values of u satisfying:

$$\begin{cases} \mathbf{z_j} = \boldsymbol{\varphi}[(\mathbf{x_j}, u_j)^T] \\ \mathbf{z_j} \in \mathbf{Z_j} \\ u_\mathbf{j} \in \boldsymbol{U_j} \end{cases} \quad (7)$$

Then, the idea is to remove parts of the search domain $\boldsymbol{U_j}$ for the model input that is inconsistent with the measured data y_j and their derivatives up to order p. In this work, the measurement derivatives are computed using HOSM differentiators (Levant [1998, 2001]). Let $f(t)$ be the signal to be differentiated and $z_0, z_1 \ldots z_n$ some estimates for the signal $f(t)$ and its derivatives. The n^{th}-order HOSM differentiator is given by:

$$\begin{cases} \dot{z}_0 = v_0, v_0 = -\alpha_0 |z_0 - f(t)|^{\frac{n}{n+1}} sign(z_0 - f(t)) + z_1 \\ \dot{z}_1 = v_1,\ v_1 = -\alpha_1 |z_1 - v_0|^{\frac{n-1}{n}} sign(z_1 - v_0) + z_2 \\ \quad \ldots \\ \dot{z}_i = v_i,\ v_i = -\alpha_i |z_i - v_{i-1}|^{\frac{n-i}{n+1}} sign(z_i - v_{i-1}) + z_{i+1} \\ \quad \ldots \\ \dot{z}_{n-1} = v_{n-1}, \\ v_{n-1} = -\alpha_{n-1} |z_{n-1} - v_{n-2}|^{\frac{1}{2}} sign(z_{n-1} - v_{n-2}) + z_n \\ \dot{z}_n = -\alpha_n sign(z_n - v_{n-1}) \end{cases} \quad (8)$$

Coming back to the problem at hand, the main assumption in this paper is that the measurement error e is bounded with a prior known bound $\overline{e}$. Thus, y domain is given by:

$$y \in [y_m - \overline{e}, y_m + \overline{e}] \quad (9)$$

where y_m is the measurement. The derivatives are estimated via the n^{th}-order HOSM differentiator (8). It has been proved in (Levant [2001]) that the i^{th} derivative estimate accuracy is proportional to $acc = \overline{e}^{(\frac{n+1-i}{n+1})}, i = 0, \ldots, n$ when the Lipschitz constant of the n^{th} derivative of the clear-off-noise signal is bounded by a certain constant. Hence, the derivative domain is: $y^{(i)} \in [y_{est}^{(i)} - acc, y_{est}^{(i)} + acc]$ where $y_{est}^{(i)}$ is the derivative estimate.

The following CSP algorithm sums up the constraint satisfaction methodology.

Algorithm CSP Estimator (Inputs: $y(t_i)$, i=1..N, **I.D***: $[u_0]$)

1. Flatness modelling (eq. 6)
2. For i=1 to N do,
   Estimate the derivatives $y^{(q)}$, q=1,2..p+1 (eq. 8)
   Estimate the bound *acc* and construct the domains of $y(t_i)$ and $y^{(i)}(t_i)$
   Solve CSP to obtain $[u(t_i)]$ (eq. 7)

* I.D: Initial search Domain

In the following paragraph, two numerical examples are presented to illustrate the efficiency of the proposed approach for detection of sensor and actuator faults. Note that both examples are under feedback control and the input is constructed from the state feedback.

4. FAULT DETECTION PERFORMANCE

4.1 Sensor faults

Equation (2) shows that the input u only depends on the flat output and its derivatives up to an order $(p+1)$. A sensor fault appearing on the measurements will cause erroneous derivative computation and finally a wrong estimate will be calculated for the input u. Comparing both estimated input set and the expected input u, an empty intersection would denote the occurrence of a fault. Consider the following system:

$$\begin{cases} \dot{x}_1 = e^{x_2} u \\ \dot{x}_2 = x_1 + e^{x_2} u \\ \dot{x}_3 = x_1 - x_2 \\ y = x_3, z = x_2 \end{cases} \quad (10)$$

y and z are the measured outputs.

It is easy to prove that:

$$\begin{cases} x_1 = -\ddot{x}_3 = -\ddot{y} \\ x_2 = -(\ddot{x}_3 + \dot{x}_3) = -(\ddot{y} + \dot{y}) \\ x_3 = y \\ u = \dfrac{\dot{x}_1}{e^{x_2}} = -\dfrac{\dddot{x}_3}{e^{-(\ddot{x}_3 + \dot{x}_3)}} = -\dfrac{\dddot{y}}{e^{-(\ddot{y}+\dot{y})}} \end{cases}$$

Thus, the system (10) is flat and y is the flat output. To illustrate the consequences of a sensor additive fault on the input u estimation, we can write the contaminated measurement as:

$$y_f = y + fault. \quad (11)$$

The most common sensor error fault is an offset bias. When the output signal slowly changes independent of the measured property, it can be modelled as drift. Finally, a sensor can be subject to an abnormal external noise.

Bias: Adding a bias to the measurement y at the instant t_b leads theoretically to the following input estimation:

$$u = \frac{\frac{d^3}{dt^3}(y+b)}{e^{-\left[\frac{d^2}{dt^2}(y+b) + \frac{d}{dt}(y+b)\right]}}$$

$$u = \frac{[\dddot{y} + \dddot{b}]}{e^{-([\ddot{y}+\ddot{b}]+[\dot{y}+\dot{b}])}} \quad (12)$$

The bias b is assumed to be a step function appearing at $t = 2s$. Using the erroneous measurement $y_f = y + 0.25$, the last equation becomes:

$$u_{est} = \frac{\dot{x}_1}{e^{x_2}} = -\frac{\dddot{x}_3}{e^{-(\ddot{x}_3 + \dot{x}_3)}} = -\frac{\dddot{y}_f}{e^{-(\ddot{y}_f + \dot{y}_f)}} \quad (14)$$

Moreover, we suppose that the measurement y^m (contaminated or not) belongs to the interval $[y^m - \overline{e}, y^m + e]$, where e=0.004 is the *a priori* known measurement error. The initial search domain for the input is taken as: $[u_0] = [-2; 4]$.

The input estimation gives the following set (black) and the expected (dashed line) as shown in figure 1.a. The upper and lower bounds of the residual are depicted in figure 1.b.

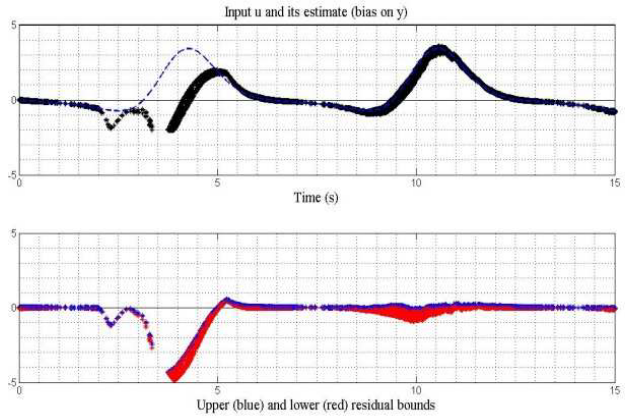


Figure 1: Input u estimation and residual in presence of bias fault.

The simulation results indicate a very short detection time. The residual upper and lower bounds describe an interval containing zero except between $t = 2s$ and $t = 5s$. The transient behaviour (the effect of fault does not persist beyond $t = 5s$) is due to the fact that the simulation is done in a closed-loop feedback configuration.

Drift: A slow ramp function is used to illustrate the drift d on the measurement y. It is given by:

$$\begin{cases} d = 0 \; \forall \; t < t_d = 2s \\ d = 0.2t \end{cases} \quad (15)$$

and $y_f = y + d$. The same assumptions as above are made on the measurement enclosure and the way to obtain the derivatives.

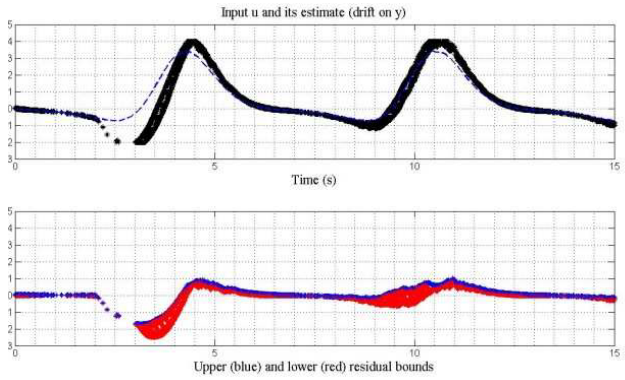


Figure 2: Input u estimation and residual in presence of drift.

Note that the behavior of the residual is very similar to the previous case. The fault effect is only clearly visible between $t = 2s$ and $t = 5.5s$. However, a deeper analysis reveals some interesting features which can be explored further for fault isolation, for example one can see that in this case the fault effect persist during a bigger time range.

Random noise: A random noise with a variance $V = 0.5$ is here used to simulate a sensor fault. The observer is still very sensitive and the detection time is very short (see figure 3).

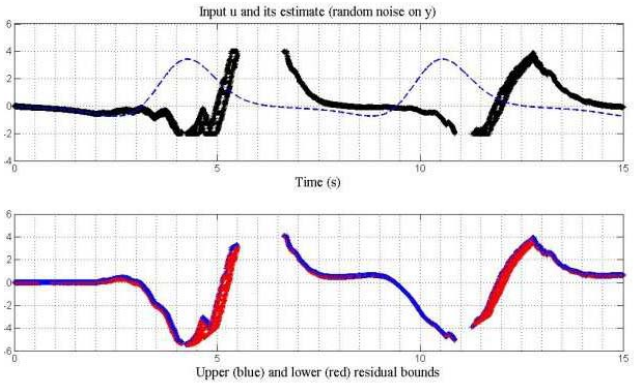


Figure 3: Input u estimation and residual in presence of random noise.

Besides the $3s$-gap between the expected input value (dashed line) and its estimated domain, one can notice that in this case the fault effect does persist on the residual. To confirm this observation, a white noise has been also simulated (instant $t_{wn} = 2s$).

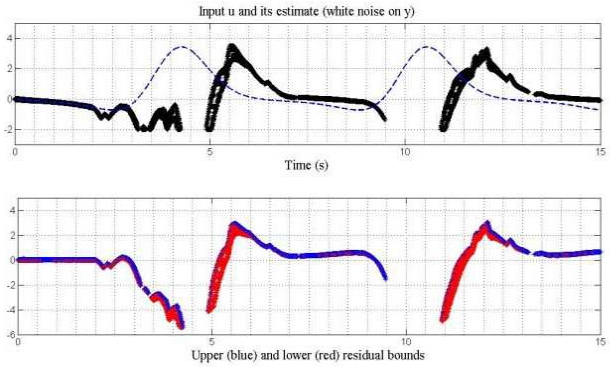


Figure 4: Input u estimation and residual in presence of white noise.

As it can be seen from figure 4, the same conclusions as in the latter case can be made. Note that the important blank space (between $t = 4.5s$ and $t = 5s$, $t = 9s$ and $t = 11s$) in the estimate and thus the residual corresponds to the non-solution part in the initial input search domain which means that the domain of the input u does not contain any admissible value. Note also that the pseudo-oscillations before the first sinusoidal form in the estimate are more important in the white noise case.

4.2 Actuator faults

In this section, actuator faults are considered using the previous example and methodology. Additive faults on the input u are simulated and an estimation of the input from the resulting measurements is performed.

Intermittent fault signal on the input

Consider the following signal μ:

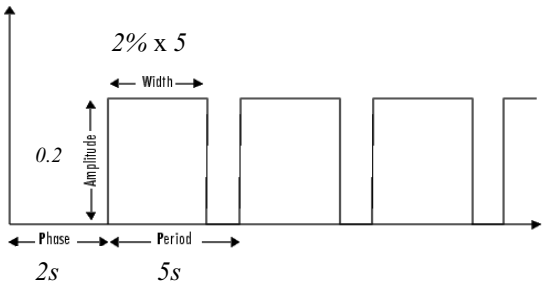


Figure 5: Pulse signal description

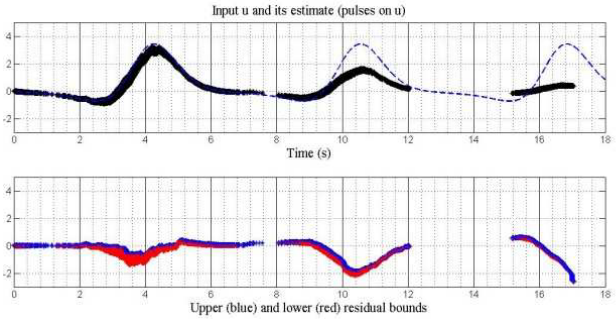


Figure 6: Input u estimation and residual in presence of intermittent (pulses) signal.

The fault is detected at $t = 3s$, i.e with a $1s$-delay time, however the residual signal form slightly changes from $t = 2s$. The effects of the second pulse (occurring at $t = 7s$) are detectable from $t = 7.5s$. Note the estimated signal amplitude decrease with time, i.e the deviation of the estimated set from the expected input u value becomes more evident.

White noise signal on the input

The same white noise signal properties as for the sensor fault are used in this test.

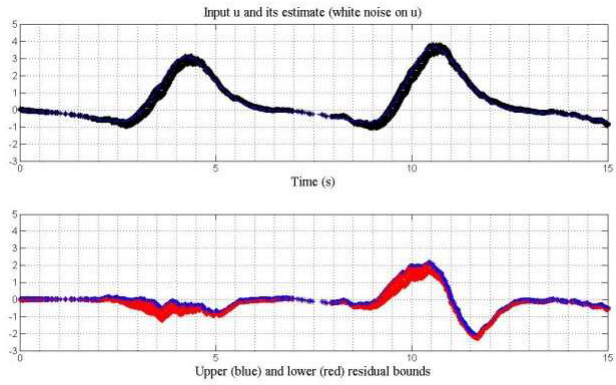


Figure 7: Input u estimation and residuals in presence of white noise.

Unlike the case where the white noise disturbs the output y, there is no phase shift between the expected input signal and its estimated set. Moreover, the delay in the detection is a bit more important since it takes $0.5s$ before residual signal form changes and quite $1s$ ($t = 3s$) before the interval residual clearly denotes a fault occurrence.

5. CONCLUSION

A constraint satisfaction-based technique has been proposed in this paper for robust input estimation of flat systems. Based on this estimation, a robust fault detection strategy has been proposed by applying a consistency test on the residuals generated from the difference between the input estimated set (faulty case) and its expected value (fault-free case). The robustness of the proposed scheme has been demonstrated through simulation examples using both sensor and actuator faults. The technique is appropriate detection of incipient faults when the permissible time window for detection is narrow. Further investigations are necessary to analyze the isolation capability and to study formally the effect of feedback control on fault diagnosis performance. This is a topic of our current research.

REFERENCES

Agrawal, S. K. and Sira-Ramirez, H. (2004). *Differentially flat systems*. Crc Press.

Benhamou, F. and Granvilliers, L. (2006). Continuous and interval constraints. In P. van Beek F. Rossi and T.Walsh. *Handbook of constraint programming*, 571-604. *Elsevier*.

Ding, S. X. (2008). *Model-based fault diagnosis techniques: design schemes, algorithms and tools*. Springer-Verlag New York, LLC

Fliess, M., Lévine, J., Martin, P. and Rouchon, P. (1992). *Sur les systèmes non linéaires différentiellement plats. Elsevier*, Paris, FR.

Frank, P.M. (1992). Principles of model-based fault detection. In *Annual Review in Automatic Programming* 17, 213-220. Artificial Intelligence in Real-time Control, IFAC/IFIP/IMACS Symposium.

Goldsztejn, A. (2006). A branch and prune algorithm for the approximation of non-linear AE-solution sets. *Proceedings of the 2006 ACM Symposium on Applied Computing*. Dijon, FRANCE.

Hansen, R. E. (2004). *Global optimization using interval analysis, second edition*. CRC.

Jaulin, L. Le Bars F., Sliwka J., Xiao K., (2009). Combining flatness with of interval analysis for state estimation. Journée MEA Paris.

Jaulin, L. (2009). Interval contractors and their applications. Ecole JN-MACS.

Khan, A., Abou, S. C. and Sepehri, N. (2005). Nonlinear observer-based fault detection technique for electro-hydraulic servo-positioning systems. *Mechatronics* 15(9), 1037-1059.

Kobi, A., Nowakowski, S. and Ragot, J. (1994). Fault detection isolation and control reconfiguration. *Mathematics and Computers in Simulation* 37, 111-117.

Kumar, S., Sinha, S., Kojima, T. and Yoshida, H. (2001). Development of parameter-based fault detection and diagnosis technique for energy efficient building management system. *Energy Conversion and Management* 42(7), 833-854.

Kumar, V. (1992). Algorithms for constraint satisfaction problems: a survey. In *AI Magazine* 13(1), 32-44.

Levant, A., (2001). Higher order sliding modes and arbitrary order exact robust differentiation. Proceedings of the European Control Conference.

Levant, A. (1998). Robust exact differentiation via sliding mode technique. *Automatica* 34(3), 379-384.

Louembet, C., Cazaurang, F.and Zolghadri, A. (2010). Motion planning using positiveB-splines : an LMI approach. *Automatica* 46(8), 1305-1309.

Martin, Ph., Murray, R. M. and Rouchon, P. (1997). Flat systems. In G. Bastin and M. Geverts, editors, Plenary lectures and mini-courses, 211–264. 4th European Control Conference, Brussels/ Belgium.

Medvedev, A. (1995). Fault detection and isolation by a continuous parity space method. *Automatica* 31(7), 1039-1044.

Moore, R. E. (1966). *Interval analysis*. Prentice Hall, Englewood Cliffs, NJ, USA.

Martin, Ph., Murray, R. M. and Rouchon, P. (2001). Flat systems, equivalence and feedback. In Banos A., Lamnabhi-Lagarrigue F. and Montoya F. J. editors, *Advances in the control of nonlinear systems*, Lecture Notes in Control and Inform. Sci., 3–32. Springer-Verlag.

Neumaier, A. (2004). Complete Search in Continuous Global Optimization and Constraint Satisfaction. *Acta Numerica*. Cambridge: University Press.

Norvig, P. and Russell, S. (2010). *Artificial Intelligence: A Modern Approach*, 3rd edition. Prentice Hall.

Rouchon, P. (2008). Systèmes différentiellement plats. JNCF, CIRM.

Sproesser, T. and Gissinger, G. L. (1992). A method for fault detection using parameter and state estimation. In *Annual Review in Automatic Programming* 17, 41-247. Artificial Intelligence in Real-time Control, IFAC/IFIP/IMACS Symposium.

Van Hentenryck P., (1989). *Constraint satisfaction in logic programming*. MIT Press.

Waltz, D.L. (1972). Generating semantic descriptions from drawings of scenes with shadows. Technical Report, AI-TR-271, MIT Artificial Intelligence Laboratory, Cambridge, MA.

Waltz, D.L. (1975). Understanding line drawings of scenes with shadows. In *The Psychology of Computer Vision,* McGraw-Hill, pp.19-91.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Communication sequence design in networked control systems with communication constraints: a graphic approach

Sinuhe Martinez-Martinez, Hossein Hashemi-Nejad, Dominique Sauter *

* *Research Center on Automatic Control of Nancy, France (e-mail: hossein.hashemi@cran.uhp-nancy.fr).*

Abstract: This paper proposes a graphical strategy for finding all communication sequences that ensure reachability of a Networked Control system (NCS) in which a linear time invariant (LTI) plant communicates with a controller over a shared medium. The medium supports a limited number of simultaneous connection between controller and actuators. The proposed method is based on a graph-theoretic approach and it needs only the knowledge of the systems structure.

Keywords: Networked Control Systems, Medium Access Constraint, Reachability, Graph Theory, Maximal Matching.

1. INTRODUCTION

In classical control theory a perfect information exchange is assumed. But the progress in communication, control and real time computation has enabled the development of large scale systems which sensors and actuators exchange information with feedback controller through a shared network. Control systems having this configuration have been termed Networked Control Systems (NCSs). Introduction of networks in control loop adds some limitation in data exchange and it brings new problems and challenges such as networked induced delay, packet dropout and constraints in access to the medium. As consequence, the classical control theories must be revised to be adapted in NCSs. For instance stabilization problem of NCS which is studied in [Shousong and Qixin (2003); Halevi and Ray (1988)]. The networked induced delays which may degrade the performance of closed loop system were investigated in [Tipsuwan and Chow (2003); Yang et al. (2006)]. Important surveys about recent results in NCSs are given in [Hristu-Varsakelis and Levine (2005); Hespaha et al. (2007)].

Access constraints are one of the major obstacles in control system design. It occurs when capacity of communication medium for providing simultaneous medium access channels for its user is limited. As a consequence only limited number of sensors or actuators is allowed to communicate with controller at each time instant. Moreover, if a Fault detection and isolation (FDI) module exists, its connexion to the network has not access to measurement of all sensors simultaneously. Recently there was a number research activity in this field. As an example, an LQG design method for NCS which are subject to medium access constraints was presented in [Zhang and Hristu-Varsakelis (2005)]. Problem of fault detection (FD) with communication constraints in linear systems [Wang et al. (2009); Zhang and Ding (2006)] and in non-linear systems [Mao et al. (2009)] were considered. The reachability and observability of an NCS with limited communication was studied in [Zhang and Hristu-Varsakelis (2006)].

Basic properties such as reachability and observability are important if we are interested in the design of controllers or a FD module for instance. Moreover, it is well known that these basic properties depend strongly on the structure of the system, see for example [Lin (1974); Reinschke (1988)]. A study on observability for LTI structured systems is carried out in [Commault et al. (2005)] and in [Boukhobza et al. (2007)] in the context of NCS systems. A complete survey on structural methods can be found in [Dion et al. (2003)]. From the structural view point, the problem of the existence of sequences which preserve the observability/reachability of an NCS with limited communication is treated in [Ionete and Cela (2006)]. The structural approach is a powerful tool for the systems analysis which main advantage is the low complexity of its algorithms when combined with the graphical approach (see [Martinez-Martinez et al. (2007)]).

This paper presents a graphical strategy to find all communication sequence for given networked access constraint on input channel that preserve reachability of the system. Comparing with previous works that studied design of communication sequence, this method is simpler and complex mathematical computation is not necessary. Also, it enables us a) to verify existence of a communication sequence that preserve reachability of system for given channel limit ω_ρ . b) Considering channel limitation, Find all communication sequences which guarantee system's reachability.

The rest of this paper is structured as follows. In section 2, constraints of communication and communication se-

quences in model of system are taken into account and an extended linear time varying (LTV) system is presented. Graphical representation of linear systems is studied in section 3 then in section 4, a graphical method for finding all communication sequence that preserve reachability of extended system with respect to communication constraints is proposed.

2. PROBLEM STATEMENT

Suppose that the model of plant connected to the network with communication constraints in figure 1 is described by following discrete-time LTI system:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \tag{1}$$

Where A,B and C are matrices of appropriate dimensions, $x(k) \in \mathcal{R}^n$ $u(k) \in \mathcal{R}^m$ and $y(k) \in \mathcal{R}^p$ are the state, input and the output of the system.

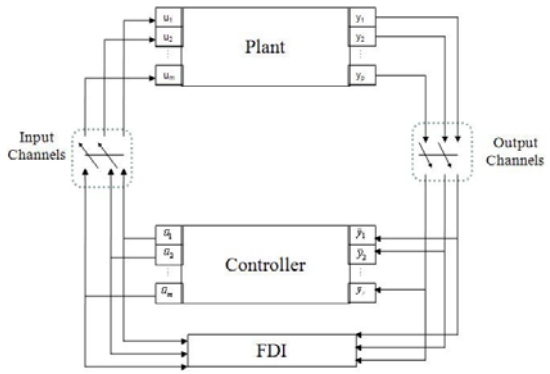


Fig. 1. *NCS with communication constraints*

In the figure 1 output (or input) channel is illustrated. It is referred to communication link that enables data transmission from sensors to controller (or controller to actuators). To focus on the effects of communication constraints, the following statements are assumed:

- transmissions are instantaneous,
- the communication channels are noiseless,
- there are not packet dropouts.

Due to communication constraints in input channels, the shared communication medium can simultaneously provide ω_ρ inputs with $1 \leq \omega_\rho \leq m$. In other words, at each sampling instant k, only ω_ρ of the actuators are allowed to access to the network. Only the commands of ω_ρ actuators are available for controller and for the system input respectively.

For all $i = 1, \ldots, p$, a binary-valued function $\rho_i(k)$ is defined as the medium access status for actuator i at sample time k. Then if the *i-th* actuator has access to the network at instant k, $\rho_i(k) = 1$, otherwise $\rho_i(k) = 0$. Whenever an actuator j loses its access to the communication medium, the control signal generated at the controller for the actuator will be unavailable and hence $u_j = 0$ for the plant until actuator j recovers accessibility.

The instantaneous medium access status of m actuators at sample time k is hence represented by a m-to-ω_ρ communication sequence [Zhang and Hristu-Varsakelis (2005)]

$$\boldsymbol{\rho}(k) = [\rho_1(k), \ldots, \rho_m(k)]^T \tag{2}$$

and then in its matrix form $\mathcal{M}_\rho(k)$ defined by

$$\mathcal{M}_\rho(k) \triangleq \text{diag}(\boldsymbol{\rho}_i(k)) \tag{3}$$

Let $\bar{u}(k)$ the available input actually used by the system whereas $u(k)$ be the signal generated by the controller. We can state

$$\bar{u}(k) = B\bar{\mathcal{M}}_\rho(k)u(k),\ \bar{\mathcal{M}}_\rho(k) \in \mathcal{R}^{\omega_\rho \times p} \tag{4}$$

Where $\bar{\mathcal{M}}_\rho$ is obtained by deleting the zero rows of $\mathcal{M}_\rho$

Therefore "from the controller point of view " NCS will behaves as a time-varying system with input $\bar{u}$. The system is represented as follows

$$\begin{aligned} x(k+1) &= Ax(k) + \bar{B}(k)u(k);\ \bar{B}(k) = B\bar{\mathcal{M}}_\rho(k) \\ y(k) &= Cx(k) \end{aligned} \tag{5}$$

Equation (5) incorporates the dynamic of the plant together with access of communication medium and we call it extended plant.

Then, dynamic of the extended plant (5) depends on communication policy $\boldsymbol{\rho}(k)$. The reachability is an important property of system (1) which may be lost when communication constraint are imposed. For verifying acceptability of a communication sequence, we can look for sequences which preserve reachability of underlying system (5). The choice of a sequence is not obvious and requires a specific knowledge of the system. In order to understand these ideas the following example is analysed in which reachability of the system may be lost depending on the sequence selection.

Example 1. Let the following matrices represent the linear model of a plant connected to a network with communication constraints:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, B = [b_1, b_2] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{6}$$

Considering the reachability matrix for a the extended discrete-time linear system 6 given by

$$\begin{aligned} \mathcal{R}(0, k_f) = [&A^{k_f - 1}[b_1, b_2]\mathcal{M}_\rho(0), \ldots, \\ &[b_1, b_2]B\mathcal{M}_\rho(k_f - 2), [b_1, b_2]\mathcal{M}_\rho(k_f - 1)] \end{aligned} \tag{7}$$

Notice that the function of matrix $\mathcal{M}_\rho$ is to select the inputs which have access to the communication medium. The interest is to find how the communication medium has access to the controller's outputs in order to preserve the reachability of the overall system and then the full rank of the reachability matrix $\mathcal{R}(0, k_f)$ for a given k_f. The system is originally reachable considering that no communication constraint exists. Now, with a communication restriction fixed to one channel let us suppose that a communication sequence is chosen as the the 2-periodic sequence

$$\{\boldsymbol{\rho}(0), \boldsymbol{\rho}(1), \ldots\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ldots \right\}$$

then the matrix $\mathcal{R}(0, k_f)$ *contains either* b_1 *or* b_2 *depending on whether* k_f *is an odd or even number, and it loses the rank. But, if the* 1*-periodic sequence*

$$\{\boldsymbol{\rho}(0), \boldsymbol{\rho}(1), \dots\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \dots \right\}$$

is chosen, the reachability matrix $\mathcal{R}(0, k_f)$ *will remain a full rank matrix and the reachability will be preserved in the extended plant.*

The interest of *Example 1* is twofold. First of all it shows that the choice of a sequence is not trivial and that it depends strongly on the structure of the system. Secondly, it presents a system with a certain redundancy which allows it to work with only one of its inputs. It is worth noting that the latter is not always the case.

Reachability of the extended plant is given by the following definition.

Definition 1. The extended plant 5 is reachable on $[k_0, k_f]$ if given any x_f, there exists an input $\bar{u}(k)$ that steers 5 from $x(k_0) = 0$ to $x(k_f) = x_f$.

Existence of sequences has already treated in precedent works, we concentrate in finding periodic communication sequences that ensure reachability of the NCS. In the general case, a discrete-time communication sequence $\boldsymbol{\eta}(\cdot)$ is called T-periodic if $\boldsymbol{\eta}(k) = \boldsymbol{\eta}(k + T)$ for all k. If a communication sequence exists for a given value of communication constraint ω_η, it is possible that such communication sequence not be unique. In such case two questions arise:

- How can we find all communication sequences with period T such that each of them preserves the reachability of the extended plant (5)?
- What is the minimal size of ω_ρ which guarantees the preservation of the reachability of the extended plant (5)?

Before giving answer to these questions, we introduce the graphic approach which will be useful to find the communication sequences.

3. GRAPHIC REPRESENTATION OF LINEAR SYSTEMS

In this part we will consider the graphical representation of a linear system and then its generalization to the extended system.

3.1 Directed graph associated to a linear system

In this part we will present how it is possible to associate a directed graph to a linear discrete system. As the main interest is upon the input side we do not consider outputs in the sequel of this communication. Consider the following linear discrete-time system:

$$\Sigma_\Lambda \{x(k+1) = Ax(k) + Bu(k) \tag{8}$$

We will refere to this system as structured because we concentrate only in its structure. That is to say, we consider matrices A and B having elements either fixed to zero or free (non-zero) parameters noted λ_i. These parameters form a vector $\Lambda = (\lambda_1, \dots, \lambda_h)^T \in \mathcal{R}^h$. We say that a property is true generically if it is true for almost all parameters values $\Lambda \in \mathcal{R}^h$. For "almost all" is to be understood as for all parameters values except those in some proper algebraic variety in the parameter space.

A digraph can be used to represent structured linear system (Σ_Λ). The digraph associated to (Σ_Λ) is noted $\mathcal{G}(\Sigma_\Lambda)$ and is constituted by a vertex set $\mathcal{V}$ and an edge set $\mathcal{E}$ *i.e.* $\mathcal{G}(\Sigma_\Lambda) = (\mathcal{V}, \mathcal{E})$. The vertices are associated to the state and the inputs of (Σ_Λ) and the edges represent links between these variables. More precisely, $\mathcal{V} = \mathbf{X} \cup \mathbf{U}$. Hence, $\mathcal{V}$ consists of $n + m$ vertices.

The edge set is $\mathcal{E} = A^\lambda$-edges $\cup$ B^λ-edges, where
A^λ-edges $= \{(\mathbf{x_j}, \mathbf{x_i}) \mid A(i,j) \neq 0\}$,
B^λ-edges $= \{(\mathbf{u_j}, \mathbf{x_i}) \mid B(i,j) \neq 0\}$,
Here $M^\lambda(i,j)$ is the $(i,j)th$ element of matrix M^λ and $(\mathbf{v_1}, \mathbf{v_2})$ denotes a directed edge from vertex $\mathbf{v_1} \in \mathcal{V}$ to vertex $\mathbf{v_2} \in \mathcal{V}$.

In order to understand the ideas developed in next section, we introduce some important definitions in the context of graph approach for structured systems.
Considering the associate graph $\mathcal{G}(\Sigma_\Lambda) = (\mathcal{V}, \mathcal{E})$. For an edge $e = (\mathbf{v_i}, \mathbf{v_f}) \in \mathcal{E}$, $\mathbf{v_i}$ (respectively $\mathbf{v_f}$) is the begin (respectively the end) vertex of e.

We denote path $\mathbf{P}$ containing vertices $\mathbf{v_{r_0}}$, $\mathbf{v_{r_1}}$, $\dots$, $\mathbf{v_{r_i}}$ by $\mathbf{P} = \mathbf{v_{r_0}} \to \mathbf{v_{r_1}} \to \dots \to \mathbf{v_{r_i}}$. The pair $(\mathbf{v_{r_j}}, \mathbf{v_{r_{j+1}}}) \in \mathcal{E}$ for $j = 0, 1 \dots, i-1$ if there is an integer l and vertices $\mathbf{v_0}, \mathbf{v_1}, \dots, \mathbf{v_l} \in \mathcal{V}$ such that $(\mathbf{v_{i-1}}, \mathbf{v_i}) \in \mathcal{E}$ for $i = 1, 2, \dots, l$. Then, path $\mathbf{P}$ is of length l. When $\mathbf{v_0} = \mathbf{v_i}$, $\mathbf{P}$ is a *cycle*. Some paths are called disjoint if they have no common vertex. A path P is a $\mathbf{U}$-*rooted* path if its begin vertex is an element of $\mathbf{U}$. A $\mathbf{U}$-*rooted* path family consist of disjoint simple $\mathbf{U}$-*rooted* paths. If such a family contains a path or a cycle which covers a vertex $\mathbf{v}$ it is said to cover such vertex. A system Σ_Λ is input connected if in its associated graph $\mathcal{G}(\Sigma_\Lambda)$ for every state vertex $\mathbf{x}_i$ there exists a direct path from the input set $\mathbf{U}$.

For the sake of clarity we present an example of the graphic representation of an structured system and we show some of the concepts developed in this section.

Example 2. Consider an structured linear system given by the following matrices:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 \\ 0 & 0 & 0 & \lambda_4 & 0 \end{bmatrix}, \text{ and } B = \begin{bmatrix} \lambda_5 & 0 \\ 0 & \lambda_6 \\ 0 & 0 \\ 0 & 0 \\ 0 & \lambda_7 \end{bmatrix} \tag{9}$$

The directed graph associated to the structured system (9) is presented in figure (2)

One can notice that the graphic representation is rather intuitive. The vertices are associated to the states and inputs and the edges represent links between them. For example, the two simple paths starting from vertices $\mathbf{u}_2$

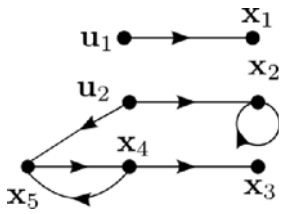


Fig. 2. *Directed graph associated to the structured system (9)*

and $\mathbf{x}_4$ *and arriving to vertex* $\mathbf{x}_5$ *represent the equation* $x_5(k+1) = \lambda_4 x_4(k) + \lambda_7 u_2(k)$.
We can also verify that it is possible to find a maximum of 2 $\mathbf{U}$*-rooted disjoint paths and that the system is input connected. We notice that such* $\mathbf{U}$*-rooted disjoint paths cannot cover completely the state. These two remarks will be important later.*

Let us recall that the main interest in this work is to find (all) T-periodic communications which preserve the reachability of the extended plant. For this aim, the method must allow us to explore all the possibilities of choice for the inputs *i.e.* $\mathcal{M}_\rho$ at every step k select every input. In other words, we deal with the nominal system (8).

Exploring all the possibilities for $\mathcal{M}_\rho$ means that availability of each input must be verified at each step k. In fact, if we analyse the evolution of an equation in 9 it follows that at every sampling time we search all the $\mathbf{U}$-*topped* path families with length equals to the sampleing time k. Indeed, if we consider that it is possible to store last information, it is worth noting that actually what is of real interest is the beginning and final vertex of such paths at each step k. This brings us to the idea of create a bipartite graph which relates the beginning vertices to the corresponding ending vertices at every sampling time. This idea help us to capture graphically the dynamic of the instantaneous medium access.

3.2 Dynamic bipartite graph association to a directed graph

In this section we introduce the bipartite graph which will be useful in the determination of the access sequences.

In order to capture graphically the dynamic of the instantaneous medium access for each step k, a particular graph will be associated to the structured system Σ_Λ called dynamic bipartite graph. The dynamic bipartite graph associated to the structured system Σ_Λ is noted $\mathcal{B}_k(\Sigma_\Lambda) = (\mathcal{U}, \mathcal{X}; \mathcal{W}_k)$. The vertex set $\mathcal{U}$ is associated to the inputs and the vertex set $\mathcal{X}$ is associated to the states. The edge set $\mathcal{W}_k$ is define as follows:

$$\begin{aligned}\mathcal{W}_k = \{&\mathbf{W}_{1,1}, \ldots, \mathbf{W}_{i,k}\}| \\ &\mathbf{W}_{i,k} = (\mathbf{u}_{i,k}, \mathbf{x}_j), \text{if there exist a path in } \mathcal{G}(\Sigma_\Lambda) \\ &\text{of lenght } k \text{ between } \mathbf{u}_i \text{ and } \mathbf{x}_j \ \}\end{aligned}$$

The index k must be fixed before constructing the dynamic bipartite graph. Regarding the reachability issue we can generate the following edge subsets for the digraph of the figure 2 for $k = 4$. The edges were grouped into different subset for the sake of clarity.

$$\begin{aligned}\{\mathbf{W}_{1,1}; \mathbf{W}_{2,1}\} &= \{(\mathbf{u}_{1,1}, \mathbf{x}_1); (\mathbf{u}_{2,1}, \mathbf{x}_2), (\mathbf{u}_{2,1}, \mathbf{x}_5)\}; \\ \mathbf{W}_{2,2} &= \{(\mathbf{u}_{2,2}, \mathbf{x}_2), (\mathbf{u}_{2,2}, \mathbf{x}_4)\}; \\ \mathbf{W}_{2,3} &= \{(\mathbf{u}_{2,3}, \mathbf{x}_2), (\mathbf{u}_{2,3}, \mathbf{x}_5), (\mathbf{u}_{2,3}\mathbf{x}_3)\}; \\ \mathbf{W}_{2,4} &= \{(\mathbf{u}_{2,4}, \mathbf{x}_2), (\mathbf{u}_{2,4}, \mathbf{x}_4)\}.\end{aligned}$$

For this kind of bipartite graph attention must be paid to those edges having the same vertices and belonging to the same edge subset $\mathbf{W}_{i,k}$. The dynamic bipartite graph generated with the edges subsets calculated above is depicted in figure 3 for $k = 4$.

In the case of reachability

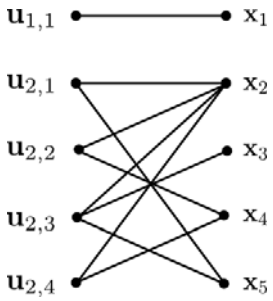


Fig. 3. $\mathcal{B}_4$ *dynamic bipartite graph associated to the structured system (9)*

For this bipartite graph some definitions must be introduced in order to understand the ideas developed in following sections.

- A *matching* is an edge set $M \subseteq \mathcal{W}$ such that the edges in M are disjoint.
- A ω_ρ*-matching* is a matching taking at most ω_ρ disjoint edges in each edge subset $\mathbf{W}_{i,j}$ for $i = 1, \ldots, p$ and j fixed.
- The cardinality of a matching M is the number of edges in it,
- In case $|\mathcal{X}| = |\mathcal{U}|$ a ω_ρ-matching that covers all the vertices is a *complete* ω_ρ-matching.
- A ω_ρ-matching M is maximal if it has a maximum cardinality.
- A *vertex sequence* is a vertex subset $\mathcal{S} \subseteq \mathcal{Z}$ such that $\mathcal{S} = [s_1, \ldots, s_j]$ where

$$s_j = \left\{ \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_i \end{bmatrix}_j \mid \mathbf{u}_i \in \mathbf{W}_{i,j} \subseteq M \right\}$$

  for $j = 1, \ldots, k$. Notice that only one sequence can be generated by a maximal ω_ρ-matching

thus, in figure 3 ($k = 4$) for example, edges $(\mathbf{u}_{2,2}, \mathbf{x}_2) \in \mathbf{W}_{2,2}$ and $(\mathbf{u}_{2,3}, \mathbf{x}_2) \in \mathbf{W}_{2,3}$ are not disjoint. On the contrary edges $(\mathbf{u}_{2,1}, \mathbf{x}_2) \in \mathbf{W}_{2,1}$ and $(\mathbf{x}_2, \mathbf{y}_{2,1}) \in \mathbf{W}_{i,1}$ are disjoint. Moreover, a maximal 1-matching, (only one communication channel at a time, $\omega_\rho = 1$), may be constituted of the following edges

$$\begin{aligned}&\{(\mathbf{u}_{1,1}, \mathbf{x}_1), (\mathbf{u}_{2,1}, \mathbf{x}_5), (\mathbf{u}_{2,2}, \mathbf{x}_2), \\ &(\mathbf{u}_{2,3}, \mathbf{x}_3), (\mathbf{u}_{2,4}, \mathbf{x}_4)\}\end{aligned}$$

The vertex sequence generated by this maximal 1-matching is

$$\mathcal{S} = \left\{ \begin{bmatrix} \mathbf{u}_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix} \right\}$$

Moreover, there exists a close relation between maximal ω_ρ-matching in a dynamic bipartite graph and the maximal number of disjoint paths in a directed graph as it is stated in the next Lemma (this relation is also proved for a maximal matching in a normal bipartite graph, see [Murota (1987)])

Lemma 3. Let the system Σ_Λ be the linear structured system defined by (8) with its associated directed graph $\mathcal{G}(\Sigma_\Lambda)$ and dynamic bipartite graph $\mathcal{B}_k(\Sigma_\Lambda)$. Following statements are equivalents:

- there exists a family of disjoint **U**-*rooted* paths and a cycle family covering all the state vertex set in the associated directed graph $\mathcal{G}(\Sigma_\Lambda)$,
- There exists a maximal ω_ρ-matching of size n with $\omega_\rho = m$ in $\mathcal{B}_k(\Sigma_\Lambda)$ for some $k \neq 0$.

Proof. Suppose that there exists a complete matching M on $\mathcal{B}_k$ for some k. Then the cardinality of the matching M is equal to the number of vertices it covers, n. For each $\mathbf{x}_j \in \mathcal{X}$ $(1 \leq j \leq n)$, there is a unique sequence of disjoint edges

$$(\mathbf{u}_{i_1,1}, \mathbf{x}_{j_1}), (\mathbf{u}_{i_1,2}, \mathbf{x}_{j_2}), \dots (\mathbf{u}_{i_1,l}, \mathbf{x}_{j_p}), \dots, (\mathbf{u}_{i_1,k}, \mathbf{x}_{j_q})$$

contained in the matching which form a disjoint path $P = \mathbf{u}_{i_1} \to \mathbf{x}_{j_1} \to \cdots \to \mathbf{x}_{j_p}$, and a cycle family $P_c = \mathbf{x}_{j_{p+1}} \to \cdots \to \mathbf{x}_{j_q} \to \mathbf{x}_{j_{p+1}}$ in the associated directed graph $\mathcal{G}$. Thus a complete matching on $\mathcal{B}_k$ determines a family of disjoint **U**-*rooted* paths $P_j = \mathbf{u}_{i_1} \to \mathbf{x}_{j_1} \to \cdots \to \mathbf{x}_{j_p}$ and cycle families $P_{c,j} = \mathbf{x}_{j_{p+1}} \to \cdots \to \mathbf{x}_{j_q} \to \mathbf{x}_{j_{p+1}}$ for $j = 1, \dots, n$ and $i = 1, \dots, m$ on $\mathcal{G}$

Conversely, suppose that there exists a family of disjoint **U**-*rooted* paths covering all the state vertices on $\mathcal{G}$. Then, by the definitions given above concerning the disjoint edges and the construction of the edge set $\mathcal{W}$ on $\mathcal{B}_k$, it is possible to construct a sequences of disjoint edges which form a complete matching on $\mathcal{B}_k$ for some $k \neq 0$.

△

4. SEQUENCES WHICH PRESERVE REACHABILITY

In this section we give a graphical method to search the communication sequences allowing the extended plant to preserve reachability along a period of time T. We recall the graphic conditions to guarantee reachability in a structured system defined in (8) without communication constraints. After this, we state the condition to preserve the reachability in the case of a communication constraint. Finally we explain how the communication sequences may be chosen.

Let us studying the reachability of the extended plant 5. Suppose that $x(0) = 0$ and that the extended plant 5 evolves from $k = 0$ to $k = k_f$. Then

$$x(k_f) = \mathcal{R}(0, k_f) \cdot [\, \bar{u}(0) \;\; \bar{u}(1) \;\; \dots \;\; \bar{u}(k_f - 1) \,]^T,$$

where

$$\mathcal{R}(0, k_f) = [A^{k_f-1} B M_\rho(0) \;\; A^{k_f-2} B M_\rho(1) \;\; \dots \;\; B M_\rho(k_f - 1) \;\;]^T$$

The extended plant 5 is reachable on $[0, k_f]$ if

$$\mathbf{rank}(\mathcal{R}(0, k_f)) = n \tag{10}$$

Actually, this is what is called the controllability *from de origin* in discrete-time linear systems. Controllability has been already studied on the context of structured linear systems. In fact, the condition for a structured linear system to be controllable states as follows [Lin (1974)]

Theorem 4. Let Σ_Λ be the linear structured system defined by (8) with associated graph $\mathcal{G}(\Sigma_\Lambda)$. The system (in fact the pair (A, B)) is structurally controllable if and only if:

- **a**. the system Σ_Λ is input connected,
- **b**. there exists a family of disjoint **U**-*rooted* paths and a family of cycles covering all the state vertex set in the associated directed graph $\mathcal{G}(\Sigma_\Lambda)$

Such conditions remain unchangeable assuming the condition of equation 10. As the system is considered originally reachable, input connection is assumed. Consequently we concentrate in condition (**b**) of Theorem 4 which can be expressed in terms of a maximal matching in the associated bipartite graph $\mathcal{B}_k$ of system Σ_Λ according to Lemma 3.

Proposition 5. Let Σ_Λ be the linear structured system defined by (8) with associated dynamic bipartite graph $\mathcal{B}_k(\Sigma_\Lambda)$. The system is structurally reachable in $[k_0, k_f]$, for $k_0 = 0$, if an only if in $\mathcal{B}_k(\Sigma_\Lambda)$ there exists a maximal and complete ω_ρ-matching of size n.

Proposition 5 can be easily proved considering results of lemma 3 and those found in [Zhang (2005)] reformulated in section 2. It is worth noting that in the dynamic bipartite graph a maximal and complete ω_ρ-matching of size n could not be unique. As a consequence different sequences $\mathcal{S}_j$ may be generated for every maximal and complete ω_ρ-matching found.

Then, according to proposition 5 every combination of disjoint edges which form a maximal ω_ρ-matching of size n, preserves the reachability of the system. Now, before select a communication sequence we have to fix the constraint of the communication medium ω_ρ. Now, we can propose the following algorithm:

Algorithm 1. Let $\mathcal{G}(\Sigma_\Lambda)$ the directed graph associated to the structured system (8):

1. From the directed graph $\mathcal{G}(\Sigma_\Lambda)$ determine the size k of the maximal **U**-*rooted* path ,
2. Build the dynamic bipartite graph $\mathcal{B}_k(\Sigma_\Lambda)$ for k (k might be infinity, in such case fix $k = n$),
3. Set the constraint ω_ρ of the communication medium,
4. If there exists in $\mathcal{B}_k(\Sigma_\Lambda)$ a maximal ω_ρ-matching M of size n
   - 4.1. then the output communication sequence $\boldsymbol{\rho}(k)$ is given by the associated vertex sequence $\mathcal{S}$ formed with the maximal ω_ρ-matching M
   - 4.2. Else the system can not preserve the reachability with the given communication constraint. If it is possible put $\omega_\rho = \omega_\rho + 1$ and return to step 3.

In order to illustrate the selection of the communication sequences we give the next example.

Example 6. Let the dynamic bipartite graph of figure 3 be associated to the structured system 8 for $k = 4$. Suppose the communication constraint imposed for this system is $\omega_\sigma = 1$ only one channel access at a time. It is clear by figure 3 that a maximal 1-matching of size 5 can be found. In fact, the 1-matching of size 5 is given by

$$\{(\mathbf{u}_{1,1}, \mathbf{x}_1), (\mathbf{u}_{2,1}, \mathbf{x}_5), (\mathbf{u}_{2,2}, \mathbf{x}_2), (\mathbf{u}_{2,3}, \mathbf{x}_3), (\mathbf{u}_{2,4}, \mathbf{x}_4)\}$$

Then, vertex sequence associated to this maximal 1-matching is

$$\mathcal{S} = \left\{ \begin{bmatrix} \mathbf{u}_1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix}, \begin{bmatrix} 0 \\ \mathbf{u}_2 \end{bmatrix} \right\}$$

This is indeed the only acceptable sequence for this system with the constraint of only one channel.

5. CONCLUSION

This paper dealt with the generation of communication sequences which preserve reachability of a system with communication constraints. To generate such sequences an original method based on the structural analysis was presented. Different communication sequences may be generated for different communication restrictions and in all the cases the minimal admissible restriction is calculated. The structural approach presented allows to get more insight into the system's limitations and possibilities to generate successful communication sequences.

As it is well known, the reachability property of a system is the dual of the observability property. Consequently, same results are expected when we deal with a communication constraint on the output side. In the FDI context, the structural approach can give more insight when dealing with the detectability of system and the generation of communication sequences that preserves it.

REFERENCES

Boukhobza, T., Hamelin, F., and Martinez-Martinez, S. (2007). State and input observability for structured linear Systems: a graph-theoretic approach. *Automatica*, 43(7), 1204–1210.

Commault, C., Dion, J.M., and Trinh, D.H. (2005). Observability recovering by additional sensor implementation in linear structured systems. In *IEEE Conference on Decision and Control, and the European Control Conference*. Seville, Spain.

Dion, J.M., Commault, C., and van der Woude, J.W. (2003). Generic properties and control of linear structured systems: a survey. *Automatica*, 39(7), 1125–1144.

Halevi, Y. and Ray, A. (1988). Integrated communication and control systems: part i-analysis. *ASME Journal of dynamic systems, measurement and control*, 110(4), 367–373.

Hespaha, J., Naghshtabrizi, P., and Xu, Y. (2007). A survey of recent results in networked control systems. In *Proceedings of the IEEE*, volume 95, 138–162.

Hristu-Varsakelis, D. and Levine, W. (2005). *Handbook of Networked and Embedded Control Systems*. Control engineering. Birkhauser, Boston, MA.

Ionete, C. and Cela, A. (2006). Structural properties and stabilization of ncs with medium access constraints. In *In Proceedings of the 45th IEEE Conference on Decision and Control (CDC)*, 1141–1146.

Lin, C.T. (1974). Structural controllability. AC-19(3), 201–208.

Mao, Z., Jiang, B., and Shi, P. (2009). Protocol and fault detection design for nonlinear networked control systems. *IEEE Transactions on circuits and systems*, 56(3), 255–259.

Martinez-Martinez, S., Mader, T., Boukhobza, T., and Hamelin, F. (2007). LISA: a linear structured system analysis program. In *IFAC Symposium on System, Structure and Control*. Foz do Iguaçu, Brésil.

Murota, K. (1987). *System Analysis by Graphs and Matroids*. Springer-Verlag, New York, U.S.A.

Reinschke, K.J. (1988). *Multivariable Control. A Graph Theoretic Approach.* Springer-Verlag, New York, U.S.A.

Shousong, H. and Qixin, Z. (2003). Stochastic optimal control and analysis of stability of networked control systems with long delay. *Automatica*, 39(11), 1187–1884.

Tipsuwan, Y. and Chow, M.Y. (2003). Control methodologies in networked control systems. *Control Engineering Practice*, 11(10), 1099–1111.

Wang, Y., Ye, H., Ding, S., and Wang, G. (2009). Fault detection of networked control systems subject to access constraints and random packet dropout. *Acta Automatica Sinica*, 35(9), 1230–1234.

Yang, F., Wang, Z., Hung, Y., and Gani, M. (2006). H_∞ control for networked systems with random communication delays. *Transaction on Automatic Control*, 51(3), 511–518.

Zhang, L. (2005). *Access scheduling and controller design in networked control systems*. Ph.D. thesis, University of Maryland.

Zhang, L. and Hristu-Varsakelis, D. (2005). Lqg control under limited communication. In *Proceedings of the 44th IEEE Conference on Decision and Control and the European Control Conference*.

Zhang, L. and Hristu-Varsakelis, D. (2006). Communication and control co-design for networked control systems. *Automatica*, 42(6), 953–958.

Zhang, P. and Ding, S.X. (2006). Fault detection of networked control systems with limited communication. In *6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Comparison of control allocation methods in the presence of Failures for the High Altitude Performance Demonstrator

V. Scordamaglia*, M. Mattei, C. Calabrò*, A. Sollazzo***, F. Corraro*****

** Department D.I.M.E.T, University of Reggio Calabria, Reggio Calabria, Italy*
*** Department of Aerospace and Mechanical Engineering, Seconda Università degli Studi di Napoli, Aversa, Italy*
**** Guidance Navigation and Control Department, C.I.R.A- Italian Aerospace Research Center,Capua, Italy*

Abstract: This paper deals with the application of control allocation concepts to the High Altitude Performance Demonstrator (HAPD) unmanned aircraft studied at CIRA (Centro Italiano Ricerche Aerospaziali). Three different techniques aimed at preserving control performance on the three axes in the presence of multiple actuators, and possible faults are compared. An equivalent classical set of three virtual command surfaces is obtained and a Proportional Integral flight control system with H-infinity performance is first designed. A control allocator is then added to obtain a suitable distribution of the control action on the available surfaces also in the presence of failures. A first technique is based on the off-line calculation of a bank of control allocation matrices taking into account a family of aircraft linearized models which are representative of the operating conditions; a second technique is based on the on-line solution of a Quadratic Programming problem taking into account also control input saturations. The use of H-infinity controllers, scheduled with the possible actuator faults, is finally analyzed which directly assumes all the healthy control surfaces available to guarantee a suitable effort distribution. A comparison of the performance exhibited by the three techniques is made by means of numerical simulations involving the nonlinear mathematical model of the HAPD aircraft.

Keywords: Control Allocation, Reconfigurable Control, H-infinity control, Flight Control, Fault-Tolerant system

1. INTRODUCTION

The idea of using reconfigurable control schemes to cope with actuator failures or control surface damages guaranteeing stability and limited performance degradation is not new to the literature see for example Tao et al. (2002), Kim at al. (2003), Pashilkar et al. (2006), Shin et al. (2004), Shin et al. (2006) and Suresh at al. (2005). Effectiveness of reconfigurable flight control schemes relies on an effective onboard fault detection, isolation and identification (FDI) system to provide accurate and timely fault information (Fig.1) and to the availability of control effectiveness after the fault occurrence. To counteract possible faults, but also to improve control efficiency, modern advanced aircrafts are often configured with redundant aerodynamic control surfaces and actuators. Control allocation in flight control system is aimed at managing how to distribute deflections of multiple control surfaces to generate required control efforts on pitch, roll and yaw. With the increase of the number of redundant actuators, the problem of allocating controls to achieve desired moments on aircraft becomes more complex. The degrees of freedom introduced by actuators redundancy can be used at the control design stage to optimize some performance indexes, like minimum control effort, or to prioritize among the actuators. The literature on flight control with redundant actuators is quite rich having its first applications to canard-elevator configurations, Papageorgiou et al. (1997) and Thrust-Vectored Control, Reiner et al. (1996). Many control allocation algorithms have been developed in recent years including direct control allocation method (Durham, 1993), pseudo inverse based methods (Poonamallee, et al. 2004), daisy chaining method (Bolender et al. 2005), linear programming methods (Harkegard, 2002), quadratic programming based methods (Dorsett et al., 1996), (Nocedal et al. 1999), (Shtessel et al., 2002).

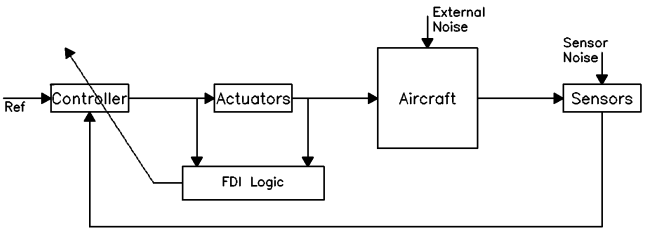


Fig. 1 General Schematic of a reconfigurable control system

In this paper, we consider the High Altitude Performance Demonstrator (HAPD, see Fig.2) aircraft which has been designed at CIRA (Centro Italiano Ricerche Aerospaziali). On this aircraft, having twelve aerodynamic surfaces, we compare three possible reconfiguration strategies taking into account the high level of redundancy offered by the aerodynamic surfaces, in the presence of possible control saturations and/or faults.

The first approach is based on a two stage control design strategy usual for this kind of problems. First a set of equivalent virtual surfaces controlling the aircraft on the three axes is identified and a robust flight control law is designed assuming such a set of virtual inputs. Then a linear and static allocation method based on optimal projection is adopted. A novelty is the use of a family of linearized models,

representative of the aircraft operating envelope, for the design of the allocation matrix. The final implementation of this technique requires a time invariant linear controller and an allocation matrix both scheduled with the fault occurrence.

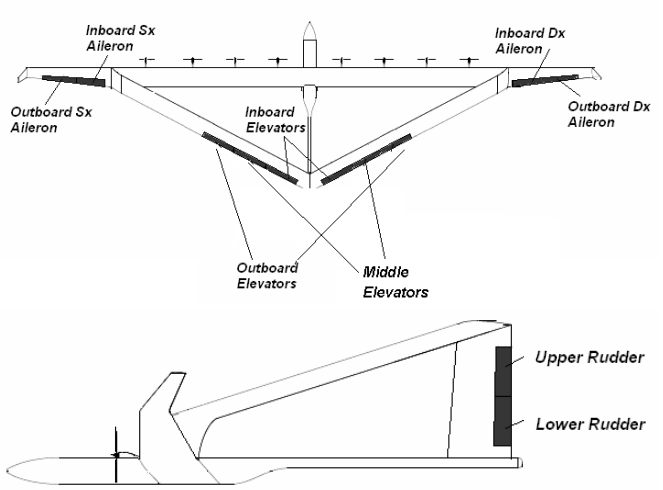


Fig. 2 The HAPD and its twelve control surfaces

The second approach is based on a so-called dynamic control allocation method which requires the on-line solution of a quadratic programming problem to define the distribution control effort over the redundant surfaces. By adding linear constraints to the optimization problem, it is also possible to account for saturations.

The last approach is based on the design of a bank of full authority H-infinity controllers. Each controller of the bank makes use of all healthy control surfaces available. The overall control scheme requires a scheduling of the controller gains with the fault occurrence.

The paper is organized as follows. The general control design problem and methodology adopted to design Proportional Integral H-infinity controllers is dealt with in Section II. The proposed reconfiguration strategies are described in Section III. Finally Section IV provides a description of the numerical simulations carried out to compare performance of the three reconfiguration strategies.

2. THE CONTROLLER STRUCTURE AND DESIGN

It is common to model aircraft as an LPV system to account both for nonlinearities and parametric uncertainties (Mattei and Scordamaglia, 2008). The presence of a parameter vector can account for the dependence of the aircraft linearized model both on the state and inputs, and on possible physical parameters. Hereinafter we consider the problem of designing a flight control law based on a fixed proportional-integral (PI) structure, integrated into a model following control scheme.

We assume the following general model for the aircraft:

$$\begin{aligned} \dot{x}_p &= A_p(\pi)x_p + B_{wp}(\pi)w_p + B_{up}(\pi)u_p \\ y_p &= C_p(\pi)x_p + D_{wp}(\pi)w_p \end{aligned} \tag{1}$$

where $x_p \in \Re^{n_p}$ is the state vector, $u_p \in \Re^{m_p}$ and $w_p \in \Re^{d_p}$ are the control and disturbance input vectors respectively, $y_p \in \Re^{l_p}$ is the vector of measured outputs, and $\pi \in P \subseteq \Re^q$ is a vector of parameters. We assume the following LTI reference model

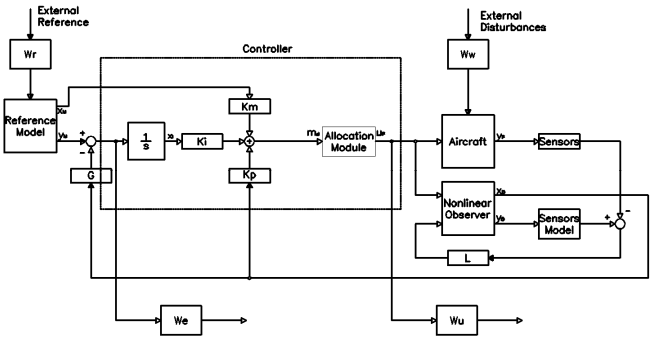


Fig.3 Closed loop System

$$\begin{aligned} \dot{x}_M &= A_M x_M + B_M r \\ y_M &= C_M x_M + D_M r \end{aligned} \tag{2}$$

where $x_M \in \Re^{n_M}$ is the model state vector, $r \in \Re^{m_M}$ is the reference signal input vector. We also define a low dimensional virtual control input $m_d \in \Re^{h_p}$ to be used as essential input for the flight control design, demanding the role of distributing effort on the available surfaces to a, possibly parameter depending, memory-less redistribution function

$$u_p = M(m_d, \pi) \tag{3}$$

which is called control allocation function.

We fix the structure of the PI multivariable controller, assuming full state accessibility

$$\begin{aligned} \dot{x}_I &= C_M x_M + D_M r - G_I x_p \\ m_d &= K_p x_p + K_M x_M + K_I x_I \end{aligned} \tag{4}$$

$x_I \in \Re^{n_I}$ being the state of the multi-integrator, G_I a matrix selecting the controlled outputs, and K_P, K_M , K_I constant controller gains with proper dimensions.

In case of partial state accessibility, we also consider a nonlinear observer of the state in the form:

$$\begin{aligned} \dot{\tilde{x}}_p &= A_p(\pi)\tilde{x}_p + B_{up}(\pi)u_p + L(\tilde{y}_p - y_p) \\ \tilde{y}_p &= C_p(\pi)\tilde{x}_p \end{aligned} \tag{5}$$

Systems (1)-(5) are connected as shown in Fig.3.

To design a full envelope flight control system, an LPV H_∞ approach with pole clustering can be adopted. LPV control allows to account for nonlinearities, parametric variations, and/or uncertainties. H_∞ requirements are given to account for the presence of atmospheric disturbances and neglected dynamics; pole clustering helps avoiding high frequency and low damped modes in the closed loop.

The controller design problem, including H_∞ dynamic weighting filters, is approached on an enlarged plant that can be rewritten, with clear meaning of matrices as:

$$\begin{aligned} \dot{x} &= A(\pi)x + B_w(\pi)w + B_u(\pi)M(\pi)m_d \\ y &= C_y x, \qquad z_\infty = C_{z\infty}(\pi)x + D_{z\infty}(\pi)w \end{aligned} \qquad \pi \in P \tag{6}$$

under the assumption that a control allocation policy (3) has been already chosen and linearized.

Defining with $\mathcal{D}(\alpha_{\min}, \zeta_{\min}, \omega_{n\max})$ the sub-region of the complex plane determined by a maximum natural frequency $\omega_{n\max}$, a minimum damping coefficient $\zeta_{\min}$ and a minimum decay rate $\alpha_{\min}$, the control design problem can be formulated as follows.

Control Design Problem: Given system in (6), find a static output (system and filters states are excluded from measured outputs) feedback virtual control action in the form

$$u_p = \begin{bmatrix} K_p & K_M & K_I \end{bmatrix} \begin{bmatrix} \tilde{x}_p \\ x_M \\ x_I \end{bmatrix} = K_y \cdot y \tag{7}$$

guaranteeing uniform exponential stability of the closed loop against all the time-varying realizations of the parameters $\pi \in P$, guaranteeing an H_∞ performance level γ on the $w - z_\infty$ I/O channel, and guaranteeing that the linearized closed loop poles belong to $\mathcal{D}(\alpha_{\min}, \zeta_{\min}, \omega_{n\max}) \ \forall \pi \in P$.

With a certain degree of conservatism Problem 1 can be solved adopting the approach proposed in Mattei and Scordamaglia (2008), whereas the observer gain matrix *L* can be designed using the H_∞ approach illustrated in Mattei et al. (2005).

When the redistribution function *M* is assumed to be the identity, allocation is directly left to the flight controller (*direct allocation method*). It has been recognized however that this approach may suffer of control surfaces coordination problems.

A possible simple solution to improve control effort distribution is based on an optimization method making use of pseudo-inversions (*pseudo-inversion method*). In facts it is quite natural to choose m_d to be three dimensional vector controlling angular rates along the coordinates axes. If we extract from B_{up} rows related to angular rates, say B_{pqr} the extracted matrix, we can then assume $M = B_{pqr}^{\dagger}$. This choice provides a solution to the following optimization problem.

$$u_p = \arg\min_{u} \ \left\| J \cdot u - N \cdot m_d \right\|_2^2 \tag{8}$$

with $J = B_{pqr}$, $N = I_3$.

Since the input matrix can depends on the vector π, a reasonable choice to obtain a parameter independent redistribution matrix all over the operating envelope can be to compute it as

$$M = J^{\dagger} \cdot N \tag{9}$$

with $J = \left[B_{pqr}^T(\pi_1) \ldots B_{pqr}^T(\pi_p) \right]^T$, $\pi_1, \ldots, \pi_p$ being a representative set of values of the parameters vector $N = \left[I_3 \ldots I_3 \right]^T$.

3. ACTUATORS FAILURES ACCOMODATION USING REDUNDANCY: THREE DIFFERENT METHODS

3.1 Method 1 – Pseudo inversion with scheduled distribution matrix

In the event of actuator faults, aircraft dynamics suddenly change and a closed loop performance deterioration may take place very soon. If a given number N_{SF} of possible faults are identified a-priori, a set of faulted dynamics can be evaluated and corresponding LPV model can be written in the form

$$\begin{aligned} \dot{x}_p &= A_p^{f_k}(\pi) x_p + B_{wp}^{f_k}(\pi) w_p^{f_k} + B_{up}^{f_k}(\pi) u_p^{nf_k} \\ y_p &= C_p(\pi) x_p + D_{wp}^{f_k}(\pi) w_p^{f_k} \end{aligned} \tag{10}$$

being $u_p^{nf_k}$ the vector of healthy control inputs (which depends on the faulted condition), and $w_p^{f_k}$ the vector of disturbances including those possibly caused by faults.

If for each fault scenario we find a redistribution matrix based on the pseudo-inversion strategy (9), we finally approach to a bank of allocation matrices

$$M^{f_k} = J_{f_k}^{\dagger} \cdot N \tag{11}$$

with $J_{f_k} = \left[B_{pqr}^{f_k\,T}(\pi_1) \ldots B_{pqr}^{f_k\,T}(\pi_p) \right]^T$, scheduled on the basis of the fault occurrence identified.

It is worth to notice that also the structure of model reference (4) and the controller gains can be adapted to account for failures. In this case the complete controller-allocator system has to be scheduled.

3.2 Method 2 –on line solution of the allocation problem

The second solution does not require scheduling at the price of some on-line optimizations. It is assumed that the vector of control inputs can be partitioned into two vectors, namely u_{nf} (vector of healthy control inputs) and u_f (vector of faulted control inputs). We denote as H_{nf} (H_f) the matrix mapping u_{nf} (u_f) onto the time derivatives of the body frame angular rates *p,q,r*. *H* denotes H_{nf} in the absence of faults. We also assume that faulted inputs are measurable, and that *M* is the redistribution calculated with no faults according to (9).

When the flight controller generates values of the virtual command m_d, the accommodation problem can be translated into the following quadratic programming problem subject to linear constraints

$$\begin{aligned} \min_{u_{nf}} \quad & \left\| H \cdot M \cdot m_d - H_f \cdot u_f - H_{nf} \cdot u_{nf} \right\|_2^2 \\ s.t. \quad & \underline{u} \le u_{nf} \le \overline{u} \end{aligned} \tag{12}$$

$\underline{u}$ and $\overline{u}$ being the lower and upper bounds on u_{nf} respectively. Actuators rate saturations can also accounted for as shown in Harkegard (2004). QP problem (12) can be efficiently solved by means of several reliable numerical tools.

3.3 Method 3 – direct allocation with controller scheduling

The third approach is based on direct allocation which means that the controller itself takes care of the distribution of effort among control surfaces. In order to increase resilience of the controller to fault occurrence, the following approach is adopted for the controller design.

Assume the fault dependent family of LPV dynamics (10). For a given fault scenario and a given $\pi \in P$, the vector of virtual commands acting on the aircraft is

$$m_d = B_{pqr}^{f_k}(\pi) u_p^{nf_k} \tag{13}$$

The vector of input can then be decomposed in two terms

$$u_p^{nf_k} = u_0^{nf_k} + u_{\perp}^{nf_k} \tag{14}$$

where $u_0^{nf_k} = B_{pqr}^{f_k\,\dagger}(\pi) m_d$ is minimum norm control vector generating m_d. From (13) and (14) it follows that

$$u_\perp^{nf_k} = \left(I - B_{pqr}^{f_k\,\dagger}(\pi) B_{pqr}^{f_k}(\pi)\right) u_p^{nf_k} \quad (15)$$

It's worth to notice that $u_p^{nf_k} \to u_0^{nf_k}$ as $\left\|u_\perp^{nf_k}\right\|$ approaches to zero. A reasonable choice to improve distribution of control effort among healthy redundant surfaces is to minimize the following H_∞ performance index at the controller design stage

$$\left\|u_\perp^{nf}\right\|_2 / \left\|w_p^f\right\|_2 \quad (16)$$

where w_p^f may include the effect of faulted inputs.

4. NUMERICAL RESULTS

We consider the HAPD over-actuated aircraft shown in Fig.2. This is a demonstrator aircraft designed by CIRA in view of a possible high altitude long endurance flight (Cicala, 2009). Its main characteristics are reported in Tab.1. A complete nonlinear model of the aircraft (Cicala, 2008) has been used to test performance and robustness of the proposed allocation strategies.

We assumed as controlled variables ϕ, θ and β which are mainly driven by ailerons, elevators and rudders respectively. The reference models chosen for the three control channels are parameter independent: $W_{\theta_r-\theta} = (s+1)^{-2}$, $W_{\phi_r-\phi} = (s+1)^{-2}$, $W_{\beta_r-\beta} = (2s+1)^{-2}$ where ϕ_r, θ_r and β_r are the requested values of the controlled variables.

Reference models were assumed to be LTI; however, if needed, they could be scheduled with actuators faults to cope with reduced aircraft manoeuvrability. Dynamic weighting W_e filters were chosen to specify the frequency range of interest for the H-infinity performance. In particular first-order low-pass filters with a cut-off frequency of 2 Hz were adopted.

To take into account the dynamic response of the actuators and possible limitations of the control system hardware time response, we chose as pole clustering region $\mathcal{D}(0,0.7,17)$. With such a choice no constraint on the maximum time constant was imposed (α_{min}=0), whereas a desired speed of response on the I/O channels of interest was imposed by the reference models. As for the maximum natural frequency and the minimum damping coefficient, these are limited by 17 rad/s and 0.7 respectively, to avoid problems in the numerical implementations of the controller and low damped closed-loop modes.

We considered 18 design points in the operating envelope with a true air speed within the interval [15,25] m/s and an altitude between 300 m and 1000 m.

Table 1. Specification of Aircraft HAPD

Parameters	Value	Units
Wing Area	13.5	m^2
Wing Span	16.55	m
Mean Chord	0.557	m
Mass	184.4	kg
Maximum altitude	1000	m
Maximum speed	30	m/s

Numerical simulations with the full nonlinear model of the aircraft were performed starting from several trimmed forward flight conditions.

We report results obtained considering 10 deg doublets on the controlled variables.

Figs.4(a-c) shows the response obtained for the whole family of plants considered in the operating envelope, in the absence of faults.

Dashed line represent the reference input. Diamonds denote the output of the reference model. Black solid lines are the results obtained applying the three allocation methods proposed in Section 3. A substantial equivalence of the performance offered by the three methods can be observed.

Fig.5 shows a comparison of the closed loop responses in the presence of a pitch maneuver and of an abrupt fault. In particular at time t = 2s both outboards and middle-boards elevators suddenly return to zero position.

Numerical simulations start from trim condition at an altitude of 300 m and a true air speed of 17 m/s. Dashed-dot line represent the reference model behavior; black solid line represent the response obtained in the absence of any failure accommodation strategy; squares, triangles and circles are obtained considering as accommodation strategy Method 1, Method 2 and Method 3 respectively.

Figs.6 (a-n) show surfaces deflections during the above mentioned pitch maneuver adopting Method 1 (squares) , Method 2 (triangles) and Method 3 (circles). Black solid lines denote surface deflections in absence of any failure accommodation strategy. Dashed lines denote deflection limits.

It is worth to notice from the analysis of Fig. 5 that the controller was not able to compensate the fault maintaining acceptable levels of performance without an accommodation strategy. On the other hand all the failure accommodation schemes considered allow obtaining satisfactory levels of maneuverability in the presence of the simulated fault.

In particular methods based on scheduling proposed in Section 3 exhibit performance which are practically equivalent, whereas the accommodation strategy based on the on-line optimization offers best performance due to the possibility to manage control saturations.

ACKNOWLEDGMENTS

This work has been partially funded by Italian MIUR grant 2008CSS4W3_004

REFERENCES

Bolender, M. A., and Doman, D. B. (2005). Nonlinear control allocation using piecewise linear functions: A linear programming approach. *Journal of Guidance, Control, and Dynamics*, 28 (3), 558-562.

Cicala, M., Sollazzo, A. (2008). HAPD – Modello di Velivolo Elastico orientato al controllo. Technical Report. CIRA-CF-08-0346.

Cicala, M. (2009). HAPD - Analisi Preliminare di Meccanica del Volo e Prestazioni. Technical Rep., CIRA-CF-08-1362.

Dorsett, K., and Mehl, D. (1996). Innovative Control Effectors (ICE), Technical Rep., WL-TR-96-3043.

Durham, W.C. (1993). Constrained control allocation. *Journal of Guidance, Control, and Dynamics*, 16 (4), 717-725.

Harkegard, O. (2002). Efficient active set algorithms for solving constrained least squares problems in aircraft control allocation. *41st IEEE conference on decision and control*, 2, 1295-1300. IEEE, New York.

Harkegard, O. (2004). Dynamic control allocation using constrained quadratic programming. *Journal of Guidance, Control, and Dynamics,* 27 (6), 1028-1034.

Kim, K.S., Lee, K.J., and Kim, Y. (2003). Reconfigurable flight control system design using direct adaptive method. *Journal of Guidance, Control, and Dynamics*, 26 (4), 543-550.

Mattei, M., Paviglianiti, G., and Scordamaglia, V. (2005). Nonlinear identity observers with Hinf performance for sensor FDI: an LMI design procedure. *Control Engineering Practice*,13, 1271–1281.

Mattei, M., Scordamaglia, V. (2008). Full Envelope Small Commercial Aircraft flight control design using multivariable proportionale-integral control. *IEEE Control system technology*, 16 (1), 169-176.

Nocedal, J., and Wright, S. (1999). *Numerical optimization*. Springer Verlag, NewYork.

Papageorgiou, G., Glover, K., and Hyde, R. A. (1997). The $H\infty$ Loop-Shaping Approach. J.-F. Magni, S. Bennani, and J. Terlouw (ed.), *Robust Flight Control: A Design Challenge*, 464-483. Springer-Verlag, London.

Pashilkar, A.A., Sundararajan, N., and Saratchandran, P. (2006). Adaptive back-stepping neural controller for reconfigurable flight control systems. *IEEE Trans. on Control Systems Technology*, 14 (3), 553-561.

Poonamallee, V. L., Yurkovich, S., Serrani, A., Doman, D. B., and Oppenheimer, M. W. (2004). A nonlinear programming approach for control allocation. *Proceedings of the American Control Conference*, 2, 1689-1694.

Reiner, J., Balas, G. J., and Garrard, W. L. (1996). Flight Control Design Using Robust Dynamic Inversion and Time-Scale Separation. *Automatica*, 32 (11), 1493–1504.

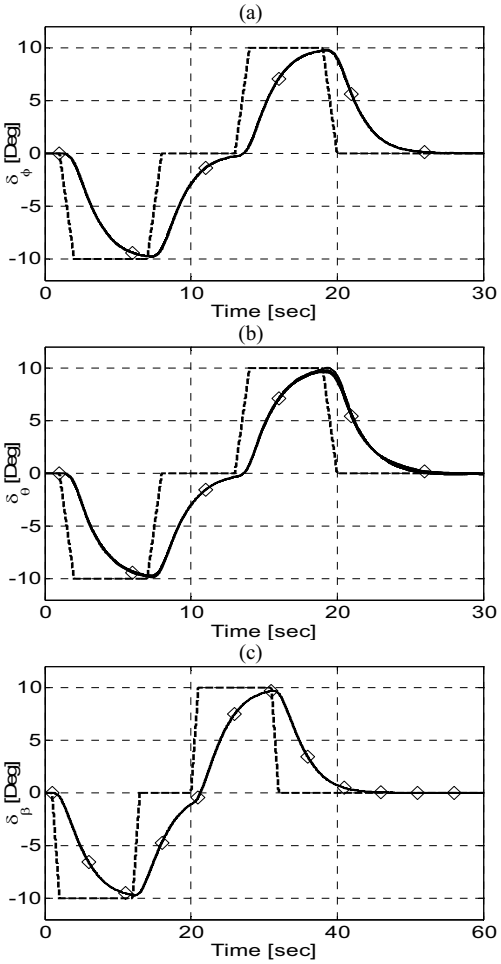
Shin, D.H., Kim, Y. (2004). Reconfigurable flight control system design using adaptive neural networks. *IEEE Trans. on Control Systems Technology*, 12 (1), 87-100.

Shin, D.H., Kim, Y. (2006). Nonlinear discrete-time reconfigurable flight control law using neural networks. *IEEE Trans. on Control Systems Technology*, 14 (3), 408-422.

Shtessel, Y., Buffington, J., and Banda, S. (2002). Tailless aircraft flight control using multiple time scalere configurable sliding modes. *IEEE Transactions on Control Systems Technology*, 10 (2), 288-296.

Suresh, S., Omkar, S.N., Mani, V. et al (2005). Nonlinear adaptive neural controller for unstable aircraft. *Journal of Guidance, Control, and Dynamics*, 28 (6), 1103-1111.

Tao, G., Chen, S., and Joshi, S.M. (2002). An adaptive actuator failure compensation controller using output feedback. *IEEE Trans. On Automatic Control*, 47 (3), 506-511.



Figs.4(a-c) Time Responses for ϕ during roll maneuver, for θ during pitch maneuver and for β during yaw maneuver (results of three different simulations)

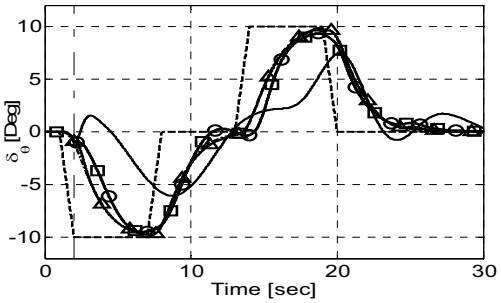
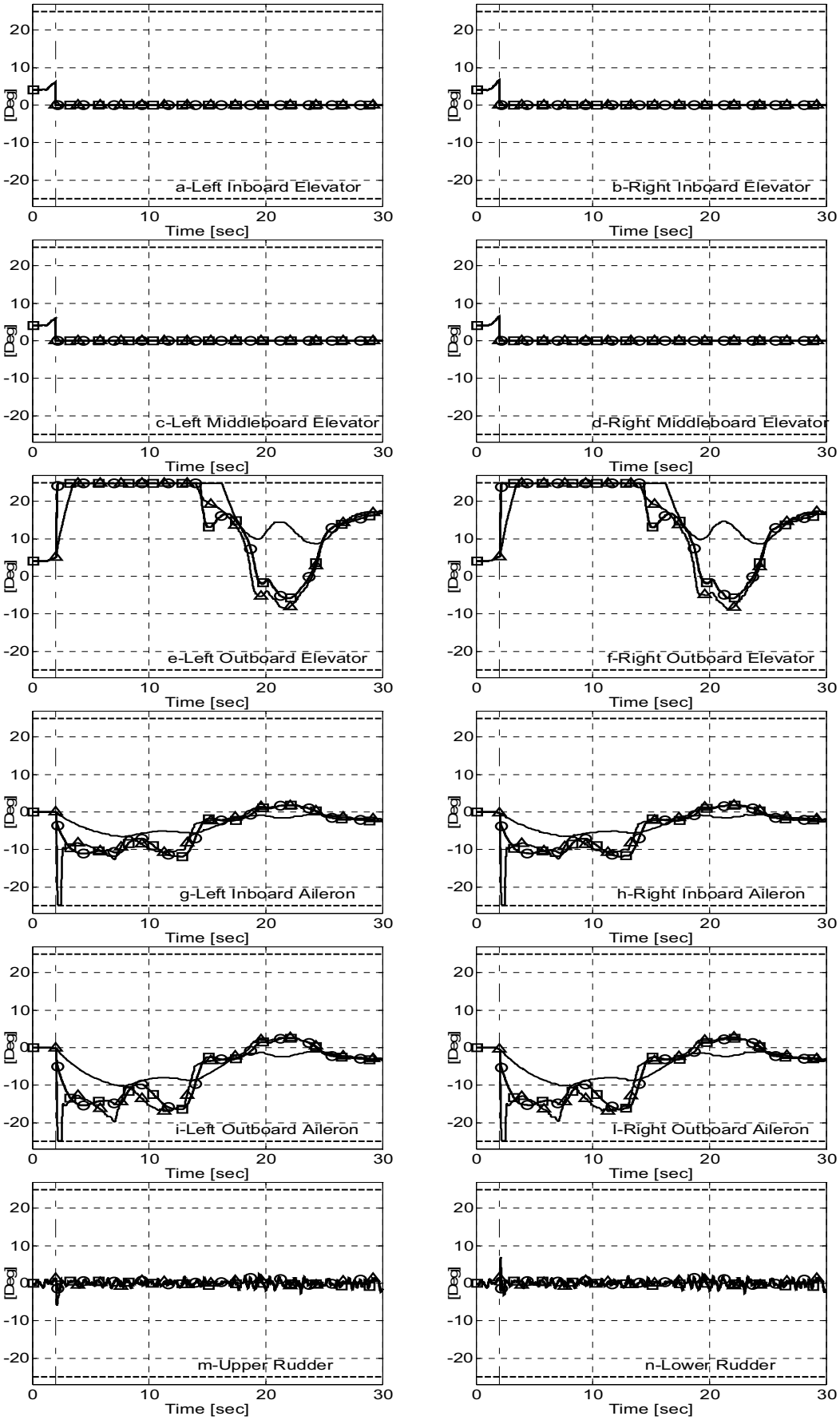


Fig.5 Comparison of time responses for θ during pitch maneuver in presence of Faults on both outboards and middle-boards elevators



Figs.6(a-n) Comparison on surfaces deflections during pitch maneuver in presence of faults on both outboards and middle-boards elevators

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Temporal Reliability Analysis of Embedded Systems

A. Ghenai*. M. Benmohammed**

**LIRE Laboratory, Mentouri University, Constantine, 25000 Algeria (Tel: +213-772-774-899; e-mail: afifa.ghenai@ gmail.com).*
***LIRE Laboratory, Mentouri University, Constantine, 25000 Algeria (e-mail: ben_moh123@yahoo. com)*

Abstract: The criticality of embedded systems requires to guarantee a suitable level of reliability. One of the principal problems encountered when we study the reliability of these systems is the taking into account efficiently and in realistic way of time constraints to which they are subjected. In this article, we propose a reliability analysis approach of embedded systems which is pressed on a formal framework based on equivalence between reachability in Time Petri-nets and the provability of a TCTL formula. The translation of the Time Petri-net into TCTL enables us to propose a new formal definition of the concept of scenario which takes the system towards a feared (dangerous) state, from a normal functioning state, in order to understand the reasons of the drift which can have dramatic consequences for the system and the user.

Keywords: Embedded systems, Reliability, Time constraints, Time Petri-nets, TPN-TCTL.

1. INTRODUCTION

An embedded system is a system controlled by a calculator combining various technologies (mechanics, hydraulics or electric). It must answer several requirements of which the criticality, which, requires to guarantee the major challenge: a suitable level of reliability.

The traditional methods of reliability reach quickly their limits: The combinative methods (failure trees, reliability diagrams) only make it possible to identify and evaluate the events combinations leading to the catastrophe occurrence. They do not hold account about occurrence of the events which compose them. This excludes any possibility of taking into account the dependence and times between events. Methods based on discrete events systems, answer well the problems of order between events (Petri-nets) but are limited by the combinative explosion problem (because of the use of the reachability graph) [Sadou 2007].

To circumvent the problems arising from an enumerative approach founded on the markings graph, the approach suggested by Khalfaoui [Khalfaoui 2003], is based on the linear logic which makes it possible to build a partial order of transitions firings and uses directly the Petri-net model without generating the associated reachability graph, to extract scenarios taking the system towards a critical condition (called : *feared scenarios*) which are indeed, unknown during the design phase of embedded systems. To implement the approach, an algorithm of search for feared scenarios is proposed [Khalfaoui 2003], by coupling the Petri-net model describing nominal operation, the failures and reconfiguration mechanisms, with differential equations representing the evolution of continuous variables of the system energy part. The limits of this algorithm are due to the fact that it operates only on the discrete aspect of the model and that many incoherent scenarios with respect to continues dynamics are generated.

Moreover, the occurrence order, due to this continuous dynamics of events is not taken into account [Sadou 2007].

In [Medjoudj 2006], Medjoudj took again the approach of Khalfaoui by working out a new version of the algorithm to determine more precisely, by a time Petri-net model, the exact conditions of the feared event occurrence. The continues part is partially taken into account by time abstractions of continuous dynamics, which makes it possible to eliminate a number of scenarios generated in the first version and which are incoherent with continues dynamics, but not totality.

2. FEARED SCENARIOS GENERATION APPROACH

In front of the limits of these approaches of embedded systems design, Sadou in [Sadou 2007] proposed an approach in which feared events define reliability requirements and taking them into account must lead to a design of a system able to avoid these events. The determination of feared scenarios and their analysis make it possible to propose solutions and to evaluate them. The method of search for feared scenarios is pressed on a formal framework which is linear logic [Girard 1987]. A quantitative analysis makes it possible to determine a partial order of transitions firings and thus to extract directly feared scenarios from the Petri-net model. Linear logic allows to focus the analysis on the interesting parts of the system from a reliability point of view, thus avoiding the exploration of the complete system and the eternal combinative explosion problem [Sadou 2007]. At the occurrence of an event which can endanger users life,

some system requirements are carried out in order to maintain the system in a degraded but sure state. It is possible that the configuration fails leading the system in a state called *feared state* which can have dramatic consequences for the system and the user [Sadou and al 2006a].

2.1 Scenario Formalization with Petri-nets and Linear Logic

To translate Petri-nets into linear logic, the fragment MILL (Multiplicative Intuitionnist Linear Logic) contains the necessary connectors.

The connector ⊗ is used to represent resources accumulation (the formula a ⊗ b ⊗ b expresses the availability of a copy of the resource 'a' and two copies of the resource 'b') and the linear implication (represented by the symbol —o) allows to take into account the production and the consumption of the resources. For example, the formula 'a—ob' represents the consumption of the resource 'a' to produce the resource 'b' [Sadou 2007].

A Petri-net transition is represented by an implicative proposition which can be consumed during the proof, which will indicate that the transition is actually firing. A logical formula is associated with each marking and each transition. A marking corresponds to the simultaneous presence of tokens in a set of places, to each place corresponds an atomic proposition.

A transition expresses a relation of causality between two marking formulas. It is translated by an implicative formula (the connector —o). The left side of the formula establishes minimal marking to fire the transition, while the right side represents the marking reached after the firing of this transition from minimal marking [Medjoudj 2006]. The translation of a Petri-net in linear logic is done in the following way:

- An atomic proposition P is associated with each place P of the Petri-net.
- A monomial, using the multiplicative conjunction ⊗, is associated with each marking, Pre( ) Pre-condition and Post ( ) Post-condition of transitions.
- For each transition t of the Petri-net, a implicative formula is defined in the form:

$$t : \bigotimes_{i \in Pre(pi,t)} p_i \multimap \bigotimes_{o \in Post(po,t)} p_o$$

A scenario can be represented by one sequent of linear logic. Reachability between two markings M_0 and M_f is represented by this sequent. The left part of sequent contains the list of all firings of transition allowing to reach marking M_f from M_0 marking. This part of sequent contains also the formula representing initial marking. The right part of sequent (the conclusion) contains the formula representing final marking. Sequent expresses reachability between M_0 and M_f markings. It is written in the form : $M_0, t_1, \ldots, t_n \vdash M_f$ [Sadou 2007].

Proving a sequent consists in showing that it is syntactically correct. Since there is equivalence between reachability in a Petri-net and the provability of sequent in linear logic, the proof of sequent can be expressed by the construction of a proof tree [Medjoudj 2006].
Each node of the tree is one sequent premise simpler than its conclusion. A sequent is provable if and only if there exists a proof of which it is the root. The construction of a proof tree of sequent is an iterative step which consists in eliminating in each stage each drawn transition after checking that the atoms necessary to its firing are available (produced). This stage must be carried out once with each firing of a transition. It makes it possible to determine the relations of precedence imposed by the structure and the marking of the Petri-net. For each atom, the application of the iterative stage determines the transition which produced the atom and which consumed it [Sadou 2007].

2.2 Proof Tree Labelling

The proof of a sequent gives a proof tree for each sequence of firing transitions. The partial order between these firings is implicitly present in the proof tree and defined formally by a scenario. To clarify it, Nicolas Rivière in [Rivière 2003] has introduced the following labelling process:

Annotation of the rules: Each time the rule: —oL (which eliminates an instance of formula describing a firing of transition) is applied, a label with the name of the transition corresponding to the eliminated implicative formula is associated. When a transition is fired several times, we put by exposing an index equal to the number of firings carried out. Thus the label t^j_i means that it is the j^{th} elimination of a formula associated with the transition t_i. That makes it possible to differentiate the firings of the same transition.

Annotations of the atoms: Each atom of a current stage is labelled differently according to whether it is on the left or on the right of sequent to prove. When this atom is on the left, this label is that of the rule which produced it. When it is on the right, it takes the label of the rule which consumed it. So that all the atoms have a label, the atoms associated with initial marking of the scenario are labelled by events I^i (representing firings having produced the corresponding tokens) and the atoms of final marking by F^j events (representing crossings having consumed the corresponding tokens). These added labels have the advantage of allowing the composition of scenarios [Rivière 2003].

2.3 Illustration of the Annotated Proof Tree

2.3.1 Scenario

The definition of a scenario is based on the concept of event and the relations between the events.

Definition 1. (Event): Let be a Petri-net (P, T, Pre, Post), M_0 its initial marking. An event is a particular firing of a transition $t \in T$. the set of events is noted E. For example, if during an evolution of the Petri-net from M_0 the transition t_i is fired for the j^{th} time, we will say that it is the occurrence of the event e_i^j.

Definition 2. (Scenario): A scenario sc, noted $sc = (l, \prec_{sc})$ associated with the Petri-net P and the couple M_0 and M_F markings, is a set of events *l provided with a strict partial*

order $\prec_{sc}$ *defined on the events of l.* If for $e_1, e_2 \in l$ we have $e_1 \prec e_2$, that wants to say that the event e_1 precedes the event e_2 in the scenario *sc* [Sadou and al 2006b].

2.3.2 Example

Let be the Petri-net of figure (Fig. 1) presented in [Sadou 2007]. The figure (Fig. 2) shows the construction of the proof tree, for sequent: $P_1 \otimes P_2$, $t_a \otimes t_b \otimes t_c \vdash P_3 \otimes P_4$. This tree corresponds to a particular firing of the sequence (*a, b, c*).

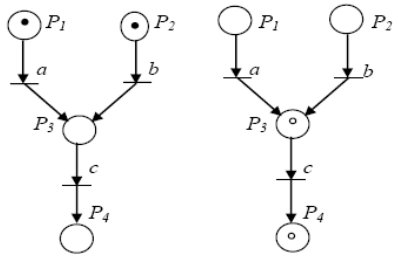


Fig. 1. Petri-net model: initial and final markings

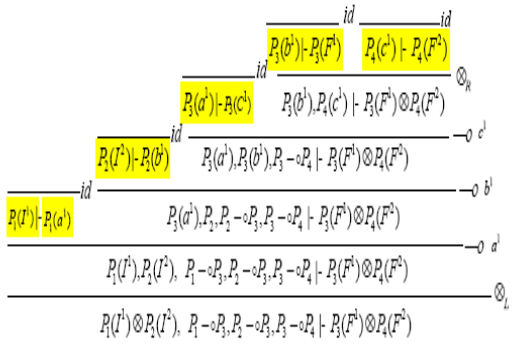


Fig. 2. Proof tree

3. OUR APPROACH

From all these reports, and to circumvent the combinative explosion problem of the reachability graph of an embedded system, The proposed approach in [Khalfaoui 2003], improved and implemented in [Medjoudj 2006] and [Sadou 2007] and which is based on the extraction of feared scenarios from a Petri-net model, seems to us well adapted to face the increasing complexity of embedded systems. Moreover, natural parallelism of Petri-nets is not preserved by the traditional analysis methods such as model-checking which is one of the reasons which can bring to the combinative explosion problem.

The description of the scenarios which take the system towards the feared state from a normal functioning state makes it possible to understand the reasons of the drift in order to envisage the necessary configurations which make it possible to avoid them.

In this article, we propose a new formal definition, of this concept of scenario. Our approach of search for feared scenarios is based on a formal framework which is TCTL logic, a new representation of the Petri-net model. For a good taking into account of time constraints, Time Petri-nets model (TPN) is selected. The representation of the scenario in logical form is done by a TCTL formula. The translation of the Time Petri-net into a TCTL formula will then enable us to formalize the concept of scenario and its properties.

3.1 TCTL : Timed Computation Tree Logic

In the real time field, time requirements are divided more precisely into two categories:
- Requirements where time is expressed in a qualitative (or logic) way. It is not considered whereas a partial order between events.
- Requirements where time is expressed in a quantitative way. We consider in this case the order of the events but also the time distances between them.

TCTL is an extension of CTL which integrates quantitative time constraints. TCTL is a temporal language (causality relations) and time (quantification of time besides the causality relations).

3.2 Time Petri-nets (TPN)

The checking of the time constraints is in particular difficult to realize by traditional tests, since it would be necessary in theory to test infinity of different time sequences. An alternative is then to build a time-lag formal model of the system. The sure design and the development of complex dynamic systems require in particular their modelling according to a rigorous and nonambiguous formalism. To this end, many formal languages were developed since the middle of last century. Among those, Time Petri-nets (TPN) constitute a powerful tool of design and analysis, particularly adapted to the description of dynamic systems. The language rises from the traditional Petri-nets by the addition of time constraints, in the form of an interval, on the occurrence of the events. More precisely, we place ourselves here within the framework of T-time Petri-nets (i.e. associating temporal information with transitions), in dense time. In TPN, places generally model a state of the system whereas transitions represent events or the validation of conditions; system evolution is modelled by the transit of tokens between the places. A marking of the network is then a vector $M \in \mathbb{N}^P$ such as for any place $p \in P$, M (p) represents the number of tokens in the place p., a transition t is sensitized by marking M if any place upstream of t contains at least as many tokens as indicated by the weight of the arc connecting this place to t. The transition t can then be fired if it is continuously sensitized since a duration at least equal to a(t) (its date of firing as soon as possible). The choice of strong semantics imposes that t must be fired at the latest at the date b(t), its date of firing at the latest (unless being desensitized meanwhile by the firing of a transition).

A transition t is sensitized by marking M if any place upstream of t contains at least as many tokens as indicated by the weight of the arc connecting this place to t : $M \geq^{\bullet} t$. It is noted in this case: $t \in enabled(M)$.

The transition t is lately sensitized by the firing of a transition t_f since marking M, which is noted $\uparrow enabled(t, M, t_f)$, if it is sensitized by marking $M - {}^\bullet t_f + t_f{}^\bullet$ but was not it by marking $M - {}^\bullet t_f$. Formally:

$$\uparrow enabled(t, M, t_f) = \begin{cases} {}^\bullet t \leq M - {}^\bullet t_f + t_f{}^\bullet \\ (t = t_f) \vee ({}^\bullet t > M - {}^\bullet t_f) \end{cases}$$

By extension, the set of transitions lately sensitized by the firing of the transition t_f since marking M is noted $\uparrow enabled(M, t_f)$. It defines the set of transitions whose clocks are given to zero by the firing of t_f.

TPN semantics is defined as a TTS (Timed Transition System) whose states are formed by the association of a marking M and valuation of clocks v, v(t) representing the time passed since the last sensitizing of the transition t [Seidner 2009].
In TPN semantics, the transition relation is made up of:
The relation of discrete transition defined for all $t_f \in T$ by:

$$(M,v) \xrightarrow{t_i} (M',v') \text{ iff } \begin{cases} M \geq {}^\bullet t_i \\ M' = M - {}^\bullet t_i + t_i{}^\bullet \\ \alpha_i \leq v_i \leq \beta_i \\ \forall k \in [1,n],\ v'_k = \begin{cases} 0 \text{ if } t_k \in \uparrow enabled(M, t_i) \\ v_k \text{ otherwise} \end{cases} \end{cases}$$

The relation of continues transition defined for all $\delta \in \mathbb{R}_{\geq 0}$ by :

$$(M,\nu) \xrightarrow{\delta} (M,\nu') \Leftrightarrow \begin{cases} \nu' = \nu + \delta \\ \forall t \in T,\ t \in enabled(M) \Rightarrow \nu'(t) \leq b(t) \end{cases}$$

3.3 TPN-TCTL Logic

The report which we can make today is that no direct method (i.e. without translation into timed automata) was proposed for the checking of quantitative time properties on TPN. Moreover, no tool of model-checking (i.e. allowing from the model of a system S and a property φ, to decide S |= φ) is available whereas several effective methods and tools allowing the checking of properties expressed with TCTL logic exist on timed automata: Uppaal, Kronos.

In [Boucheneb and al 2009], a TCTL for Time Petri-nets (TPN-TCTL) is defined. The time intervals represent time constraints on a sequence of firings transitions.

Syntax.(TPN-TCTL) The syntax of TPN-TCTL formulas is given inductively by the following grammar:

$$\varphi := P \mid \neg\varphi \mid \varphi \Rightarrow \psi \mid \exists\varphi\ U_I\ \psi \mid \forall\varphi\ U_I\ \psi$$

where: $I \in \mathbf{I}(Q^+)$ is a time interval, $P \in PR$, and $PR = \{P \mid P : M \rightarrow \{true, false\}\}$ is the set of the propositions on markings of the TPN. The following abbreviations are used:
$\exists\blacklozenge_I \varphi = \exists true\ U_I \varphi$, $\forall\blacklozenge_I \varphi = \forall true\ U_I \varphi$,
$\exists\ \ _I \varphi = \neg\forall\blacklozenge_I \neg\varphi$, and $\forall\ \ _I \varphi = \neg\exists\blacklozenge_I \neg\varphi$.

We also define the bounded answer formula by:

$$\varphi \rightsquigarrow_I \psi = \forall\ (\varphi \Rightarrow \forall\blacklozenge_I \psi).$$

which expresses that when the formula φ becomes true, the formula ψ must become true in a time interval I. Where:
For the operator ♦ : «finally», the formula is true if it is checked in a future state of the way.
For the operator : «globally», the formula is true if it is checked in all the following states of the way. [Traounouez 2009].

3.4 Formalization of Scenario with TPN and TPN-TCTL Logic

In our approach, we propose that a logical formula is associated with each marking and each transition of the TPN. A marking corresponds to the simultaneous presence of tokens in a set of places, to each place corresponds an atomic proposition. A transition expresses a relation of causality between two marking formulas. The left side of the transition establishes minimal marking to fire the transition, while the right side represents the marking reached after the firing of this transition from minimal marking.

The translation of a TPN into TCTL is done in the following way:
- A proposition M(P) is associated with each place P of the TPN, it is the marking of the place P.
- Each transition t_i of the TPN will be represented by the formula:

$$\mathbf{M[P_i(J)] = k \Rightarrow O_{Ii}\ M[P_o(J')] = k'}$$

With:
J, J' : label of the rule which produced the token, J=I for initial marking and J=F for final marking.
k, k' : tokens number (copies number of the same resource)
$\mathbf{I}_i$: time interval associated with the transition t_i.
O_{Ii} : « Next » operator of TCTL logic.

- A scenario will be considered as an execution ρ (a sequence of transitions: continuous transitions make it possible to run out time whereas discrete transitions correspond to firings of a transition of the Time Petri-net). $\rho^*(r)$ is the state reached in the sequence ρ after a time of *r* units of time.

The execution ρ (*the scenario*) can be represented by a bounded answer formula. Reachability between two markings M_0 (the initial state) and M_f (the final state) is represented by this formula:

$$\varphi \rightsquigarrow_I \psi = \forall\ (\varphi \Rightarrow \forall\blacklozenge_I \psi)$$

The left part of the formula can be translated during the proof of the scenario, by the list of all firings of transitions making it possible to reach marking M_f (the formula Ψ) from M_0 marking (the formula φ) in a time interval **I**.

We propose to lead the proof of the bounded answer formula by the construction of a proof tree. A TPN-TCTL formula is provable if and only if there exists a proof of which it is the

root. The construction of a proof tree is an iterative step which consists in eliminating in each stage each transition fired after checking of its time interval and that atoms necessary to its firing are available (produced). It makes it possible to determine the relations of precedence imposed by the structure and the marking of the TPN and to determine the execution time of the scenario. For each token, the application of the iterative step determines the transition which produced the token.

3.4.1 Illustration of the Proof Tree

3.4.1.1 Scenario

The definition of a scenario is based on the concept of event and the relations between the events.

Definition 1. (Event): Let be a Time Petri-net (P, T, Pre, Post), M_0 its initial marking. An event is a particular firing of a transition $t \in T$ in a time interval **I**. the set of events is noted E. For example, if during an evolution of the time Petri-net from M_0 the transition t_i is fired for the j^{th} time, we will say that it is the occurrence of the event $e_i^j(\mathbf{I}_i)$.

Definition 2. (Scenario): A scenario sc, noted $sc = (l, \prec_{scI})$ associated with the Time Petri-net P and the couple M_0 and M_f markings, is a set of events l provided with a strict partial order $\prec_{scI}$ defined on the events of l. If for $e_1, e_2 \in l$ we have $e_1 \prec_{\mathbf{I}} e_2$, that wants to say that the event e_1 precedes the event e_2 in the scenario *sc* and that e_1 occurs **I** units of time before e_2.

3.4.1.2 Example

Let us take again the Petri-net of figure (Fig. 1). In order to extend this Petri-net to a time Petri-net (TPN), we associate the time intervals I_1, I_2 and I_3 respectively with the transitions a, b, and c. With: $I_1 < I_2 < I_3$, as shown in the Figure (Fig. 3).

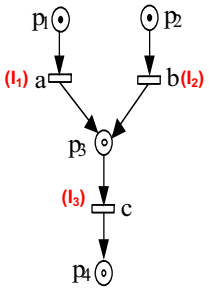


Fig. 3. Time Petri-net

We note:
$\mathbf{I_1} = [t_{min}(a) , t_{max}(a)]$.
$\mathbf{I_2} = [t_{min}(b) , t_{max}(b)]$.
$\mathbf{I_1} = [t_{min}(c) , t_{max}(c)]$.

We note also:
$t_{min}(a) < t_{min}(b) < t_{min}(c)$, and ,
$t_{max}(a) < t_{max}(b) < t_{max}(c)$.

- The transition *a* can be fired in $\boldsymbol{T_a}$ units of time, with:

$$\mathbf{t_{min}(a) \le T_a \le t_{max}(a)} \ ;$$

- The transition *b* can be fired in $\boldsymbol{T_b}$ units of time, with:

$$\mathbf{t_{min}(b) \le T_b \le t_{max}(b)} \ ;$$

- The transition *c* can be fired in $\boldsymbol{T_c}$ units of time, with:

$$\mathbf{t_{min}(c) \le T_c \le t_{max}(c)} \ .$$

We choose strong semantics of Time Petri-nets (TPN) which imposes that a transition *t* must be fired at the latest at its date of firing at the latest : $\mathbf{t_{max}(t)}$.

- The transition *a* of the TPN will be represented by the TCTL formula :

$$\mathbf{M[P_1(I)] = 1 \Rightarrow O_{Ta}\, M[P_3(a)] = 1}$$

- The transition *b* of the TPN will be represented by the TCTL formula :

$$\mathbf{M[P_2(I)] = 1 \Rightarrow O_{Tb}\, M[P_3(b)] = 1}$$

- The transition *c* of the TPN will be represented by the TCTL formula :

$$\mathbf{M[P_3(a,b)] = 2 \Rightarrow O_{Tc}\, M[P_4(c)] = 1}$$

The figure (Fig. 4) shows the construction of the proof tree, for the following TPN-TCTL formula:

$$\mathbf{((M[P_1(I)]=1)\wedge(M[P_2(I)]=1)) \rightsquigarrow_{Ta+Tb+Tc} ((M[P_3(F)]=1) \wedge (M[P_4(F)]=1))}$$

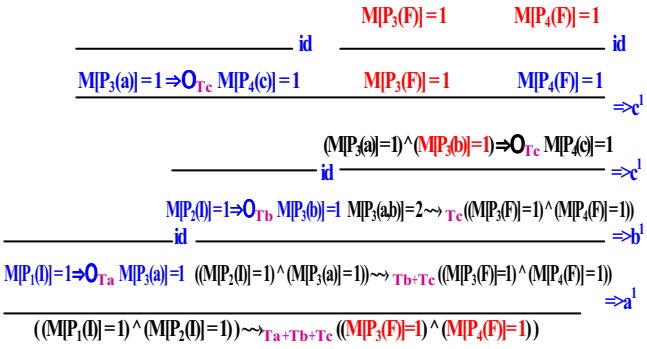


Fig. 4. Proof tree

This tree corresponds to a particular firing of the sequence $(a(I_1), b(I_2), c(I_3))$. The execution time of the scenario can thus be estimated: $\mathbf{T_a + T_b + T_c}$ units of time.

4. CONCLUSION

In this article, we presented the limits of the principal approaches of sure design of embedded systems, we explained our approach to take into account time constraints to which are subjected the embedded systems, in a realistic and effective way, in order to undertake a reliability study on these systems.

Our approach is based on TCTL logic as formal framework, which makes it possible to consider the order of events but

also the time distances between them. This formal framework is based on equivalence between reachability in Time Petri-nets (TPN) and the provability of a TPN-TCTL formula. The translation of the Time Petri-net into TCTL enabled us to propose a new formal definition of the concept of scenario which takes the system towards a feared (dangerous) state from a normal functioning state and to estimate its (dense) time of execution.

REFERENCES

Boucheneb, H., Gardey G., and H.Roux, O. (2009). TCTL model checking of time Petri-nets. *Journal of Logic and Computation*.

Girard, J.Y. (1987). Linear logic. *Theoretical Computer Science*, 50.

Khalfaoui, S. (2003). Méthode de recherche des scénarios redoutés pour l'évaluation de la sûreté de fonctionnement des systèmes mécatroniques du monde automobile, thesis, Institut National Polytechnique, Toulouse.

Medjoudj, M. (2006). Contribution à l'analyse des systèmes pilotés par ordinateurs : extraction de scénarios redoutés et vérification de contraintes temporelles, thesis, Paul Sabatier University, Toulouse.

Rivière, N. (2003). Modélisation et analyse temporelle par réseaux de Petri et logique linéaire, thesis, INSA, Toulouse.

Sadou, N., Demmou, H., Pascal, JC., and Valette, R. CIFA (2006a). Fiabilité dynamique des systèmes hybrides : Approche basée scénarios.

Sadou, N., Demmou, H., and Valette, R. Fac (2006b). Minimalité des scénarios dans le cadre des réseaux de Petri.

Sadou, N. (2007). Aide à la conception des systèmes embarqués sûrs de fonctionnement, thesis, Toulouse III University- Paul Sabatier.

Seidner, C. (2009). Aide à la vérification des EFFBDs : Model checking en ingénierie système, thesis, Nantes University.

Traounouez, L.M. (2009). Vérification et dépliages de réseaux de Petri temporels paramétrés, thesis, Nantes University.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Data–Driven and Model–Based Fault Diagnosis of Wind Turbine Sensors

Silvio Simani [*,⋆] **Paolo Castaldi** [**] **Marcello Bonfè** [*]

[*] *Department of Engineering, University of Ferrara. Via Saragat 1E. 44122 Ferrara (FE), Italy, (Tel/fax: +390532974844; e-mail: {silvio.simani, marcello.bonfe}@unife.it).*
[**] *Aerospace Engineering Faculty, University of Bologna. Via Fontanelle, 40. 47100 Forlì (FC), Italy. (e-mail: paolo.castaldi@unibo.it).*

Abstract: In order to improve reliability of wind turbines, it is important to detect faults in their very early occurrence, and to handle them in an optimal way. This paper focuses on the pitch sensors of the turbine blade system, as they are mainly used for wind turbine control, in order to maximise the power production, and the efficiency of the whole process. On the other hand, as the input–output behaviour of the system under diagnosis is nonlinear, this work suggests a modelling scheme relying on piecewise affine models, whose parameters are identified through the acquired input–output measurements affected by measurement uncertainty. Therefore, these prototypes are exploited for generating suitable residual signals, which allow the detection and the isolation of the considered sensor faults. This noise rejection scheme is used since the wind turbine measurements are not very reliable, due to the uncertainty of wind speed acting on the wind turbine, and to the turbulence around the rotor plane. A detailed benchmark model simulating the wind turbine where realistic fault conditions can be considered shows the effectiveness of both the identification and fault diagnosis techniques.

1. INTRODUCTION

The key step towards system supervision, monitoring, diagnosis, and control design is to find a suitable mathematical description of the process under investigation. In some cases system modelling based on insight on the physical laws, which govern the real process behaviour might be cumbersome and practically infeasible. On the other hand, input–output process measurements can be successfully used to infer analytical descriptions of the system in the framework of a parametric structure, which possess approximation properties with respect to the complex, nonlinear, unknown analytical functions that are amenable as candidate to describe the real behaviour of the observed process (Juditsky et al. [1995]). So, the choice of the parametric structure become an important and difficult step towards the system identification, especially when the behaviour of the target process is nonlinear, as it is common by far in real world applications (Billings and Voon [1983]). As matter of fact, nonlinear models have received great attention by researchers, as they can overcome limits of linear models to describe complex processes, often encountered in fault diagnosis–oriented applications (Chen and Patton [1999], Korbicz et al. [2004]).

The mathematical treatment of nonlinear models follows different approaches and cover topics ranging from approximation theory, estimation theory, non–parametric regression to the most modern techniques based on use of neural networks, wavelets, and fuzzy models (Sjöberg et al. [1995]). The modelling approach suggested in this work refers to a nonlinear process, namely a wind turbine, which operates at different regimes, in which distinct models can be associated to each admissible operating condition. A switching function governs the transition among different models or interpolations of models. Such mathematical descriptions are referred in current literature as *hybrid models*.

This paper suggests a fault diagnosis strategy based on dynamic, discrete–time, time–invariant, affine models describing locally the behaviour of the monitored wind turbine in its different operating regimes. This type of models have been formerly proposed by Simani, *et al.*, and used in stochastic environment for time series model identification (Fantuzzi et al. [2002]).

Concerning the fault diagnosis issue, symptoms are signals representing inconsistencies between the model and the actual system being monitored. Any inconsistency will indicate a fault in the system. Residual must, therefore, be different from zero when a fault occurs and zero otherwise. However, the deviation between the model and the plant is influenced not only by the presence of the fault but also the modelling error. Several techniques had been proposed for Fault Detection and Isolation (FDI) in dynamic systems (Chen and Patton [1999], Korbicz et al. [2004]). In particular, in this work, the hybrid modelling scheme is combined with the model–based method to formulate a FDI technique exploiting the identified piecewise affine prototype (Fantuzzi et al. [2002]) for residual generation. Under such a scheme, a number of local affine models are designed and the estimate of outputs is given by a fusion of local outputs. The diagnostic signal (symptom or residual)

⋆ Corresponding author.

is the difference between the estimated and actual system output.

In this paper, the different operating points are self–selected with a clustering method presented in (Babuška [1998]). On the basis of knowledge of the operating point regions, the identification of the structure, and the parameters of each local affine dynamic model has been performed (Fantuzzi et al. [2002]). This modelling scheme is used here since the wind turbine measurements are not very reliable, due to the uncertainty of wind speed acting on the wind turbine, and to the turbulence around the rotor plane. A detailed benchmark model simulating the wind turbine where realistic fault conditions can be considered shows the effectiveness of both the proposed identification and fault diagnosis techniques.

The remainder of this paper is organised as follows. Section 2 recalls the structure of the exploited multiple model. Section 3 shows the design of the diagnostic scheme for the FDI of dynamic systems. The application of the FDI approach to the wind turbine is described in Section 4. The example demonstrates the effectiveness of the technique proposed. Finally, some concluding remarks are included in Section 5.

2. HYBRID MODELLING AND IDENTIFICATION

The main idea underlying the mathematical description of nonlinear dynamic systems is based on the interpretation of single input–single output, nonlinear, time–invariant regression models such as:

$$y(t+n) = F\big(y(t+n-1), \cdots, y(t), u(t+n-1), \cdots, u(t)\big) \tag{1}$$

with $t = 0, 1, \ldots$. $u(\cdot)$ and $y(\cdot)$ belong respectively to the bounded input $\mathcal{U}$ and output $\mathcal{Y}$ sets, n is the finite system memory (*i.e.* the model order), and $F(\cdot)$ is a continuous nonlinear function defining a hypersurface from a $\mathcal{A}_n$ to $\mathcal{Y}$, being $\mathcal{A}_n$ the Cartesian product $\mathcal{U}^n \times \mathcal{Y}^n$. The identification of the nonlinear system can be translated to the approximation of its mathematical model given by (1) using a parametric structure that exhibits arbitrary accuracy interpolation properties. A piecewise model defined through the composition of simple models having local validity is the natural candidate to perform this task, as it combines function interpolation properties with mathematical tractability.

In the following, the proposed piecewise structure is recalled and briefly discussed.

2.1 Piecewise Affine Structure

The piecewise model is formed by a collection of parametric submodels of the type:

$$y(t+n) = \sum_{j=0}^{n-1} \alpha_j^{(i)} y(t+j) + \sum_{j=0}^{n-1} \beta_j^{(i)} u(t+j) + b^{(i)}, \quad t = 0, 1, \ldots \tag{2}$$

in which the system operating point is described by the input and output samples $y(t+n-1), \cdots, y(t)$ and $u(t+n-1), \cdots, u(t)$, that can be collected with a vector $\mathbf{x}_n(t) = [y(t), \cdots, y(t+n-1), u(t), \cdots, u(t+n-1)]^T$. The *switching* function $\chi_i\big(\mathbf{x}_n(t)\big), i = 1, \ldots, M$ is

$$\chi_i\big(\mathbf{x}_n(t)\big) = \begin{cases} \chi_i\big(\mathbf{x}_n(t)\big) = 1 \text{ if } \mathbf{x}_n(t) \in \mathcal{A}_n^{(i)} \\ \chi_i\big(\mathbf{x}_n(t)\big) = 0 \text{ otherwise} \end{cases} \tag{3}$$

where $\{\mathcal{A}_n^{(1)}, \ldots, \mathcal{A}_n^{(M)}\}$ is a partition of $\mathcal{A}_n$, whose structure will be characterised in the following.

Thus, the output $y(t+n)$ of the nonlinear dynamic system described by (1) can be approximated by the *piecewise affine model* $f(\cdot)$ in the form:

$$y(t+n) = f\big(\mathbf{x}_n(t)\big) = \sum_{i=1}^{M} \chi_i\big(\mathbf{x}_n(t)\big)[\mathbf{x}_n(t), 1]^T \mathbf{a}_n^{(i)} \tag{4}$$

where the model parameters are collected in the vector $\mathbf{a}_n^{(i)} = [\alpha_0^{(i)}, \ldots, \alpha_{n-1}^{(i)}, \beta_0^{(i)}, \ldots, \beta_{n-1}^{(i)}, b^{(i)}]^T$. It is worthwhile noting that the model is affine in each $\mathcal{A}_n^{(i)}$, $\mathbf{a}_n^{(i)}$ being the affine submodel parameters.

2.2 Local Model Identification

It is assumed that the input–output data $u(t)$ and $y(t)$, $(t = 0, 1, , \ldots, L_i)$ generated by a system of the type of (2) are available. Restricting the investigation also to find order n and parameters $\mathbf{a}_n^{(i)}$ for local model in the form of (2) in region $\mathcal{A}_n^{(i)}$, the following matrix should be defined:

$$X_k^{(i)} = \begin{bmatrix} y(k) & \mathbf{x}_k^T(0) & 1 \\ y(k+1) & \mathbf{x}_k^T(1) & 1 \\ \vdots & \vdots & \\ y(k+N_i-1) & \mathbf{x}_k^T(N_i-1) & 1 \end{bmatrix}$$
$$\Sigma_k^{(i)} = \Big(X_k^{(i)}\Big)^T X_k^{(i)} \tag{5}$$

with $k + N_i - 1 \leq L_i$ and N_i is chosen so that $k + N_i - 1$ is large enough to avoid unwanted linear dependence relationships due to limitations in the dimension of the vector spaces involved.

To determine the model order n in region $\mathcal{A}_n^{(i)}$, it is possible to consider the sequence of increasing–dimension positive definite or positive semidefinite $\big((2k+2) \times (2k+2)\big)$ symmetric matrices

$$\Sigma_2^{(i)}, \quad \Sigma_3^{(i)}, \quad \ldots \quad \Sigma_k^{(i)}, \quad \ldots \tag{6}$$

testing their singularity as k increases. As soon as a singular matrix $\Sigma_k^{(i)}$ is found then $n = k$, and the parameters $\mathbf{a}_n^{(i)}$ describe the dependence relationship of the first vector of $\Sigma_n^{(i)}$ on the remaining ones as

$$\Sigma_n^{(i)} \begin{bmatrix} -1 \\ \mathbf{a}_n^{(i)} \end{bmatrix} = 0 \tag{7}$$

It is worth noting that the vectors $\mathbf{x}_n(0), \mathbf{x}_n(1), \ldots, \mathbf{x}_n(N_i-1)$ in (5) must belong to the region $\mathcal{A}_n^{(i)}$ according to the partition defined in (3). Note also that in the presence of noise the above procedure described to determine order and model parameters would obviously be useless since matrices Σ_k would always be non–singular (positive definite).

In order to solve the problem in a mathematical framework, it is necessary to characterise the noise affecting

the input-output data. Following common assumptions (Kalman [1982], Beghelli et al. [1990]), the noises $\tilde{u}(t)$ and $\tilde{y}(t)$ are assumed additive on input–output data $u^*(t)$ and $y^*(t)$ and region independent, so that:

$$\begin{cases} u(t) = u^*(t) + \tilde{u}(t) \\ y(t) = y^*(t) + \tilde{y}(t). \end{cases} \quad (8)$$

Obviously, only $u(t)$ and $y(t)$ are available for the identification procedure, and moreover every noise term $\tilde{u}(t)$ and $\tilde{y}(t)$ is modelled with a zero–mean white process and is supposed to be independent of every other term. These structures are also commonly known as "Error-In-Variables" models. Under these assumptions, and $\bar{\tilde{\sigma}}_u$ and $\bar{\tilde{\sigma}}_y$ being the input and output noise variances respectively, the generic positive definite matrix $\Sigma_k^{(i)}$ associated with the input-output noise-corrupted sequences can always be expressed as the sum of two terms $\Sigma_k^{(i)} = \Sigma_k^{*(i)} + \bar{\tilde{\Sigma}}_k$ where

$$\bar{\tilde{\Sigma}}_k = \mathrm{diag}[\bar{\tilde{\sigma}}_y I_{k+1}, \bar{\tilde{\sigma}}_u I_k, 0] \geq 0. \quad (9)$$

Thus, it is again possible to determine the order and parameters of the model in region $\mathcal{A}_n^{(i)}$ from the analysis of the sequence of increasing–dimension $\big((2k+2)\times(2k+2)\big)$ symmetric positive definite matrices:

$$\Sigma_2^{(i)}, \quad \Sigma_3^{(i)}, \quad \dots \quad \Sigma_k^{(i)}, \quad \dots \quad (10)$$

The solution of the above identification problem requires the computation of the unknown noise covariances $\bar{\tilde{\sigma}}_u$ and $\bar{\tilde{\sigma}}_y$, that can be achieved solving the following relation:

$$\Sigma_k^{*(i)} = \Sigma_k^{(i)} - \tilde{\Sigma}_k \geq 0. \quad (11)$$

in the variables $\tilde{\sigma}_u, \tilde{\sigma}_y$, where $\tilde{\Sigma}_k = \mathrm{diag}[\tilde{\sigma}_y I_{k+1}, \tilde{\sigma}_u I_k, 0]$. It is worthy to note that the set of values of variables $\tilde{\sigma}_u, \tilde{\sigma}_y$ which make matrix $\Sigma_k^{*(i)}$ positive semidefinite forms a curve.

If the noise characteristics are common to all the regions $\mathcal{A}_n^{(i)}$, since the physical nature of the process generating the noise is independent of the model structure and of the partition of $\mathcal{A}_n$, and all assumptions regarding the Frisch scheme are fulfilled, a common point $(\bar{\tilde{\sigma}}_y, \bar{\tilde{\sigma}}_u)$ in the noise plane exists for the singularity curves. In real applications, we are forced to relax these assumptions, thus no common point can be determined among curves $\Gamma_n^{(i)} = 0$ in the noise plane and a unique solution to the identification problem can be obtained only by introducing a criterion to select a different noisy point for each region as best approximation of the ideal case (Fantuzzi et al. [2002]).

With reference to the identification of the system order n in the i–th region $\mathcal{A}_n^{(i)}$, it must be noted that the $\Gamma_{n+1}^{(i)} = 0$ curve has a single point in common with the $\Gamma_n^{(i)} = 0$ curve in ideal conditions, which corresponds to a double singularity of the matrix $\Sigma_{n+1}^{*(i)}$. In real cases, the order n can be computed finding the point $(\tilde{\sigma}_u, \tilde{\sigma}_y) \in \Gamma_{n+1}^{(i)} = 0$ that makes $\Sigma_{n+1}^{*(i)}$ closer to the double singular condition (*i.e.* minimal eigenvalue equal to zero and the second minimum eigenvalue near to zero). As n is unknown, increasing system orders k must be tested, and the value of k associated to the minimum of the second eigenvalue of the matrix $\Sigma_{k+1}^{*(i)}$ corresponds to the order n. This criterion is consistent as it leads to the common point of the curves when the assumptions of the Frisch scheme are not violated.

Note that since the order n of the piecewise model described by (4) is region independent, it can be determined by choosing k that fulfil the following inequality:

$$\max_{i=1,\dots,M_k} \lambda_k^{(i)} < \epsilon \quad (12)$$

when ϵ is an arbitrary positive constant and $\lambda_k^{(i)}$ is the minimal eigenvalue different from zero of matrix $\Sigma_{k+1}^{*(i)}$. This result led to derive an algorithm for the selection of the model order (Fantuzzi et al. [2002]).

Once the model order n is selected, the parameters $\mathbf{a}_n^{(i)}, i = 1, \dots, M$ can be computed considering for each region a different noise $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$. The values $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$ can be computed by solving an optimisation problem which minimises both the distances between $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$ and $(\bar{\tilde{\sigma}}_u^{(j)}, \bar{\tilde{\sigma}}_y^{(j)})$ with $i \neq j$ and the continuity constraints (Fantuzzi et al. [2002]):

$$\begin{aligned} &J((\bar{\tilde{\sigma}}_u^{(1)}, \bar{\tilde{\sigma}}_y^{(1)}), \dots, (\bar{\tilde{\sigma}}_u^{(M)}, \bar{\tilde{\sigma}}_y^{(M)})) = \\ &= d\left((\bar{\tilde{\sigma}}_u^{(1)}, \bar{\tilde{\sigma}}_y^{(1)}), \dots, (\bar{\tilde{\sigma}}_u^{(M)}, \bar{\tilde{\sigma}}_y^{(M)})\right) + \\ &+ (C_n A_n)^T H C_n A_n \end{aligned} \quad (13)$$

H being a definite positive weighting matrix, and $d(\cdot)$ a distance defined as:

$$\begin{aligned} &d\left((\bar{\tilde{\sigma}}_u^{(1)}, \bar{\tilde{\sigma}}_y^{(1)}), \dots, (\bar{\tilde{\sigma}}_u^{(M)}, \bar{\tilde{\sigma}}_y^{(M)})\right) = \\ &= \sum_{i=1}^{M} \sum_{j=i+1}^{M} \sqrt{(\bar{\tilde{\sigma}}_u^{(i)} - \bar{\tilde{\sigma}}_u^{(j)})^2 + (\bar{\tilde{\sigma}}_y^{(i)} - \bar{\tilde{\sigma}}_y^{(j)})^2}. \end{aligned} \quad (14)$$

It is worthwhile observing that the matrix A_n collects the parameters $\mathbf{a}_n^{(i)}, i = 1, \dots, M$ which depend on $(\bar{\tilde{\sigma}}_u^{(i)}, \bar{\tilde{\sigma}}_y^{(i)})$.

3. FDI BASED ON IDENTIFIED HYBRID MODELS

The problem treated in this section regards the detection and isolation of sensor faults of the process under diagnosis on the basis of the knowledge of the measured uncertain sequences $u(t)$ and $y(t)$.

In the following, it is assumed that the monitored system, depicted in Fig. (1), is described by a model of the type of (4). The term $y(t) \in \Re^m$ is the system output vector, and $u(t) \in \Re^r$ the control input vector. The signal $\varepsilon(t)$ takes into account the modelling error, which is due to process noise, parameter variations, nonlinearities, etc. According to Eqs. (8), in realistic situations the variables $u^*(t)$ and $y^*(t)$ are measured by means of sensors, whose outputs are affected by noise. Neglecting sensor dynamics, faults affecting the measured input and output signals $u(t)$ and $y(t)$ are modelled as:

$$\begin{cases} u(t) = u^*(t) + f_u(t) \\ y(t) = y^*(t) + f_y(t) \end{cases} \quad (15)$$

in which, the vectors $f_u(t) \in \Re^r$ and $f_y(t) \in \Re^m$ are composed of additive signals, which assume values different from zero only in the presence of faults.

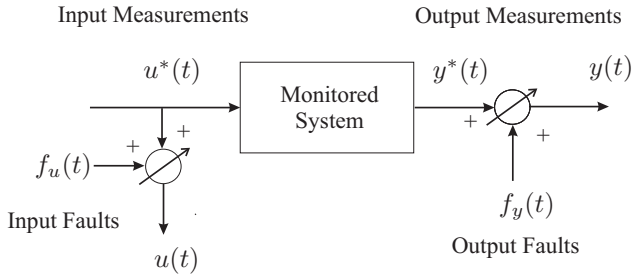


Fig. 1. The structure of the monitored system, with sensor faults.

There are different approaches to generate the diagnostic signals, residuals or symptoms, from which it will be possible to diagnose the considered fault cases. In this work, a model–based approach is used to estimate the outputs of the system from the input–output measurements. As depicted in Fig. (2), residuals can be generated by the comparison of the measured and the estimated outputs:

$$r(t) = \hat{y}(t) - y(t). \tag{16}$$

The symptom evaluation refers to a logic device which

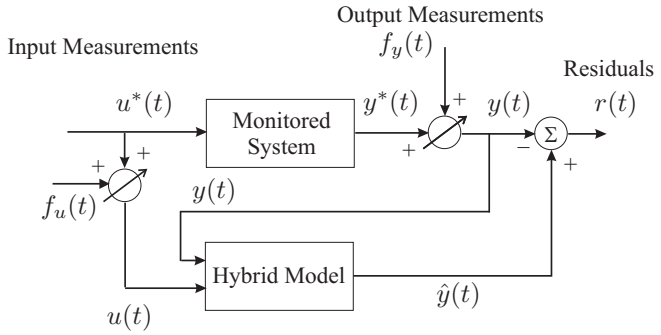


Fig. 2. The residual generation scheme.

processes the redundant signals generated by the first block in order to estimate when a fault occurs, and to univocally identify the unreliable sensors. Faults can be detected by using a simple thresholding logic:

$$|r(t)| \begin{cases} \leq \text{Threshold} & \text{, in fault–free conditions,} \\ > \text{Threshold} & \text{, in faulty conditions.} \end{cases} \tag{17}$$

and, in more detail, according to the following relations:

$$\begin{cases} \bar{r} - \nu\,\sigma_r \leq r(t) \leq \bar{r} + \nu\,\sigma_r & \text{, in fault–free conditions;} \\ \\ r(t) < \bar{r} - \nu\,\sigma_r \\ \text{ or } \\ r(t) > \bar{r} + \nu\,\sigma_r & \text{, in faulty conditions.} \end{cases} \tag{18}$$

where $\bar{r}$ and σ_r represent the mean and the standard deviation values of the fault–free residual $r(t)$, respectively. Due to the presence of modelling errors, ν has to be properly selected in order to achieve the best performances in term of false alarm and missed fault rates (Patton et al. [2009]). In practice, as shown in Section 4, the value of ν can be fixed according *e.g.* to the three–sigma rule.

4. WIND TURBINE MODELLING AND FDI

The three blade horizontal axis turbine considered in this paper works by the principle that the wind is acting on the blades, and thereby moving the rotor shaft. In order to upscale the rotational speed to the needed one at the generator, a gear box is introduced. A more accurate description of the benchmark model can be found in (Odgaard et al. [2009], Odgaard and Stoustrup [2009]). The rotational speed, and consequently, the generated power can be regulated by means of two control strategies: the converter torque and the pitch angle of the turbine blades. In partial load of the wind turbine is controlled to generate as much power as possible. This is achieved by keeping a specific ratio between the tip speed of the blades and the wind speed, which in turn is regulated by controlling the rotational speed and by adjusting the converter torque. In the full power region the converter torque is kept constant and the rotational speed is adjusted by controlling the pitch angle of the blades, which changes the aerodynamic power transfer from the wind to the blades (Odgaard et al. [2009]). The wind turbine model is illustrated in Fig. 3, according to the nomenclature defined in (Odgaard et al. [2009]).

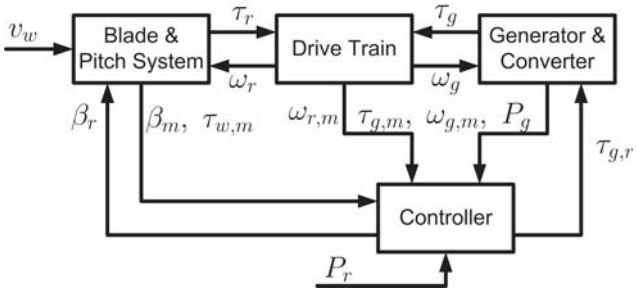


Fig. 3. Logic diagram of the monitored wind turbine.

From the wind turbine model considered here, a number of measurements are considered for identification and FDI purposes. In particular, $\omega_r(t)$ represents the measurement of the rotor speed, $\tau_{gen}(t)$ the torque of the generator controlled by the converter, which is provided with the torque reference, $\tau_{ref}(t)$. The estimated aerodynamic torque is defined as $\tau_{aero}(t)$, $v_{hub}(t)$ is the wind speed $v(t)$ measured at hub position, whilst $\beta_i(t)$ is the pitch angle measurement of the i–th blade ($i = 1, 2, 3$). It is worth noting how the estimate of $\tau_{aero}(t)$ clearly depends on the wind speed $v(t)$, which unfortunately is very difficult to measure correctly. A very uncertain measurement is normally provided, which is used to provide 10 minutes mean values, since this measurement is heavily influenced by measurement noises. In this benchmark, a sample time of $T_s = 0.01$ s. is used (Odgaard et al. [2009]).

4.1 Wind Turbine Model Description

The wind turbine model is briefly recalled in this section in the continuous–time, and subsequently described as identified hybrid prototype in the discrete–time. The following model equations describe the benchmark model presented in (Odgaard et al. [2009]).

The aerodynamic model is defined as in (19):

$$\tau_{aero}(t) = \frac{\rho\,A\,C_p\,(\beta(t),\,\lambda(t))\;v^3(t)}{2\,\omega_r(t)} \tag{19}$$

where ρ is the density of the air, A is the area covered by the turbine blades in its rotation, $\beta(t)$ is the generic pitch

angle of the blades, $v(t)$ the wind speed, whilst $\lambda(t)$ is the tip–speed ratio of the blade, defined as:

$$\lambda(t) = \frac{\omega_r(t)\, R}{v(t)} \tag{20}$$

with R the rotor radius. C_p represents the power coefficient, here described by means of a two–dimensional map (look–up table) (Odgaard et al. [2009]). Equation (19) is used to estimate $\tau_{aero}(t)$ based on an assumed estimated $v(t)$, and measured $\beta(t)$ and $\omega_r(t)$. Due to the uncertainty of the wind speed, the estimate of $\tau_{aero}(t)$ is considered affected by an unknown measurement error, which can be estimated by means of the approach described in Section 2. Moreover, the nonlinearity represented by the relations (19) and (20) motivates the modelling approach suggested in Section 2.

A simple one–body model is used to represent the drive train, in the following form (Odgaard and Stoustrup [2009]):

$$\dot{\omega}_r(t) = \frac{1}{J}\ (\tau_{aero}(t) - \tau_{gen}(t)) \tag{21}$$

where:

$$\dot{\tau}_{gen}(t) = p_{gen}\ (\tau_{ref}(t) - \tau_{gen}(t)) \tag{22}$$

The generator torque $\tau_{gen}(t)$ and the reference $\tau_{ref}(t)$ are in this context transformed to the low speed side of the drive train (rotor side).

These assumptions yield the continuous–time dynamic model of the system under diagnosis in the following form:

$$y(t) = F_c\,(u(t)) \tag{23}$$

with:

$$u(t) = [\tau_{ref}(t),\, v_{hub}(t),\, \beta_i(t)]^T\,, \quad y(t) = \omega_r(t) \tag{24}$$

where $u(t)$ and $y(t)$ are the input and the monitored output measurements, respectively. $F_c\,(\cdot)$ represents the continuous–time nonlinear function representing the discrete–time unknown function F in the form (1), which will be approximated with the discrete–time hybrid prototype (4) from N sampled data $u(t)$ and $y(t)$, with $t = 1, 2, \ldots N$, and using the procedure presented in Section 2.

Finally, the model parameters and the map $C_p\,(\beta,\, \lambda)$ are chosen such that they represent a realistic turbine, which is used as case study, as shown in (Odgaard et al. [2009]).

4.2 Wind Turbine FDI

The proposed methodology was applied to the identification and fault diagnosis of the wind turbine described in Section 4.1. The considered process is shown in Fig. 3, where the considered $r = 3$ inputs are the the reference signal $\tau_{ref}(t)$, the wind speed $v_hub(t)$, and the pitch angle $\beta_i(t)$ measurements ($i = 1, 2, 3$), whilst the $m = 1$ output corresponds to the rotor angular speed $\omega_r(t)$. The available data from the measured inputs and outputs were 440×10^3 samples from normal operating records acquired with a sampling rate of 100 Hz. Because of the underlying physical mechanisms described in Section 4.1, and because of the switching control logic described in (Odgaard et al. [2009], Odgaard and Stoustrup [2009]), the wind turbine system has nonlinear steady state, as well as dynamic characteristics.

Two series of data were acquired from the benchmark process. The first one was used for model identification, and the second one was exploited for the validation task. According to the algorithm recalled in Section 2 for the selection of the model order, the initial value of $k = 1$ and $\epsilon = 10^{-7}$ have been fixed. Under these assumptions, as stated in Section 2, the triangulation of the input–output domain $\mathcal{U} \times \mathcal{Y}$ into simplexes was performed. The partition of the domain was obtained by exploiting the Matlab toolbox for data clustering presented in (Babuška [1998]). The partition of the domain for the wind turbine with $k = 1$ has been achieved by considering the Cartesian product of the intervals $\mathcal{I}_i^{\mathcal{U}}$ and $\mathcal{I}_i^{\mathcal{Y}}$.

A number of $M_1 = 5$ regions were considered for applying (5) and to perform the identification task. Five local affine models were therefore estimated. In this case, $\mathbf{x}_1(t) = [y(t), u^T(t)]^T$, and the data belonging to the domain $\mathcal{U} \times \mathcal{Y}$ have been clustered into the considered partition $\{\mathcal{A}_1^{(1)}, \mathcal{A}_1^{(2)}, \mathcal{A}_1^{(3)}, \mathcal{A}_1^{(4)}, \mathcal{A}_1^{(5)}\}$ ($k = 1$, $M_1 = 5$), the $\Sigma_2^{*(i)}$ matrices ($i = 1, \cdots, 5$) have been computed (Fantuzzi et al. [2002]), and the test of (12) performed. In such a case, $\max_{i=1,\cdots,5} \lambda_k^{(i)} = 2.4765 \times 10^{-9}$. This value is below the selected accuracy ϵ, so the model order can be estimated as $n = 1$. The mean square errors of the output estimation $r(t)$, under no–fault conditions, is 0.0043 with respect to the estimation data, and 0.0044 for the validation data set. The fitting capabilities of the estimated hybrid models can be expressed also in terms of the so–called Variance Accounted For (VAF) index (Babuška [1998]). In particular, the VAF value computed for the estimation data is 97.97%, whilst 89.15% for the validation data. Hence, the fuzzy multiple description seems to approximate the process under diagnosis quite accurately. It is worth noting how the identified model represents a trade–off between simulation accuracy (also dependent on the available data in each region) and structure complexity. Using this hybrid prototype, the model–based approach presented in Section 3 for fault diagnosis is exploited, and applied to the benchmark wind turbine process.

The following simulation results were obtained by considering a fault $f_u(t)$ affecting the $\beta_1(t)$ sensor, whose measurement gets stuck to the constant value 5^o for 100 s., and commencing at the instant $t = 2000$ s. On the other hand, a second fault case, $f_u(t)$, corresponding to the $\beta_3(t)$ sensor stuck at the constant value 10^o, is considered. This fault is active for 100 s., in the period between 2600 s. and 2700 s.

In general, the controller in this wind turbine simulation model, runs in two modes: power optimisation (speed controlled by converter torque), and speed control (speed controlled by pitching blades) (Odgaard et al. [2009], Odgaard and Stoustrup [2009]). A wind speed in the range from approximately $5\ \frac{m}{s}$ to $15\ \frac{m}{s}$ is simulated. This wind speed scenario is used to cover the relevant wind speed region of power optimisation including turbulence.

The considered faults cause alteration of the signals $u(t)$, and therefore of the residuals $r(t)$ given by the predictive model in the form of (4). Residuals indicate fault occurrence according to the logic (17), whether their values are lower or higher than the thresholds fixed in fault–free conditions.

As an example, Fig. 4 represents the fault–free (grey continuous line) and the faulty (black dashed line) residuals $r(t)$ related to the $\beta_1(t)$ pitch sensor fault.

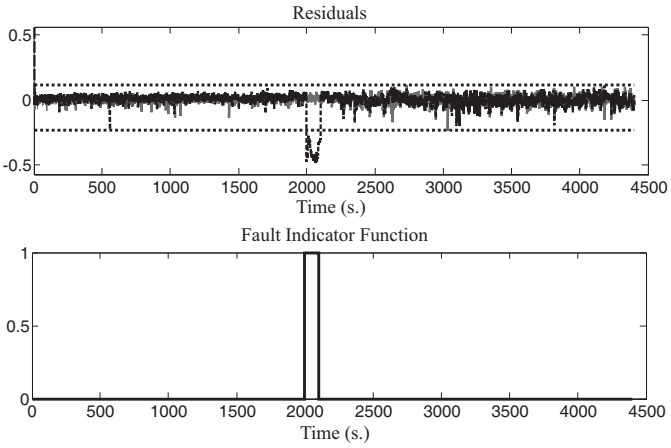


Fig. 4. $\beta_1(t)$ sensor fault residuals $r(t)$, and the fault indicator function.

The fault detection thresholds reported in the relations (17) and (18) are represented as dotted constant lines in Fig. 4. Their values were properly settled by selecting $\nu = 6$, which leads to minimise the false alarm and missed fault rates, while maximising the correct detection and isolation rates. In these conditions, the fault is correctly detected when the corresponding residual signals exceed the thresholds by more than 50 consecutive samples, as indicated by the fault indicator function depicted in Fig. 4. This function is zero in fault–free conditions, and has value one between 2000 s. and 2100 s.

On the other hand, Fig. 5 represents the fault–free (grey continuous line) and the faulty (black dashed line) residuals $r(t)$ related to the $\beta_3(t)$ pitch sensor fault.

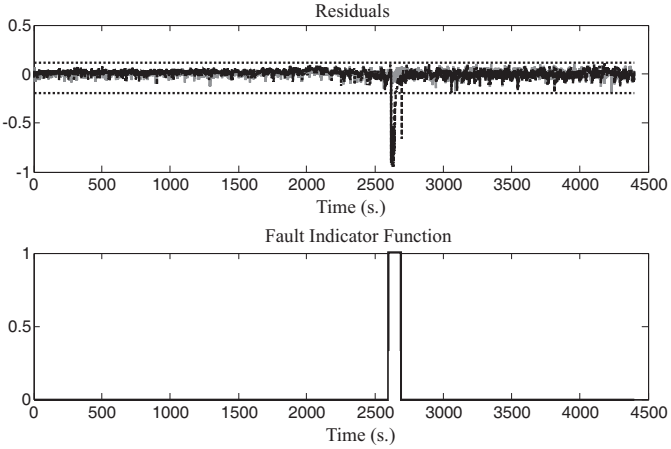


Fig. 5. $\beta_3(t)$ sensor fault residuals $r(t)$, and the fault indicator function.

The fault detection thresholds represented as dotted constant lines in Fig. 5 were obtained with $\nu = 5.5$, leading to minimise the false alarm and missed fault rates, while maximising the correct detection and isolation rates. Also in these conditions, the fault is correctly detected when the residuals exceed the thresholds by more than 50 consecutive samples, as indicated by the fault indicator function of Fig. 5. This function is one between 2600 s. and 2700 s.

Finally, it is worth noting that the developed strategy based on hybrid prototypes allows the detection and isolation of realistic faults using uncertain measurements acquired from the wind turbine simulator.

5. CONCLUSION

This paper proposed a procedure for the identification and fault diagnosis of a wind turbine model using a hybrid prototypes estimated from uncertain input–output measurements. It was assumed that the process under investigation is nonlinear, and these available measurements are normally not very reliable, due to the wind speed uncertain nature. The hybrid modelling approach considered here consists of a collection of several local affine models, each describing a different operating point of the process. The identification algorithm was based on data clustering technique, in order to determine the regions in which measured data can be approximated by local affine dynamic models. The proposed approach provided a model of the wind turbine, which was exploited to generate residuals for fault diagnosis. The effectiveness of these strategies were tested on the data acquired from the wind turbine simulated model, thus allowing the detection and isolation of the wind turbine pitch angle sensor faults.

REFERENCES

Robert Babuška. *Fuzzy Modeling for Control.* Kluwer Academic Publishers, 1998.

S. Beghelli, R. P. Guidorzi, and U. Soverini. The Frisch scheme in dynamic system identification. *Automatica*, 26(1):171–176, 1990.

S.A. Billings and W.S.F. Voon. Structure detection and model validity tests in the identification of nonlinear systems. *IEE Proc.*, 130(4):193–200, July 1983.

J. Chen and R. J. Patton. *Robust Model–Based Fault Diagnosis for Dynamic Systems.* Kluwer Academic Publishers, 1999.

C. Fantuzzi, S. Simani, S. Beghelli, and R. Rovatti. Identification of piecewise affine models in noisy environment. *International Journal of Control*, 75(18):1472–1485, December 2002. Publisher: Taylor and Francis, Ltd.

A. Juditsky, H. Hjalmarsson, A. Beneviste, L. Delyon, B. Ljung, J. Sjöberg, and Q. Zhang. Nonlinear black–box modelling in system identification: a mathematical foundation. *Automatica*, 31(12):1691–1724, 1995.

R. E. Kalman. System Identification from Noisy Data. In A. R. Bednarek and L. Cesari, editors, *Dynamical System II*, pages 135–164. Academic Press, New York, 1982.

J. Korbicz, J. M. Koscielny, Z. Kowalczuk, and W. Cholewa, editors. *Fault Diagnosis: Models, Artificial Intelligence, Applications.* Springer–Verlag, 1st edition, February, 12 2004. ISBN: 3540407677.

Peter Fogh Odgaard and Jakob Stoustrup. Unknown Input Observer Based Scheme for Detecting Faults in a Wind Turbine Converter. In *Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, volume 1, pages 161–166, Barcelona, Spain, June 30 – July 3 2009. IFAC – Elsevier. DOI: 10.3182/20090630–4–ES–2003.0048.

Peter Fogh Odgaard, Jakob Stoustrup, and Michel Kinnaert. Fault Tolerant Control of Wind Turbines

– a Benchmark Model. In *Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, volume 1, pages 155–160, Barcelona, Spain, June 30 – July 3 2009. DOI: 10.3182/20090630–4–ES–2003.0090.

R. J. Patton, F. J. Uppal, S. Simani, and B. Polle. Robust FDI applied to thuster faults of a satellite system. *Control Engineering Practice*, 2009. ACA'07 – 17th IFAC Symposium on Automatic Control in Aerospace Special Issue. Publisher: Elsevier Science. ISSN: 0967–0661. DOI:10.1016/j.conengprac.2009.04.011.

J. Sjöberg, Q. Zhang, L. Ljung, A. Beneviste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky. Nonlinear black–box modelling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Regular Paper

Central sensor cluster simulation for anti-lock-braking system validation using hardware-in-the loop

P. Kret, * **K. J. Burnham,** * **L. Koszalka,** ** **A. Mouzakitis** ***

* *Control Theory and Applications Centre, Coventry University, Priory Street, Coventry, CV1 5FB, UK (e-mail: kretp@ coventry.ac.uk)*
** *Dept. of Systems and Computer Networks, Wroclaw University of Technology, 50-370 Wroclaw, Poland*
*** *Jaguar Land Rover, Engineering Centre, Abbey Rd, Whitley, Coventry, CV3 4LF, UK*

Abstract: in modern automotive electronic stability systems the sensor cluster is a standard component. Essentially, it is able to measure yaw rate, lateral acceleration, roll rate and, optionally, longitudinal acceleration of the vehicle. The sensor cluster is directly connected to the anti-lock-brake (ABS) electronic control unit (ECU) using an independent control network area (CAN). This paper presents a solution for the sensor cluster simulation model to be used within a hardware-in-the-loop (HIL) setup. The model of the sensor cluster simulation exhibits the same functional characteristics as the actual component. The purpose of this is to replace the hardware component of the cluster sensor with an appropriate model, which will satisfy the sensor inputs to the ABS system from a HIL environment.

Keywords: Electronic control unit (ECU), Hardware-in-the-loop (HIL), Sensor cluster, stability control, yaw rate, lateral acceleration

1. INTRODUCTION

Present day vehicles are equipped with various control systems which ensure safe and confident driving. The active safety systems are designed to assist the vehicle when it drifts or deviates from the trajectory of intended direction. During travel, drivers are often unconditionally obliged to make quick steering movements in order to avoid obstacles which suddenly appear on the road. Active safety systems allow the car to be driven safely and in the intended direction and, most importantly they help to avoid all unwanted situations, which may lead to accidents. The main active safety systems in present day vehicles are the Anti-lock Braking System (ABS), the Traction Control System (TCS) and the Electronic Stability Programme (ESP). Active control systems are sophisticated and interact with each other which allows the driver to keep control of the vehicle either in a longitudinal or lateral direction. In order to function properly, the controllers within these systems require ongoing information about the vehicle dynamics. The focus in this paper is on ESP, which utilizes the sensor cluster in order to receive information about the vehicle yaw rate and lateral acceleration. CAN is a network protocol which allows communication of multiple devices on the vehicles (Dawson and Mannisto, 2003). In the automotive industry CAN is widely applied to connect sensors and electronic modules. The sensor cluster sends information to the ABS module using a private CAN bus connection. Based on the information provided by the sensor cluster appropriate control action is generated in order to keep the vehicle on the intended path. The premise of this paper is to present the sensor cluster simulation model, which reconstructs a real-time environment and operates under the same conditions as the true component.

2. VEHICLE HANDLING CHARACTERISTICS

The vehicle can be described as a six degrees of freedom rigid body with translations along the x, y, and z axes, and with rotations about these axes (Wong, 2008). These rotations can be denoted by roll (ϕ), pitch (θ) and yaw (ψ) angles, respectively, and going further, the corresponding angular velocities and accelerations may easily be formulated. Yaw, pitch and roll movements are directed according to the right hand rule of thumb. They go in directions perpendicular to the x, y and z axes, where direction of the given axis is the same as right hand thumb direction. The co-ordinate system for an exemplary vehicle (Range Rover) may be observed in Fig. 1.

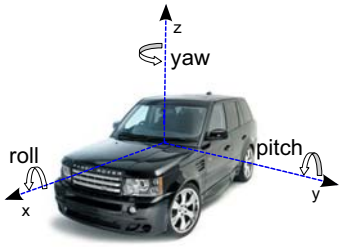


Fig. 1. Range Rover - system of co-ordinates

In this paper attention is focused on vehicle handling stability and vehicle lateral dynamics. The primary mo-

tions associated with handling stability are longitudinal, lateral and rotational motion. Lateral vehicle dynamics are conditioned by a centrifugal force (F_c) which arises while undergoing a turning manoeuvre. The centrifugal force can be defined as an inertial force arising from the natural component of acceleration outwards from the centre of the turn (Wong, 2008) and can be formulated as follows:

$$F_c = \frac{mV_x^2}{R} \tag{1}$$

The centrifugal force is basically formulated by three components: the variable R denotes the cornering radius, m the total mass of the vehicle and V_x is the vehicle forward velocity perpendicular to the centrifugal force (Muvdi et al., 1997). The centrifugal force affects vehicle handling characteristics by acting on the all vehicle tyres. For low speeds $V_x \approx 0$, the centrifugal force is negligible, Eq. 1 tends to zero. But when vehicle is turning at moderate or higher speeds, the effect of the centrifugal force becomes more noticeable. All the vehicle's tyres resist centrifugal force by developing cornering forces in the opposite direction. These are denoted F_{FL}, F_{FR}, F_{RL}, F_{RR} and correspond to the forces generated by the front left, front right, rear left and rear right tyre, respectively.

3. VEHICLE MODEL

The bicycle model is sufficiently detailed to give the basic idea about lateral dynamics and its interaction with the surrounding environment. The main objective of the presented work is to generate signals which can be measured by the sensor cluster. The body coordinate system, denoted by x, y and z with its origin at the Center of Gravity (COG) is introduced in order to describe the vehicle motion. The vehicle dynamics which are active while cornering can be formulated based on physical laws defined by Newton's equations (You and Kim, 1999). The total sum of forces, when considering a Cartesian coordinate system can be expressed as follows:

$$m\dot{V}_x = \sum longitudinal\ forces \tag{2}$$

$$m\dot{V}_y = \sum lateral\ forces \tag{3}$$

and when considering the vehicle motion in terms of the z axis, yaw motion (Ω) and angular moments around the vertical axis can be summarized as:

$$I_z\dot{\Omega} = \sum steering\ torque \tag{4}$$

Strongly related to vehicle handling chracteristics is the tyre model. A simple tyre model for the purposes of this work was sufficient. The simple tyre model neglects longitudinal forces (F_{xn}) and self-aligning moments (M_{zn}) and assumes a linear relationship between the tyre slip angle (α_n) and the cornering stiffness ($C_{\alpha n}$) in effect it considers only the lateral forces (Pacejka, 2002). Here the subscript n is used to denote rear r and front f. These forces can be given by:

$$F_{yn} = C_{\alpha n}\alpha_n \tag{5}$$

$$F_{xn} = 0 \tag{6}$$

$$M_{zn} = 0 \tag{7}$$

while F_x and M_z are equal to zero. The values for the cornering stiffness of the simple tyre model was applied based on the lookup table given in Wenzel (2005). Based on Wong (2008) the physical forces and angular moments while the vehicle is cornering can be expressed as follows:

$$m(\dot{V}_x - V_y\Omega) = F_{xf}cos\delta_f + F_{xr} - F_{yf}sin\delta_f \tag{8}$$

$$m(\dot{V}_y - V_x\Omega) = F_{yr} + F_{yf}cos\delta_f + F_{xf}sin\delta_f \tag{9}$$

$$m(I_z\dot{\Omega}) = l_1F_{yf}cos\delta_f - l_2F_{yr} + l_1F_{xf}sin\delta_f \tag{10}$$

where m represents the vehicle mass, δ_f is the steer angle of the front tyres, l_1 and l_2 distances between COG and front and rear axis, respectively. The variable I_z denotes vehicle moment of inertia around z axis. The relationship $tan\beta = \frac{V_x}{V_y}$ denotes vehicle body side slip. For the small angles there may be a simplification, i.e. $tan\beta = \beta$. The side slip angles for each wheel can be formulated by considering the velocity of the centre points. The slip angle of the front tyre (α_f) and rear tyre (α_r) are formulated based on the yaw rate and velocities in the x and y directions, respectively. The considered bicycle model is a linear model of the vehicle which assumes the low values of lateral slip angle at either the front or rear tyres. Slip angles are assumed to be contained within linear region and can be defined as follows:

$$tan(\alpha_f - \delta_f) = -\frac{l_1\Omega + V_y}{V_x} \tag{11}$$

$$tan(\alpha_r) = \frac{l_2\Omega - V_y}{V_x} \tag{12}$$

Assuming small slip angles, the simplification is valid and final side slip angles can be expressed as follows:

$$\alpha_f = \delta_f - \frac{l_1\Omega + V_y}{V_x} \tag{13}$$

$$\alpha_r = \frac{l_2\Omega - V_y}{V_x} \tag{14}$$

This particular bicycle model considers only the lateral forces acting on the tyres. The lateral forces acting on each tyre are functions of the cornering stiffness and slip angles, which can be given by:

$$F_{yf} = 2C_{\alpha f}\alpha_f \tag{15}$$

$$F_{yr} = 2C_{\alpha r}\alpha_r \tag{16}$$

The above physical laws described by mathematical equations may be now formulated in a state space representation and the representative bicycle model obtained. The formulated bicycle model representation considers two states, these are yaw rate (Ω) and lateral velocity (V_y), which are included in the state vector x (Eq. 17)

$$\dot{x} = Ax + Bu \tag{17}$$

The state space representation of the bicycle model in matrix form can be formulated as follows:

$$\begin{bmatrix} \dot{V}_y \\ \dot{\Omega}_z \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} V_y \\ \Omega \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{18}$$

where the system matrix A is created by the following coefficients:

$$a_{11} = -\frac{2C_{\alpha f} + 2C_{\alpha r}}{V_x(t)m} \tag{19}$$

$$a_{12} = -\frac{V_x^2(t) + 2l_1C_{\alpha f} - 2l_2C_{\alpha r}}{V_x(t)m} \tag{20}$$

$$a_{21} = -\frac{2l_1C_{\alpha f} - 2l_2C_{\alpha r}}{V_x(t)I_z} \quad (21)$$

$$a_{22} = -\frac{2l_1^2C_{\alpha f} + 2l_2^2C_{\alpha r}}{V_x(t)I_z} \quad (22)$$

and input matrix B may be expressed as:

$$b_1 = \frac{2C_{\alpha f}}{m} \qquad b_2 = \frac{2l_1C_{\alpha f}}{I_z} \quad (23)$$

The developed bicycle model representation is linear with time varying parameters within the matrix A. The time varying a_{11}, a_{12}, a_{21} and a_{22} specify the vehicle dynamics and are conditioned by the signals characterizing longitudinal velocity (V_x).

4. SENSOR CLUSTER SIMULATION MODEL

The sensor cluster is an inherent vehicle component which enables the measurement of yaw rate, lateral acceleration, roll rate and optionally longitudinal acceleration, where the output delivered from the sensor cluster is compatible with CAN bus requirements. The sensor cluster component is supplied by the ESP with the switched battery voltage of the car (Continental, 2007). The proposed approach is to create a sensor cluster simulation model which consists of the driver model, vehicle model, transient sensor model, CAN bus interface model and a Kalman filter. The sensor cluster model excluding the Kalman filter can be observed in Fig. 2.

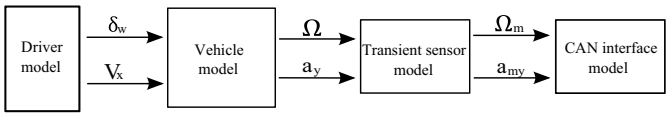


Fig. 2. Sensor Cluster Simulink model

The driver generates the steering wheel angle and longitudinal velocity. Based on these signals the vehicle model generates lateral dynamics, i.e. yaw rate and lateral acceleration. However, these calculated variables are conditioned by the vehicle structure. These calculated variables are processed by the sensor model, which is approximated by a second order system. The processed yaw rate (Ω_m) and lateral acceleration (a_{my}) are progressed to the CAN bus through a CAN interface model. The composed message is readable by a Canalyzer.

4.1 Driver model

The driver model for the off-line simulation generates signals to be sent to the vehicle, with notification that these signals were previously designed to follow a desired path. These are the steering wheel angle (δ_w) and the longitudinal velocity (V_x). The steering wheel angle generated by the driver is expressed in degrees, while longitudinal velocity is expressed in km/h. For on-line simulation when using Real Time Workshop, these signals are generated manually by the operator.

4.2 Vehicle model

The vehicle model is the main part of the overall sensor cluster simulation model. The input signals generated by the driver model and sent to the vehicle model. It is clear from Eq. 20 – 23 that the model parameters are dependent on the longitudinal velocity (V_x). The dynamics generated by the vehicle model are yaw rate (deg/s), yaw acceleration (deg/s^2), lateral velocity (m/s) and lateral acceleration (m/s^2). For the purpose of the work presented here, a Simulink model of the vehicle state space representation was implemented. Various knobs, indicators, scopes were implemented in the dSpace Control Desk to enable control over the sensor cluster simulation model.

4.3 Transient sensor model

To introduce an element of reality, there is a difference between calculated and processed variable at the sensor model. In practice the processed variables may be characterized by a measurement lag (Zimmerschied and Isermann, 2010). Based on sampling information in the system the transient sensor was modelled as a second order critically damped system with the locations of the poles at -10 in the s-plane and with a unity steady state gain, i.e.

$$G_s = \frac{100}{(s+10)(s+10)} = \frac{100}{s^2+20s+100} \quad (24)$$

A representation for a single sensor model (exemplary gyroscope) can be observed in Fig. 3.

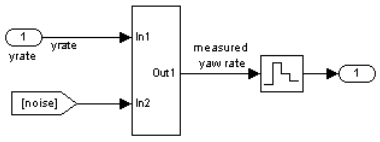


Fig. 3. Transient yaw rate sensor model

The input signals in Fig. 3 are yaw rate and simulated white noise. The output from the transient sensor model is measured discrete yaw rate. Two submodels can also be specified, i.e under non-zero and near zero yaw rate conditions noting that in practice even in the case of zero yaw rate there may be a small residual measurement. The sensor model for non-zero conditions can be observed in Fig. 4.

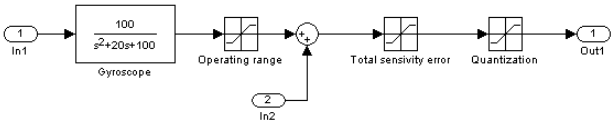


Fig. 4. Transient yaw rate sensor model 2

The signal limitations were accommodated into the model using saturation blocks. The applied saturations are responsible for yaw rate and lateral acceleration limitations and total sensitivity error.

4.4 CAN interface model

The initial version of the model was developed by Guy Curtiss (Jaguar and Land Rover) and then redesigned by

the first author in order to create the sensor cluster simulation model. The redesigned sensor cluster model allows the transmision of data from the transient sensor cluster model to the CAN bus in the form of a multimessage.

4.5 Kalman filter for noise filtering

Every real life system is affected by some external disturbances. The sensor cluster model as a reliable copy of the true component should also be affected by noise. The sensor measurements are disturbed by the effects of calibration, temperature, aging and other environmental aspects. To make the system more realistic, a zero mean Gaussian noise signal was added to the measurements. Kalman filters provide a powerful tool for filtering measurement noise and are widely applied in the automotive industry. Kalman filters not only filter the measurement noise but allow the user to estimate the states and parameters of the system. For the purposes of this work, Kalman filters in various configurations were investigated. The Kalman filter (sensor system), applied in this particular example, consists of a vehicle model and a transient sensor model. The representative system can be seen in Fig. 5.

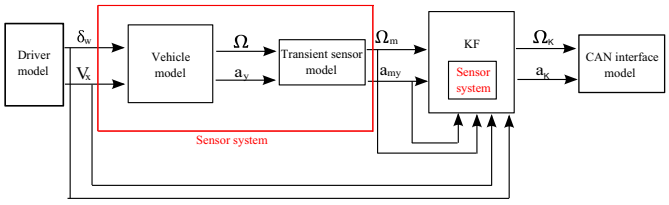


Fig. 5. Kalman filter (KF)

The processed yaw rate and lateral acceleration are fed to the Kalman filter block, which is equivalent to the system output. The input values applied to the Kalman filter are steering wheel angle and longitudinal velocity. The proposed solution is to apply the Kalman filter for each state, i.e yaw rate and lateral acceleration and filter the measurement noise.

4.6 Model integration

In order to execute Real Time Simulation, the Multi-Message Block set with the DS2211 CAN Board has been utilized. The host computer communicates with the dSpace simulator by sending the Simulink model of the sensor cluster simulation model in a C-code representaion. The dSpace simulator interacts with the ECU through the CAN bus and sends information about the vehicle measured vehicle dynamics. The integration of the host computer, dSpace Simulator and ECU can be observed in Fig. 6.

Fig. 6. Sensor cluster integration

The RTI CAN MultiMessage Blockset is an extension for Real-Time Interface and can be used either for combining dSpace systems with CAN communication networks or for configuring these CAN networks. The CAN MultiMessage Blockset may handle complex CAN setups, especially in hardware-in-the-loop applications. Numerous CAN messages can be controlled and configured at the same time just from one single Simulink block. For the purposes of the work here, the transmited message is represented in a frame. The CAN configuration can easily be read in the form of communication matrix description files such as database container files. The overall model is called the Sensor Cluster Simulation model because it consists not only of sensor cluster model but also contains the RTI CAN MultiMessage Blockset which is essential for communication with the CAN. The sensor cluster simulation model can be observed in Fig. 7.

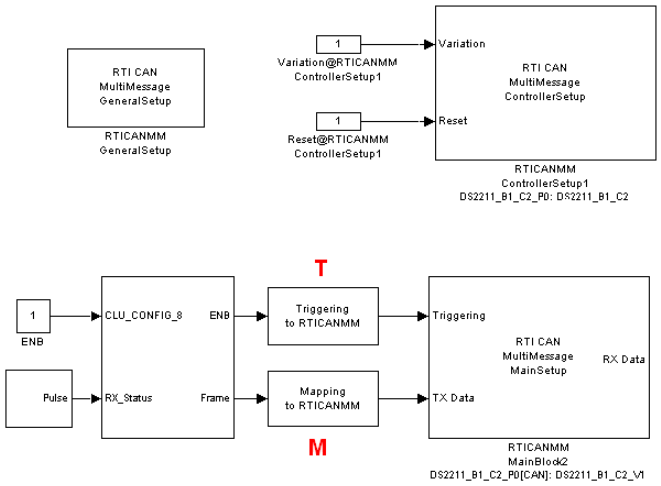


Fig. 7. Sensor Cluster Simulation model

The RTI CAN MultiMessage General Setup Block allows the user to specify the path for model roots, this is the destination folder for RTICANMM_FILES. The RTI CAN MultiMessage Controller Setup Block enables communication with other ECU in the CAN bUS. The private bus connection between the ABS module and the sensor cluster was created (500 kBit/s). MM defines all messages which are to be sent to the CAN bus. The mapping (M) and triggering (T) blocks were used respectively. The triggering block specifies mechanisms for data transmission between Sensor Cluster and CAN bus.

5. SIMULATION STUDY

The simulations during this work were carried out in two stages. Either off-line or on-line experimentation was considered. The off-line simulation was conducted exclusively in the Simulink environment. It allowed creation of the most appropriate sensor cluster model corresponding to real life. Limitations on the measured signals, ranges for the vehicle input signals and techniques for noise filtering were also investigated. Multimessage Block set enabled the sending of multimessage frame from the sensor cluster

model to the CAN bus network. All simulations in the real time environment were carried out on-line.

6. OFF-LINE SIMULATIONS

At the first stage off-line simulations were conducted in order to indicate the difference between the calculated and measured variables from the sensor cluster, with a focus on the model dynamics. An exemplary manoeuvre was presented with a fixed velocity of 70 km/h and a turning radius increasing from 0 up to 30 degrees within 20 seconds. The sensor cluster is a sensitive device and the range for steering wheel angle was chosen exclusively in order to emphasise the difference between the calculated and processed values of yaw rate and lateral acceleration. The resulting yaw rate and lateral acceleration, both generated by the vehicle model and measured can be observed in Fig. 8.

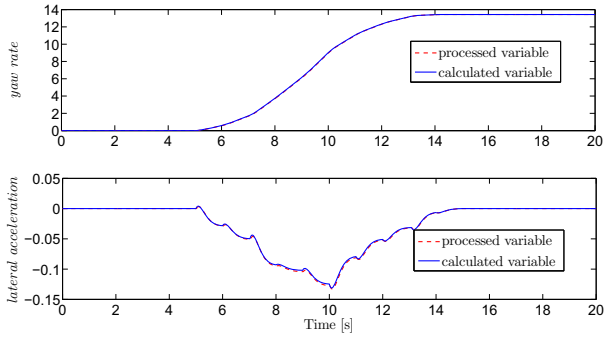


Fig. 8. Yaw rate and lateral acceleration

In this particular case the difference between measured and calculated variables is negligible. Thus, the zoom in Fig. 8 was carried out and the same variables can be observed in Fig. 9.

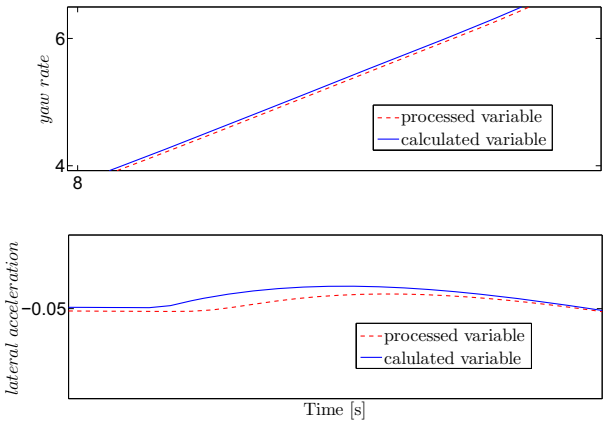


Fig. 9. Zoomed yaw rate and lateral acceleration

As can be seen in both cases the solid line (processed variable) lags the dashed line (calculated variable). The difference between measured and calculated variable can be decreased by applying a more accurate sensor. The more accurate the sensor implies on a higher bandwidth in which the poles are further along the negative axis in the s-plane. A sensor with poles at infinity is the ideal sensor. When considering measurement noise, two main combinations of Kalman filter were tested. These are the double Kalman filter for parameter and state estimation (KFP + KFS) and the extended Kalman filter (EKF). The EKF in comparison to the KFP + KFS is possibly more difficult in terms of implementation because it requires a prior knowledge about system but its advantage is that the states and parameters are estimated within one algorithm. The proposed criteria to assess various configurations of the Kalman filter and to choose the most effective set up is the minimum square error criterion, expressed as follows:

$$MSE = \frac{1}{N - t_0} \sum_{t=t_0}^{N} (y(t) - r(t))^2 \qquad (25)$$

The noise in the system is simulated, thus a noise free signal could be easily formulated. The noise free signal is assumed to be the reference signal and all filtered variables are compared against this particular signal. To indicate the most effective techniques for noise filtering an exemplary experiment is presented. The vehicle accelerates from 0 to 100 km/h within 90 seconds. The longitudinal velocity signal and steering wheel angle within this period can be observed in Fig. 10.

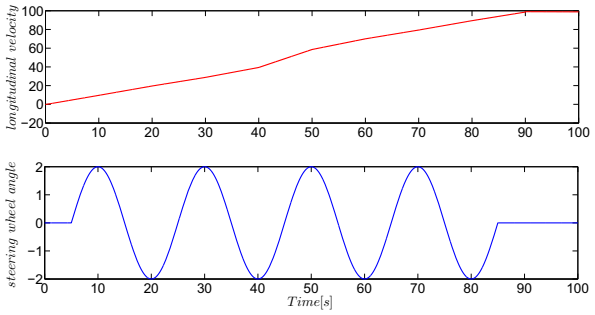


Fig. 10. Longitudinal velocity and steering wheel angle

The generated yaw rate and lateral acceleration can be observed in Fig. 11.

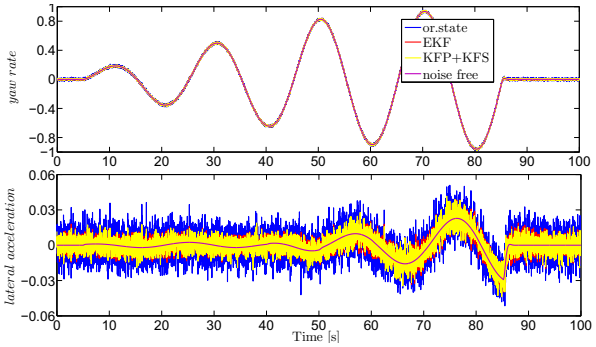


Fig. 11. Yaw rate and lateral acceleration

Examining Fig. 11 the most effective technique for noise filtering would appear to be the double Kalman filter configured for parameter and state estimation. The final validation can be carried out based on numerical analysis from Table 1

Table 1. Margin settings

	$\dot{V}_y \cdot 10^{-4}$	$\Omega \cdot 10^{-4}$
actual state	1.0079	1.0079
EKF	0.4752	0.3191
KFP+KFS	0.4201	0.2870

Based on the results obtained, the most powerful tool for noise filtering seems to be the tandem of Kalman filter for parameter and the state estimation. The extended Kalman filter also gives satisfying results close to those from double Kalman filter, but in practice this solution may be more difficult to implement. Thus, for further investigation the double Kalman filter is chosen.

7. ON-LINE SIMULATIONS

All simulations within this section are carried out on-line, i.e. in real time. There is no need to specify in advance the driver intention, but now the driver may continuously apply various steering wheel angles and longitudinal velocities. Based on these values, the vehicle generates appropriate yaw rate and lateral acceleration, which are measured by the sensor cluster and sent to the CAN bus. For simulation purposes, numerous experiments and layouts for Control Desk were created. The illustrative Control Desk for sensor cluster simulation may be observed in Fig. 7.

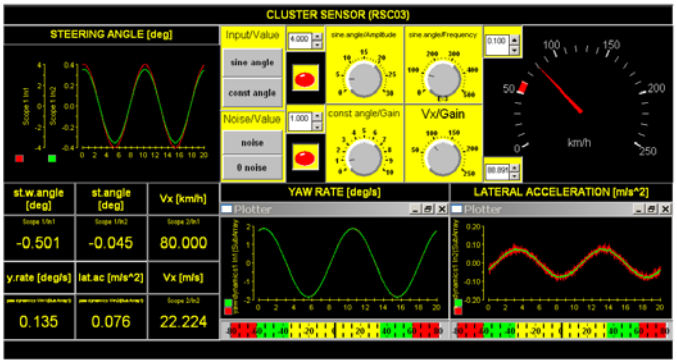


Fig. 12. Illustrative Control Desk layout

The Control Desk enables the behaviour of the driver during the manoeuvre to be reconstructed. The driver applies steering wheel angle and longitudinal velocity. Data is captured and send to the Matlab Workspace

8. CONCLUSION

The paper has presented a model for a sensor cluster simulation. The developed approach consists of a driver model, a vehicle model, a transient sensor model, a CAN interface model and a MultiMesage Blockset, which enables the sensor cluster integration with the CAN bus. The sensor cluster simulation model was developed in the Matlab/Simulink environment. Using HIL and dSpace simulator, real time conditions were reconstructed and several tests carried out. The sensor cluster simulation model communicates with the ABS module sending information about measured vehicle yaw rate and lateral acceleration in the form of a message frame. The simulations were carried out ether off-line or on-line. The off-line simulations allowed limits on the signals generated by the vehicle to be specified and indicated the most effective techniques for noise fltering. When the model was configured, the integration was achieved and on-line experiments conducted.

ACKNOWLEDGEMENTS

The first author wishes to thank supervisors; Prof. Keith J. Burnham from Coventry University, UK, Dr. Alexandros Mouzakitis from Jaguar Land Rover, UK and Dr. Leszek Koszalka from Wroclaw University of Technology, Poland.

REFERENCES

Continental (2007). Customer product specification. Sensor cluster RSC03 (yaw / roll).
Dawson, M. and Mannisto, D. (2003). An overview of controller area network (CAN) technology. *mBus*, 1–16.
Muvdi, B., Al-Khafaji, A., and McNabb, J. (1997). *Dynamics for Engineers*. Springer, New York, 3rd edition.
Pacejka, H. (2002). *Tyre and Vehicle Dynamics*. Butterworth-Heinemann, Burlington, 2nd edition.
Wenzel, T. (2005). *State and parameter estimation for vehicle dynamic control. Phd thesis*. Coventry University,UK.
Wong, J. (2008). *Theory of ground vehicles*. Wiley, 4th edition.
You, S.S. and Kim, H.S. (1999). Lateral dynamics and robust control synthesis for automated car steering. *Institution of Mechanical Engineers*, 21(D), 31–43.
Zimmerschied, R. and Isermann, R. (2010). Nonlinear time constant and dynamic compensation of temperature sensors. *Control Engineering Practice*, 300–310.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Extended Kalman Filter Approach for Road Condition Estimation: a preliminary study

Mariusz Ruta, Keith J. Burnham

Control Theory and Applications Centre, Coventry University
Priory Street, Coventry CV1 5FB, U.K. (e-mail: rutam@uni.coventry.ac.uk)

Abstract: The paper is an initial study into the context of usage of different methods for road condition estimation or monitoring. The aim of the paper is to implement and simulate one of the approaches and evaluate its usefulness for the future investigation.

The paper is focused in particular on the problem of the road condition estimation by means of the extended Kalman filter (EKF) approach, expanded by an additional proportional-integral (PI) block. The approach has been used to estimate a key model parameter, which indicates the road condition (dry, wet, etc.). The tyre/road relationship has been represented by mean of the dynamic nonlinear LuGre model.

1. INTRODUCTION

Safety aspects represent a crucial issue in the case of commercial and domestic vehicles. Modern cars are equipped with active electronic systems which support the driver in vehicle manoeuvrability. Most of the systems such as the electronic stability program (ESP), anti-lock braking system (ABS) or traction control system (TCS) can be found in the majority of current day vehicles (Matusko, Petrovic, & Peric, 2003). All of these systems are focused on increasing the efficiency of the interaction between the tyre patch and the road surface. The tyre patch is a relatively small area where the vehicle has contact with the road. The means by which the tyre interacts with the road surface has an effect on the stability of the whole vehicle. The successful interaction between these two critical interfaces depends on many aspects, i.e. not only tyre condition (tyre pressure, tread condition, tyre material) or road condition (wet or dry, asphalt or snow) but also on the vehicle itself (mass distribution in vehicle, front or/and rear drive) and of course the interaction efficiency and the efficiency of the control systems.

The major factor responsible for the successful tyre/road interaction is the so called friction coefficient, denoted μ. The friction coefficient is described as a normalized relationship between the friction force (F) and the normal force (F_n), which occurs at each wheel, i.e. $\mu = F/F_n$. The coefficient depends on a number of factors, such as roughness of the road, rubber material etc. This paper is focused on the estimation of one of these factors, denoted here θ. This factor (or rather its inverse) indicates the road condition, e.g. is the road dry, wet, snow, ice.

In a case of the vehicle wheel and tire combination the friction coefficient is commonly represented as a function of slip, denoted s. The tyre slip is defined as the difference between momentary linear wheel speed ($r\omega$) and linear speed of the wheel axle/vehicle (v), divided by the greater of two, i.e.

$$s = \frac{r\omega - v}{max(r\omega, v)} \tag{1}$$

where: ω is the angular velocity and r is the wheel radius. In other words, this quantity indicates a per unit deficiency of the tyre-road speed transmission. Note, that negative value of the nominator in the equation (1) and consequently, the negative value of the slip corresponds to braking action (i.e. when linear velocity is gather than the product of the angular velocity and the wheel radius) while the positive slip corresponds to acceleration. A typical relationship between the wheel slip and the friction coefficient for different road conditions is presented in Fig. 1.

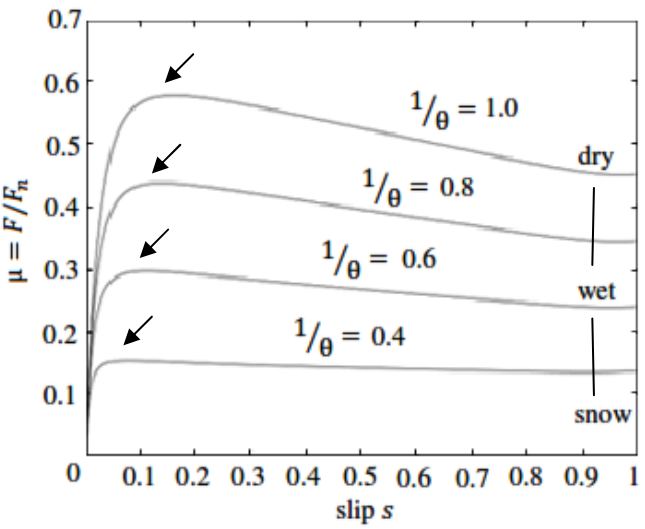


Fig. 1. Typical relationship between friction coefficient and wheel slip including different road conditions (Canudas-de-Wit & Horowitz, 1999).

Fig. 1 shows the optimal operation points, doted by arrows, at the peak values. The operation point for a given road condition corresponds to the maximum value of the friction coefficient, which ensures that the traction (or braking) force

conveyed from the wheel to the road is transferred maximally. This corresponds to the optimal wheel slip, which occurs during acceleration (or braking) and leads to an optimal tyre-road interaction. If characteristics (Fig. 1) are already established and the road conditions are known, knowledge of the tyre-road slip from the equation (1) via measurements will allow the friction coefficient, hence the friction force to be deduced. To achieve that it is necessary to be able to estimate the road condition in advanced to establish the operating curve (Fig. 1). The aim of this work is to continuously estimate the road condition on-line; this condition being defined as a model parameter, which is updated via an extended Kalman filter.

2. TYRE/ROAD FRICTION MODEL

The tyre and road surfaces are by design abrasive. At a given contact patch these two surfaces are in contact through a number of microscopic irregularities having opposing effect providing the required adhesion. This phenomenon is modelled by means of elastic bristles, which form the contact patch of two rigid bodies, i.e. the road and the tyre. The force applied to the rigid body dislocates the elastic bristles, which operate as elastic springs (Fig. 2). Increasing the force gives rise to the friction force and this causes some of the bristles slip.

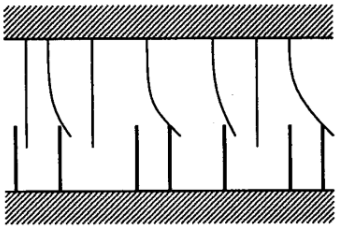


Fig. 2. Illustrating the concept of the elastic bristles model.

The model is based on the linear and rotational dynamical equations of the form (Matusko, Petrovic, & Peric, 2003):

$$m\dot{v} = F + F_n\sigma_2 v_r \tag{2}$$

$$J\dot{\omega} = -rF + u - \sigma_\omega\omega \tag{3}$$

where m is a mass $[kg]$, v is a linear velocity of the vehicle/wheel axle $[m/s]$, σ_2 is the viscous relative damping $[s/m]$, v_r is the relative velocity ($v_r = r\omega - v$) $[m/s]$, J is the moment of inertia of the wheel $[kgm^2]$, ω is the angular velocity $[rad/s]$, r is the wheel radius $[m]$, u is the input torque applied to the wheel axis $[Nm]$ and σ_ω is the coefficient of viscous friction $[Nms]$.

The equations (2) and (3) are used as a basis to formulate the extended model describing the tyre-road friction relationship. One of the most common models is a lumped friction model known as the LuGre model. The primary derivation of this model can be found in (Canudas de Wit, Olsson, Åström, & Lischinsky, 1995) and its successive modification in (Canudas de Wit & Tsiotras, 1999). The LuGre model forms basis of the model used in this work and is represented by the set of equations:

$$F = F_n(\sigma_0 z + \sigma_1\dot{z} + \sigma_2 v_r) \tag{4}$$

$$J\dot{\omega} = -rF_n(\sigma_0 z + \sigma_1\dot{z}) - \sigma_\omega\omega + u \tag{5}$$

$$\dot{z} = v_r - \theta\frac{\sigma_0|v_r|}{g(v_r)}z \tag{6}$$

$$g(v_r) = \mu_C + (\mu_S - \mu_C)e^{-|v_r/v_s|^{0.5}} \tag{7}$$

where σ_0 is the rubber longitudinal lumped stiffness $[1/m]$, z is the bristle deflection $[m]$, σ_1 is the rubber longitudinal lumped damping $[s/m]$, θ is the coefficient indicating the road condition (see Fig. 1), μ_C is the normalized Coulomb friction coefficient, μ_S is the normalized static friction coefficient and v_s is the Stribeck relative velocity $[m/s]$.

The LuGre model used here has been slightly modified. The new coefficients have been included to eliminate the steady-state error (Deur, 2001). Consequently equation (6) is modified and is represented as:

$$\dot{z} = v_r - \left[\theta\frac{\sigma_0|v_r|}{g(v_r)} + \frac{\kappa}{L}r|\omega|\right]z \tag{8}$$

where κ is a lumped model coefficient and L is the tyre-road contact patch length $[m]$.

Combining equations (4), (5), (7) and (8) the model can be rearranged to a more general state-space form:

$$\begin{bmatrix}\dot{v}\\ \dot{\omega}\\ \dot{z}\end{bmatrix} = \begin{bmatrix}0\\ 1/J\\ 0\end{bmatrix}u + \tag{9}$$

$$+\begin{bmatrix}\frac{F_n}{m}\left[\sigma_0 z + \sigma_1\left(v_r - \left[\theta\frac{\sigma_0|v_r|}{g(v_r)} + \frac{\kappa}{L}r|\omega|\right]z\right) + \sigma_2 v_r\right]\\ \frac{1}{J}\left[-rF_n\left(\sigma_0 z + \sigma_1\left(v_r - \left[\theta\frac{\sigma_0|v_r|}{g(v_r)} + \frac{\kappa}{L}r|\omega|\right]z\right)\right) - \sigma_\omega\omega\right]\\ v_r - \left[\theta\frac{\sigma_0|v_r|}{g(v_r)} + \frac{\kappa}{L}r|\omega|\right]z\end{bmatrix}$$

$$y = [0\ \ 1\ \ 0]\cdot[v\ \ \omega\ \ z]^T \tag{10}$$

where $v_r = r\omega - v$.

The model has three states, i.e. linear velocity of the vehicle v, angular velocity of the wheel ω and dislocation of the bristles z. Additionally, the model contains a time-variant parameter θ indicating the road condition. Other parameters are assumed to be time-invariant. The model is represented in continuous time, but, for the purpose of simulation a discretization procedure is implemented. The discrete model is further utilized to implement the EKF. The sampling interval was established experimentally by comparing the continuous and discrete-time model representations. The system with the appropriate sampling interval replicates the continuous representation sufficiently. The sampling interval was set to 1ms.

3. INTRODUCING THE APROACH

1.1 EKF Estimator

The Kalman filter (KF) estimates a state vector via a prediction and a subsequent correction process. The prediction is carried out between sampling instants while correction is carried out at the sampling instant, when a new measurement becomes available. The EKF is a modified linear KF estimator to deal with non-linear systems. The EKF is able to linearize the estimation around the current estimate using partial derivatives (Welch & Bishop, 2006). The non-linear system in state-space form can be represented in the general form:

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \quad (11)$$

$$y_k = h(x_k, v_k) \quad (12)$$

where $x \epsilon R^n$ is the state vector, $y \epsilon R^m$ is the measurement vector, f and h are non-linear functions and random variables w_k and v_k represent process noise and measurement noise, respectively. In practice, however, the noise represented by the variables w_k and v_k is unknown and equations (9) and (10) can be replaced by approximate equations expressed, respectively, as:

$$\tilde{x}_k = f(\hat{x}_{k-1}, u_{k-1}, 0) \quad (13)$$

$$\tilde{y}_k = h(\tilde{x}_k, 0) \quad (14)$$

where $\tilde{x}_k$ and $\tilde{y}_k$ are approximated state and measurement vectors respectively, and $\hat{x}_k$ is an a posteriori estimated state vector at time k. The estimation is carried out without noise variables ($w_k = v_k = 0$). It is a consequence of the assumption that the noise variables w_k and v_k are both zero-mean white noise with certain variance. The functions f and h remain non-linear. To overcome the problem of non-linearity the EKF estimator is applied, which makes use of partial derivatives of both non-linear functions. The derivatives are expressed as (Welch & Bishop, 2006):

$$A_{k_{[i,j]}} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}_{k-1}, u_{k-1}, 0) \quad (15)$$

$$H_{k_{[i,j]}} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\tilde{x}_k, 0) \quad (16)$$

The matrices A, H are Jacobian matrices of partial derivatives and are calculated at each time step. The linearized matrices and the modified linear KF can be subsequently used to implement the EKF as follows:

Time update equations (prediction):

$$\hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}, 0) \quad (17)$$

$$P_k^- = A_k P_{k-1} A_k^T + Q_k \quad (18)$$

Measurement update equations (correction):

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \quad (19)$$

$$\hat{x}_k = \hat{x}_k^- + K_k\big(y_k - h(\hat{x}_k^-, 0)\big) \quad (20)$$

$$P_k = (I - K_k H_k) P_k^- \quad (21)$$

The estimated state vector $\hat{x}_k^-$ is equivalent to $\tilde{x}_k$ in equation (11) and denotes the predicted states based on the *a priori* information available between sampling instants (before correction). Similarly, the notation P_k^- denotes the covariance matrix calculated from the information available between sampling instants, while P_k is the covariance matrix corrected with the information available from the measurements. The process noise and the measurement noise present in the system are assumed to be zero-mean and white with covariance represented by matrices Q_k and R_k respectively.

1.2 PI EKF Estimator

The EKF estimates three states, i.e. linear velocity of the vehicle v, angular velocity of the wheel ω and dislocation of the bristles z. By means of the PI EKF the additional parameter θ_k (indicating the condition of the road surface) is estimated in a similar manner as it was presented in (Matusko, Petrovic, & Peric, 2003). To make the estimation possible the EKF has been modified and expanded by an additional proportional-integral gain term. The modified scheme is known as the PI EKF and a detailed description can be found in (Kim, Shafai, & Kappos, 1989). Equations (17) and (20) are modified to the form:

$$\hat{x}_k^- = f\big(\hat{x}_{k-1}, u_{k-1}, \hat{\theta}_{k-1}, 0\big) \quad (22)$$

$$\hat{x}_k = \hat{x}_k^- + K_k\Big(y_k - h\big(\hat{x}_k^-, \hat{\theta}_{k-1}, 0\big)\Big) \quad (23)$$

and also additional equation is introduced:

$$\hat{\theta}_k = \hat{\theta}_{k-1} + K_I\Big(\mathrm{y}_k - h\big(\hat{x}_k^-, \hat{\theta}_{k-1}, 0\big)\Big) \quad (24)$$

where K_I is a constant proportional-integral gain. The value of the gain effects on the ability of the estimator to capture the variations of the estimated parameter $\hat{\theta}_k$. Large values of the gain produce a more dynamic tracking of the parameter, but at the same time can give rise to an oversensitive response. On the other hand low values of the gain reduce the tracking ability. Thus, the choice of the gain is a trade-off between tracking ability and measurement noise variance. Consequently, the noise variance is required in order to tune the estimator. In (Matusko, Petrovic, & Peric, 2003) an offline optimization was proposed which relied on minimization of the square errors between real parameter θ_k and its estimation. However, to satisfy that requirement the knowledge of the real parameter is necessary. If the relationship between measurement noise variance and gain K_I is available the gain could be potentially changed on-line.

4. SIMULATION RESULTS

For the simulation purpose the discretized tyre/road model represented in the continuous state-space form in equations (7) and (8) has been used. The simulation was carried out with the parameters presented in the Table 1.

The PI EKF is utilized to estimate a road condition parameter θ. During the simulation the actual θ was varied in order to simulate different road conditions. The simulation begines with dry conditions, after 2s it changes to wet and after the next 2s it starts to change linearly into snowy conditions. The input torque applied to the wheel axle was in the form of a

varying square wave function with different amplitudes. The input torque and the real road condition parameter are presented in Fig. 3.

Table 1. Parameters used during simulation.

Parameter	Value	Unit
σ_0	40	$[1/m]$
σ_1	4.9487	$[s/m]$
σ_2	0.0018	$[s/m]$
σ_ω	0.001	$[Nms]$
μ_C	0.5	
μ_S	0.9	
v_s	12.5	$[m/s]$
r	0.25	$[m]$
m	5	$[kg]$
J	0.2344	$[kgm^2]$
F_n	14	$[N]$
κ	1.1	
L	0.25	$[m]$

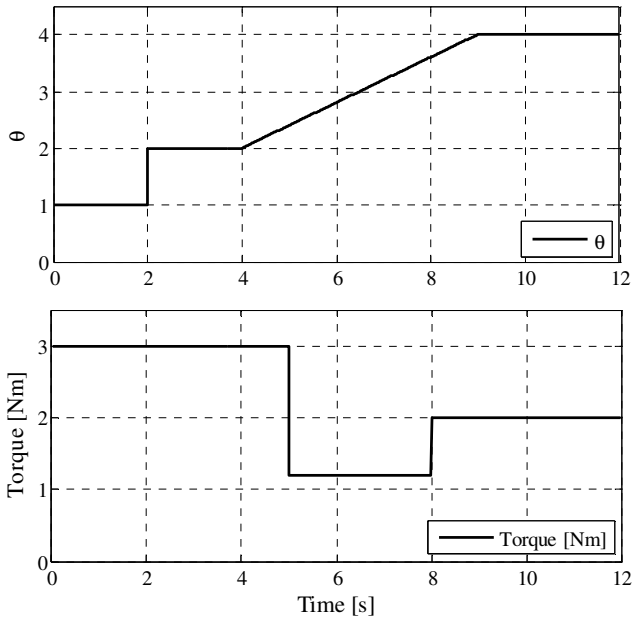


Fig. 3. The real road condition parameter (upper) and input torque (lower).

The optimization of the gain K_I is carried out offline for different values of measurement noise variance. Fig. 4 presents a few experimental data obtained for different noise variances: 0.001, 0.005, 0.01, 0.05, 0.1, 0.5 and 1. These points are subsequently used to approximate a log-linear relationship between the variance R of the zero-mean white measurement noise and the gain K_I.

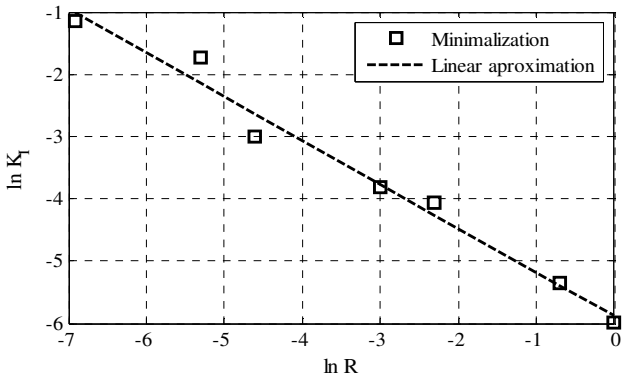


Fig. 4. Relationship between measurement noise variance and gain K_I (logarithmic scale).

The results of a single run simulation are presented in Fig. 5. The initial condition of the parameters in question were set to zero. The measurement noise covariance was set to 0.01 and the corresponding proportional-integral gain was set to 0.05. The sampling interval was set to 1ms.

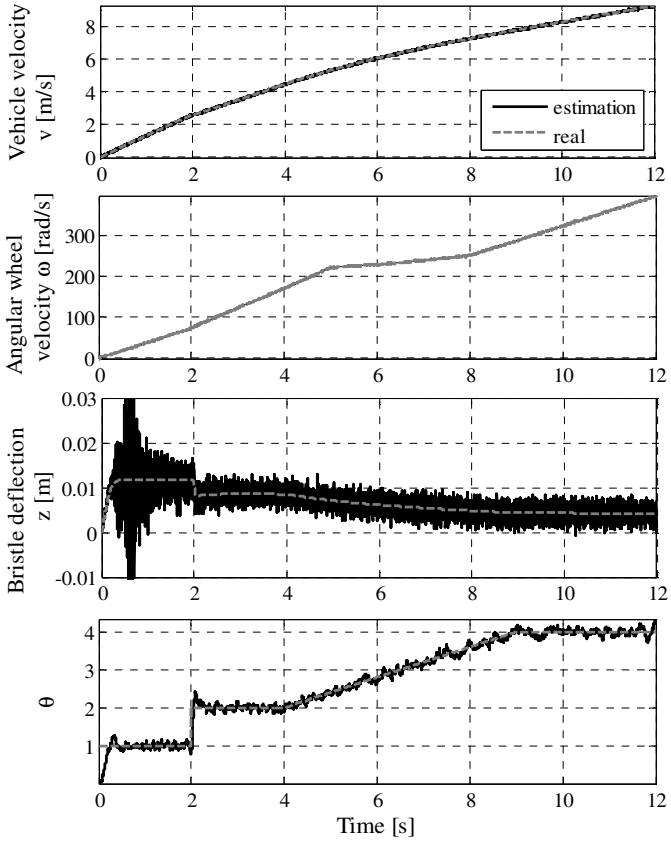


Fig. 5. Vehicle velocity, angular wheel velocity, bristles deflection and road condition parameter θ - real and estimated values (from upper plot to lower plot).

6. CONCLUSIONS

The paper has shown the potential of combining the EKF with an additional PI term in order to estimate the road condition parameter θ. The approach is able to estimate the parameter for varying road conditions with respect to the measurement noise variance. The simulation studies carried

out to data have involved a single wheel station along with a constant normal force. Further work is to include vary vehicle loads due to a dynamic loads distribution and drive cycles involving variable road conditions and different measurement noise. Future work is also to include comparison studies of the benchmark method with alternative approaches.

REFERENCES

Canudas de Wit, C., & Tsiotras, P. (1999). Dynamic Tyre Friction Models for Vehicle Traction Control. *38th IEE Conference on Decision and Control* (pp. 3746-3751). Phoenix, Arisona, USA.

Canudas de Wit, C., Olsson, H., Åström, K. J., & Lischinsky, P. (1995). A New Model for Control of Systems with Friction. *IEEE Transactions On Automatic Control* , 491-425.

Canudas-de-Wit, C., & Horowitz, R. (1999). Observers for tyre/road contact friction using only wheel angular velocity information. *38th IEEE Conference on Decision and Control*, (pp. 3932-3937). Phoenix, Arizona , USA .

Deur, J. (2001). Modeling and Analysis of Longitudinal Tyre Dynamics Based on the LuGre Friction Model. *3rd IFAC Workshop Advances in Automotive Control*, (pp. 101-106). Karlsruhe, Germany.

Kim, K., Shafai, B., & Kappos, E. (1989). Proportional Integral Estimator. *SPIE Signal and Data Processing of Smal Targets* .

Matusko, J., Petrovic, I., & Peric, N. (2003). Application of Extended Kalman Filter for Road Codition Estimation. *Automatika , 21* (3), 59-65.

Welch, G., & Bishop, G. (2006). *An Introduction to the Kalman Filter.* UNC-Chapel Hill, TR95-041.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Diagnostics of distributed faults in ball bearings by means of vibration cyclostationary indicators

G. D'Elia * **S. Delvecchio** * **M. Cocconcelli** ** **G. Dalpiaz** *

* *Engineering Department In Ferrara, University of Ferrara, Ferrara, 44122, Italy (e-mail: gianluca.delia@unife.it).*
** *Unversity of Modena and Reggio Emilia, Reggio Emilia, Italy*

Abstract: This paper deals with the detection of distributed faults in ball bearings. In literature most of the authors focus their attention on the detection of incipient localized defects. In that case classical techniques (i.e. statistical parameters, envelope analysis) are robust in recognizing the presence of the fault and its characteristic frequency. In this paper the authors focalize their attention on bearings affected by distributed faults, due to the progressive growing of surface wear or to low-quality manufacturing process. These faults can not be detected by classical techniques; in fact, in this case the signal does not contain impulses at the fault characteristic frequency, but more complex components with strong non-stationary contents. Distributed faults are here detected by means of advanced tools directly derived from the theory of cyclostationarity. In particular three metrics - namely Integrated Cyclic Coherence (ICC), Integrated Cyclic Modulation Coherence (ICMC) and Indicator of Second-Order Cyclostationarity (ICS_{2x}) - have been calculated in order to condense the information given by the cyclostationary analysis and to help the analyst in detecting the fault in a fast fault diagnosis procedure. These indicators are applied on actual signals captured on a test rig where a degreased bearing running under radial load developed accelerated wear. The results indicated that all the three cyclostationary indicators are able to detect both the appearance of a localized fault and its development in a distributed fault, whilst the usual approach fails as the fault grows.

Keywords: Vibration, Diagnostic, Condition Monitoring, Cyclostationarity, Bearings, Distribuited faults.

1. INTRODUCTION

It is well-known that the vibro-acoustic signature of a machine contains pivotal information about the machine state of health. Because bearings play a pivotal role in the rotating machine scenario, due to their ubiquity and importance, a crowd of signal processing procedures have been developed in order to extract information about incipient localized faults from the measured acceleration signals.

The study of localized failure detection in bearings started over two decades ago, embracing a large number of signal processing techniques that can be roughly subdivided with respect to their pertinence domain, i.e. time, frequency and time-frequency domain.

Concerning time domain several statistical parameters can be evaluated on the time vibration signal, i.e. root mean square (RMS), Kurtosis, Crest factor, Clearance factor and Peak values. In particular Howard (1994), Li and Pickering (1992) show how these parameters are robust to varying bearing operating conditions and good indicators of localized defects. However, Howard (1994) shows that as the defect spreads across the bearing surfaces, the values of these statistical parameters drop back to normal. In the scenario of the time domain techniques for bearing diagnostics, acoustic emission (AE) is gaining ground as a complementary diagnostic tool, thanks to its capabilities in detecting transient elastic waves. Literally, AE is defined as the transient elastic waves generated from a rapid release of strain energy. In particular for bearings, AE results from microshocks and friction between the bearing elements. Therefore, due to the high frequency content of AE signature (about 100 to 1000 kHz), typical mechanical noise (less than 20 kHz) is eliminated. Earliest works, Yoshioka and Fujiwara (1982) and Yoshioka and Fujiwara (1984), have shown that AE parameters can detect bearing defects before they appear in the vibration acceleration range. In Morhain and Mba (2003), Choudhury and Tandon (2000) and Al-Ghamd and Mba (2006) a detailed investigation on AE parameters for a range of several defect conditions in bearings is presented.

Potential defects can also be analyzed by the frequency domain spectrum of the vibration signal, Randall (1987). The most used frequency method for bearing diagnostics is the resonance technique, McFadden and Smith (1984), which deals with the spectrum of the envelope signal, filtered around a structure resonance frequency. However, the effectiveness of this method relies on a suitable choice of the frequency band around the selected resonance, which varies from case to case; this limits the ability of this

technique for automated detection. In order to overcome this drawback, the Spectral Kurtosis (SK) represents an effective and important tool as it is useful for locating the frequency bands with a high amount of impulsiveness, and also for filtering the signal to maximise that impulsiveness. The application of SK to rolling element bearings has been studied through the use of simulated and actual signals by Randall et al. (2004). Another option is the use of the Discrete Wavelet Transform (DWT) algorithm in order to discriminate the frequency range having the highest temporal kurtosis level, Parameswariah and Cox (2002). Significant potential in bearing fault diagnosis can also be achieved by time-frequency domain methods. Among all the time-frequency analysis methods, wavelets have been established as the most widespread tool in bearings diagnosis, due to their flexibility and their efficient computational implementation. The work of Mori et al. (1996) was probably one of the first to use the wavelet for the bearing diagnosis.

All the techniques herein described treat the signals as if they were statistically stationary. However, the signals of interest often consist of combinations of stationary and polyperiodic signals, which are called polycyclostationary signals. This typically requires that the random signal is modeled as cyclostationary, i.e. the statistical parameters vary in time with single or multiple periodicities. The majority of the initial works based on cyclostationary and spectral correlation, has been carried out by Gardner (1986). The work of McCormick and Nandi (1998) was one of the first to apply the principles of cyclostationarity to bearing signals. In particular, Randall et al. (2001) demonstrate the relationship between the classical envelope analysis and spectral correlation analysis. Further works, Antoniadis and Glossiotis (2001) and Antoni and Randall (2005), show the potential of the cyclostationary analysis in bearing diagnostics; in fact, it can better highlight the modulation mechanism present in the vibration response of this type of system. In particular Antoni (2007) discusses which cyclic spectral tool is the most suitable for the localized fault detection in bearings.

The outlined techniques mainly deal with localized defects. However, as a fault grows, becoming a distributed fault, no sharp impulses are generated but a more complex signal with strong non-stationary contents. The same type of signal can be produced from low-quality bearings with high roughness and geometrical irregularities. In this condition the usual signal processing techniques fail, and the characteristic fault frequencies can not be extracted from the vibration signals. On the other hand, global indicators of the bearing health can be used, i.e. the vibration energy or the root mean square value of the vibration signals. Unfortunately, these global indicators do not specify where the fault is located (e.g. on the inner race rather than the outer race).

The aim of this work is to apply cyclostationary metrics for the identification of both the appearance and the growth of distributed faults in ball bearings, in order to overcome the pitfall of the usual approaches.

The paper is organized as follows. After the brief introduction and the problem statement given in this Section, the cyclostationary metrics are outlined in the next one. Finally in Section 3, the diagnostic capabilities of the cyclostationary approach is discussed on the basis of experimental results.

2. CYCLOSTATIONARY TOOLS

Non-stationary signals can be defined as signals which satisfy a non-property, i.e. they do not satisfy the property of stationarity. It is not possible to define a general theory which treats non-stationary signals. The non-stationary behavior of each signal has to be individually evaluated.

In the case that a signal presents periodic energy variations a particular class of non-stationary signals can be defined: the cyclostationarity signals.

Mathematically, a signal $x(t)$ that satisfies the periodicity of the first two moments can be respectively defined as cyclostationary at order 1 (periodicity of mean) and order 2 (periodicity of instantaneous autocorrelation function).

Since this paper is focalized on the second-order cyclostationary content of the acquired signals we refer to the generalized expression of the autocorrelation function, Gardner (1986):

$$\mathrm{R}_x(t,\tau) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x\Big(t+nT+\frac{\tau}{2}\Big)x^*\Big(t+nT-\frac{\tau}{2}\Big) \tag{1}$$

where n is an integer and T is the period. Since, for a second-order cyclostationary signal the autocorrelation function is periodic, it can be expanded into a Fourier series. The coefficients of this decomposition are the cyclic autocorrelation function, given by:

$$\mathrm{R}_x^{\alpha}(\tau) = \lim_{T\to\infty} \int_T x\Big(t+\frac{\tau}{2}\Big)x^*\Big(t-\frac{\tau}{2}\Big)e^{-j2\pi\alpha t}dt \tag{2}$$

where α is called the cyclic frequency.

In order to understand if one signal presents second order cyclostationary behaviour it can be useful to compute the Fourier Transform of the cyclic autocorrelation function and obtain the so-called spectral correlation density (SCD) that reads as:

$$S_x(f,\alpha) = \int_{-\infty}^{+\infty} R_x^{\alpha}(\tau)e^{-j2\pi f\tau}d\tau \tag{3}$$

This function depends on two frequencies: the spectral frequency f and the cyclic frequency α. When $\alpha = 0$ the SCD is equal to the power spectral density of the signal $x(t)$, whilst at other values of α the SCD is the cross-spectral density of the signal $x(t)$ and its version shifted by frequency α.

For diagnostic purposes an interesting tool derived from the SCD is the (squared-magnitude) Cyclic Coherence (CC) defined as, Antoni (2009):

$$|\gamma_x(f,\alpha)|^2 = \frac{|S_x(f,\alpha)|^2}{S_x(f+\alpha/2)S_x(f-\alpha/2)} \tag{4}$$

As proven by Antoni this represents a correlation coefficient between two spectral frequencies with different phases.

Another tool for displaying the cyclostationary properties of a signal is the Cyclic Modulation Coherence (CMC), which is a power-normalised version of the Cyclic Modulation Spectrum (CMS) that reads, Antoni (2009):

$$P_x^{\alpha}(f;\Delta f) = \lim_{T\to\infty} \int_T |x_{\Delta f}(t;f)|^2 e^{-j2\pi\alpha t} dt \qquad (5)$$

where $x_{\Delta f}(t;f)$ is the filtered version of signal $x(t)$ through a frequency band of width Δf centered on frequency f. Therefore, the CMC is defined as:

$$CMC(f,\alpha) = P_x^{\alpha}(f;\Delta f)/P_x^{0}(f;\Delta f) \qquad (6)$$

In order to get simple indicators that can be useful for quality control purposes both CMC and CC have been reassumed in two metrics: Integrated Cyclic Modulation Coherence (ICMC) and Integrated Cyclic Coherence (ICC) derived from CMC and CC respectively. These metrics have been calculated integrating both coherences along variable f. This paper is focused on the evolution of the second-order cyclostationary content with bearing conditions. Therefore, the above defined metrics are evaluated on the residual signal, which reads as the difference between the signal and its first-order cyclostationary part. In practice, this corresponds to consider these metrics based on the second-order cumulant (i.e. the autocovariance function) instead of the second-order moment (i.e. the autocorrelation function).

The last implemented cyclostationary tool is the Indicator of Second order cyclostationarity (ICS_{2x}) outlined in Raad et al. (2008). This indicator tries to quantify the distance of a second-order cyclostationary process from the closest stationary process having a similar power spectral density giving an indication of the presence of second-order cyclostationary components within a signal. It is defined as (Antoni (2009)):

$$ICS_{2x} = \sum_{\alpha\in\mathcal{A}} \frac{\left|\lim_{T\to\infty}\int_T |x^R(t)|^2 e^{-j2\pi\alpha t} dt\right|}{\lim_{T\to\infty}\int_T |x^R(t)|^2 dt} \qquad (7)$$

where $\mathcal{A}$ is the set of all possible cyclic frequency α and $x^R(t)$ is the residual signal. As reported by Raad et al. (2008), this is a cumulant based estimator.

3. EXPERIMENTAL RESULTS

An experimental campaign was carried out on a ball bearing in order to obtain distributed fault on the outer race, by using a test-bed composed of an asynchronous 4-pole motor which moves a shaft by means of a driving belt. The shaft is supported by a couple of cone-shaped bearings, Fig. 1 (a).

The bearing under test is a double-row self-aligning type SKF 1205; it is cantilever mounted on this shaft at the opposite to the pulley. A radial external load supplied

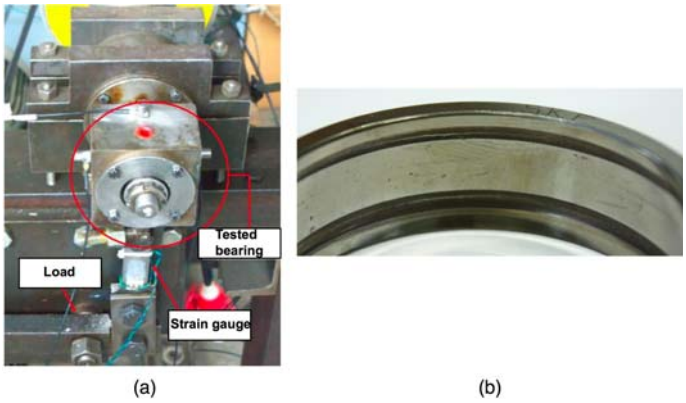


Fig. 1. (a) Test Rig; (b) Bearing outer race at the end of the test.

by a leverage system acts on the test bearing. In the present test, the bearing was externally loaded with a force of $1962\,N$, while the shaft was rotating at $26.6\,Hz$. The bearing was degreased in advance in order to accelerate the wear process and then mounted on the test machine. Three accelerometers Brüel & Kjær were used to measure the vibration signal ($0.1\,Hz \div 16.5\,kHz$) and they have been connected to a LMS acquisition board. The vibration signals were acquired each 15 minutes with an acquisition time of 2 minutes, obtaining 21 acquisitions in total. The data were later post-processed using Matlab software. The sampling frequency was $51.2\,kHz$. At the end of the test the bearing was unmounted to check the status of the surfaces. The outer race of the bearing presented a groove corresponding to the passage of the balls, see Fig. 1 (b). The groove length took more or less half of the outer race circumference.

The expected frequencies of the typical faults are computed from the usual formulae, McFadden and Smith (1984). Their values for the bearing under test are collected in Table 1. Since the damage mainly involves the outer race in the actual case, the outer race fault frequency is briefly referred as 'fault frequency' in the following.

Table 1. Characteristic fault frequencies

Description	Symbol	$[Hz]$
Rotation frequency	f_r	26.6
Outer race fault frequency	f_o	127
Inner race fault frequency	f_i	192
Cage fault frequency	f_c	10.6
Ball fault frequency	f_b	121

In this paper it is considered only one of the three acceleration signals: it deals with the vibration signal that has be proven to be less affected by the transmission path. Fig. 2 depicts that time signal captured during acquisitions n. 1 (first) and 21 (last) for a complete revolution of the shaft. As expected, the overall amplitude level strongly increases from the first to the last acquisition, but no impulsive content can be observed. Actually a distributed fault is not related to an impulsive content, but to an increase of the signal energy. Therefore, the evaluation of the RMS value can be an useful parameter for condition monitoring.

As depicted in Fig. 3 the RMS value gives an alarm on acquisition 5 where probably the condition of the bearing

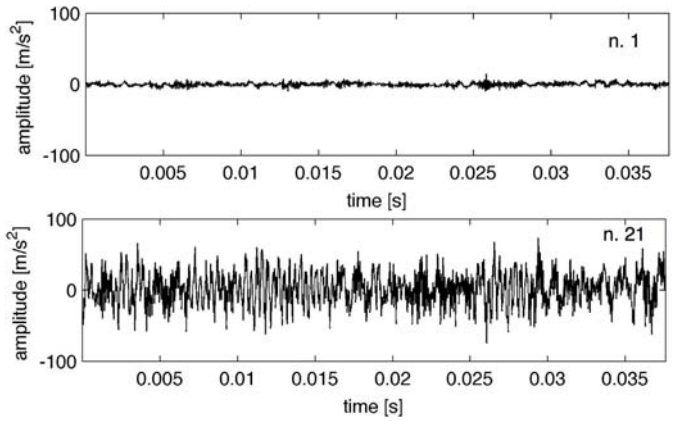


Fig. 2. Time signal for one shaft rotation: n. 1 first acquisition; n. 21 last acquisition.

is changed. However the RMS remains high until the end of the acquisitions giving no information about the evolution of the fault. In addition, due to its global nature the RMS cannot give any information about the fault position.

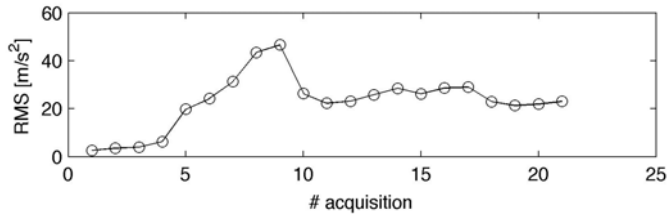


Fig. 3. RMS trend for acquisitions 1 to 21.

In order to better investigate the fault behavior, the classical envelope analysis is carried out by following these steps:

- band-pass filtering around a suitable frequency;
- computation of the analytical signal, which is a complex quantity having the acquired signal as real part and its Hilbert Transform as imaginary part;
- computation of the Spectrum of the absolute value of the analytical signal.

The first step, i.e. the choice of a proper frequency band for signal filtering, plays a pivotal role in this approach. As a matter of fact, the frequency band with the highest amount of impulsive components have to be chosen in order to proper highlight the fault signature. In this work, the choice of the optimal frequency band is done by means of the Discrete Wavelet Transform (DWT) technique in order to extract the frequency band with the maximum time domain kurtosis value (TK), which is strictly related to the impulsive content of the signal. As mentioned in Parameswariah and Cox (2002), the DWT technique provides a multi-resolution representation of signals with different time-frequency scales. The calculation is based on the Mallats algorithm (or pyramid algorithm) and it is implemented in Matlab. This is a very practical filtering algorithm that enables the wavelet transform to be computed in the form of a discrete wavelet transform. As described in Parameswariah and Cox (2002), one can look at Mallats algorithm (pyramid algorithm) either as a way of calculating wavelet coefficients at various scales or as a filter bank for processing discrete-time signals. The pyramid algorithm operates on a finite set of input data whose length is an integer power of two. These data are passed through two convolution functions that essentially act as filters. After the implementation of the DWT algorithm the band presenting the maximum temporal kurtosis has been considered as the optimal frequency band. After the application of this algorithm the $19 \div 25\,kHz$ frequency band is chosen. Once computing the spectrum of the absolute value of the analytical signal the amplitude of the faulty frequency component (i.e. $127\,Hz$) is plotted for all acquisitions (Fig. 4).

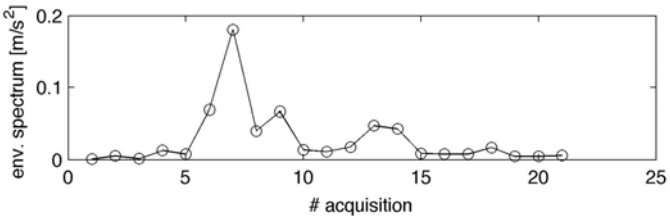


Fig. 4. Amplitude of the fault frequency component in the Envelope Spectrum of the raw signal: trend for acquisitions 1 to 21.

It can be noted that the envelope analysis confirms what found by the RMS value analysis trend. Moreover the envelope can identify the presence of a localized fault on the outer race around acquisition n. 5. In more details this localized fault gives origin to very slight but detectable impacts. The envelope is sensitive to these impacts showing the ball passing frequency on the incipient localized defect. Unfortunately this technique cannot supply further information concerning the increase of severity and extension of the wear. As a matter of fact, when the localized defect on the outer race grows, becoming a distributed fault, no impulses are generated; thus, this technique can not highlight the fault evolution.

At this stage the cyclostationarity analysis has been applied as a powerful tool in order to obtain information concerning the fault evolution. As said before this type of analysis is well suited in describing the vibration response signal captured from a ball-bearing with a distributed fault. In fact, it is reasonable that this signal has non-stationary properties with a second-order cyclostationary content, due to the periodic variation of the bearing configuration. For example, the action of the balls passing on a distributed fault on the outer race produces a cyclostationary vibration with fundamental cyclic frequency corresponding to the ball passing frequency. This is the reason why the cyclostationary tools such as Cyclic Coherence (CC) and Cyclic Modulation Coherence (CMC) were implemented. In order to condense the information derived from these two distributions in one simple metric, their Integrated versions along the frequency axis have been computed obtaining the above-described Integrated Cyclic Coherence (ICC) and Integrated Cyclic Modulation Coherence (ICMC) for all 21 acquisitions.

The use of both Coherences gives more advantages with respect of the quantities ($S_x(f,\alpha)$ and $P_x^{\alpha}(f;\Delta f)$) from which they have been derived. As a matter of a fact, the Coherences are not affected by the scale effects, so the values they assume are not dependent on the increase of the overall signal level due to the distributed fault increase.

In that circumstance only the amplitude of the cyclic frequency of the fault has an influence on the Coherence value.

Fig. 5 depicts the ICC for in the $110 \div 200\,Hz$ cyclic frequency range for all the acquisition. It is possible to notice that only the cyclic frequency of the outer race fault is present, which is a clear information concerning the fault location. Therefore, hereunder only the cyclic frequency of the outer race fault is investigated.

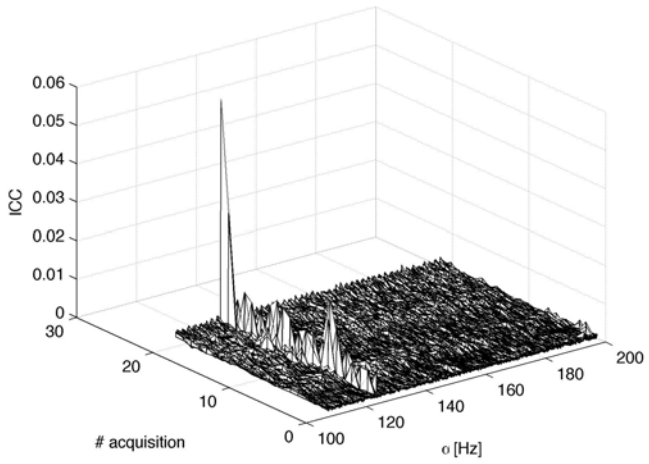


Fig. 5. Integrated Cyclic Coherence (ICC) for acquisitions 1 to 21.

The amplitude of the faulty cyclic frequency (i.e. $127\,Hz$) for both ICC and ICMC has been tracked for all 21 acquisitions and depicted in Fig. 6 and Fig. 7 respectively.

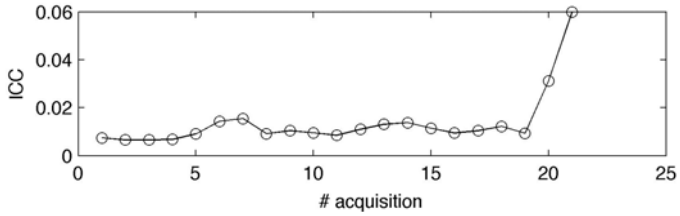


Fig. 6. Integrated Cyclic Coherence (ICC): trend of the fault cyclic frequency amplitude for acquisitions 1 to 21.

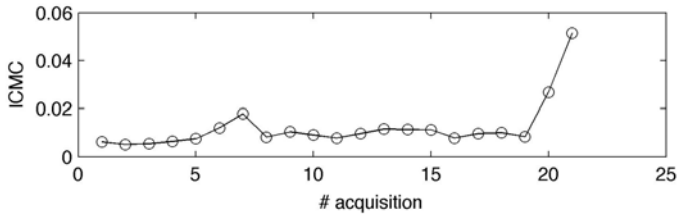


Fig. 7. Integrated Cyclic Modulation Coherence (ICMC): trend of the fault cyclic frequency amplitude for acquisitions 1 to 21.

For both indicators we can observe the same trend: a first part in which the trend increases due to the localized fault appearance (around acquisition n. 5); a second part where the indicators assume lower but almost constant values while the race area interested by the fault becomes larger and a third part (around acquisition n. 19) where the increasing of the fault severity causes an increase of the indicator amplitudes. These indicators bring extra information than the envelope analysis since they are able to monitor both the fault appearance and the growth.

Finally, the ICS_{2x} is evaluated in order to have a simple cyclostationary metric that can track the fault evolution. In particular, this metric is computed in the cyclic frequency range covering the first two fault harmonics. Fig. 8 depicts the trend of this cyclostationary metric for all the acquisitions. It is possible to notice as this metric shows the same results obtained by using ICC and ICMC, highlighting both the fault appearance and the growth. In particular this metric is simpler to implement, and therefore it can be an useful alternative compared to the previous ICC and ICMC for distributed fault diagnostics.

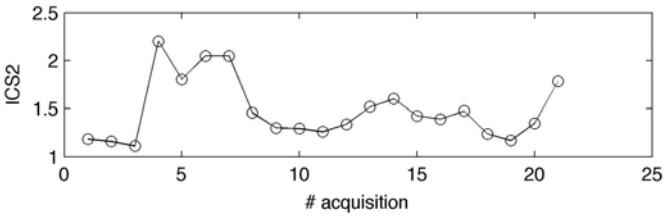


Fig. 8. ICS_{2x} trend for acquisitions 1 to 21.

4. CONCLUSION

This paper proposes a cyclostationary approach in order to identify distributed faults in ball bearings. The effectiveness of this approach is assessed through an experimental test: a degreased bearing running under radial load developed accelerated wear, while the vibration signal is periodically captured during the bearing life in order to monitor its deterioration. Classical and cyclostationary techniques are then applied to the signals. The results indicate that the usual approach can detect the appearance of the fault but can not track the successive growth. On the contrary, cyclostationary tools like ICC and ICMC are able to detect both the appearance of a localized fault and its development in a distributed fault. Therefore it seems that the use of the cyclostationarity approach can be used not only with localized faults but also with distributed ones. In addition, since the ICS_{2x} gives the same results as the other metrics and it is simpler to be implemented, it can be considered as a robust tool for applications in on-line monitoring. The implementation of this metic in on-line detection systems and its practical performance should be taken into account in the development of this research.

Future works are also needed in order to apply the detection of bearing distributed faults in more complex machines, such as gearboxes. In particular, the effectiveness of the studied cyclostationary metrics will be tested on signals acquired on gearbox during the bearing life. As a matter of fact, in these complex machines the vibration due to the faulted bearing is generally hidden by signal components rising from the tooth meshing. Therefore, more complex techniques could be required in order to extract features related to the bearing condition.

ACKNOWLEDGEMENTS

This work has been developed within the Laboratory of Advanced Mechanics (MECH-LAV) of Ferrara Technopole, realized through the contribution of Regione Emilia Romagna - Assessorato Attivitá Produttive, Sviluppo Economico, Piano telematico - POR-FESR 2007-2013, Attivitá I.1.1.

REFERENCES

Al-Ghamd, A.M. and Mba, D. (2006). A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size. *Mechanical Systems and Signal Processing*, 20, 1537–1571.

Antoni, J. (2007). Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions. *Journal of Sound and Vibration*, 304, 497–529.

Antoni, J. (2009). Cyclostationary by examples. *Mechanical Systems and Signal Processing*, 23, 987–1036.

Antoni, J. and Randall, R.B. (2005). On the use of the cyclic power spectrum in rolling element bearings diagnostic. *Journal of Sound and Vibration*, 281, 463–468.

Antoniadis, I. and Glossiotis, G. (2001). Cyclostationary analysis of rolling-element bearing vibration signals. *Journal of Sound and Vibration*, 248, 829–845.

Choudhury, A. and Tandon, N. (2000). Application of acoustic emission technique for the detection of defects in rolling element bearings. *Tribology International*, 33, 39–45.

Gardner, W.A. (1986). The spectral correlation theory of cyclostationary time-series. *Signal Processing archive*, 11, 13–36.

Howard, I. (1994). A review of rolling element bearing vibration detection, diagnosis and prognosis. *Defense Science and Technology Organization*.

Li, C.Q. and Pickering, C.J.D. (1992). Robustness and sensitivity of non-dimensional amplitude parameters for diagnosis for fatigue spalling. *Condition Monitoring and Diagnostic Technology*, 2, 81–84.

McCormick, A.C. and Nandi, A.K. (1998). Cyclostationary in rotating machine vibrations. *Mechanical Systems and Signal Processing*, 12, 225–242.

McFadden, P.D. and Smith, J.D. (1984). Vibration monitoring of rolling element bearings by the high-frequency resonance techniquea review. *Tribology International*, 17, 3–10.

Morhain, A. and Mba, D. (2003). Bearing defect diagnosis and acoustic emission. *Proceedings of the Institution of Mechanical Engineers, Part J, Journal of Engineering Tribology*, 217, 257–272.

Mori, K., Kasashima, N., Yoshioka, T., and Ueno, Y. (1996). Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals. *Wear*, 195, 162–168.

Parameswariah, C. and Cox, M. (2002). Frequency characteristics of wavelets. *IEEE Transactions on Power Delivery*, 17, 800–804.

Raad, A., Antoni, J., and Sidahmed, M. (2008). Indicators of cyclostationarity: Theory and application to gear fault monitoring. *IEEE Transactions on Power Delivery*, 22, 574–587.

Randall, R.B. (1987). *Frequency analysis*. Brüel Kjær, London.

Randall, R.B., Antoni, J., and Chobsaard, S. (2001). The relatioship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, 15, 945–962.

Randall, R.B., Antoni, J., and Sawalhi, N. (2004). Application of spectral kurtosis to bearing fault detection in rolling element bearings. In *Eleventh International Congress on Sound and Vibration*. St.Petersburg.

Yoshioka, T. and Fujiwara, T. (1982). A new acoustic emission source locating system for the study of rolling contact fatigue. *Wear*, 81, 183–186.

Yoshioka, T. and Fujiwara, T. (1984). Application of acoustic emission technique to detection of rolling bearing failure. In D. Dornfiled (ed.), *Acoustic emission monitoring and analysis in manufacturing*, 55–75. ASME, New York.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Robust Fault Detection of Nonlinear Systems using Local Linear Neuro-Fuzzy Techniques with Application to a Gas turbine Engine

Hasan Abbasi Nozari*, Mahdi Aliyari Shooredeli, Silvio Simani*****

**Department of Mechatronics, Faculty of Engineering, Azad University, Science and Research branch, Iran, Tehran (Tel: +989376414624; e-mail: H.Abbasi@ srbiau.ac.ir).*
***Khaje nasir Toosi University of Technology, Faculty of Electrical Engineering, Iran, Tehran (e-mail:aliyari@eetd.kntu.ac.ir)*
**** Department of Engineering, University of Ferrara,Via Sargat, IE-44122 Ferrara(FE), Italy (e-mail: silvio.simani@unife.it)}*

Abstract: This study proposed a model-based robust fault detection (RFD) method using soft computing techniques. Robust detection of the possible incipient faults of an industrial gas turbine engine in steady-state conditions is mainly centered. For residual generation a bank of Multi-Layer perceptron (MLP) models, is used, Moreover, in fault detection phase, a passive approach based on Modellling Error Model (MEM) is employed to achieve robustness and threshold adaptation, and toward this purpose, Local Linear Neuro-Fuzzy (LLNF) model is exploited to construct error model to generate uncertainty interval upon the system output in order to make decision whether or not a fault occurred. This model is trained using the Locally Linear Model Tree (LOLIMOT) algorithm which is an incremental tree-structure algorithm, In order to show the effectiveness of proposed RFD method, it was tested on a single-shaft industrial gas turbine prototype model and has been evaluated using non-linear simulations, based on the gas turbine data.

Keywords: fault detection, neural network, gas turbine engine, local linear neuro-fuzzy local linear model tree (LOLIMOT), system identification.

1. INTRODUCTION

Nowadays reliability is one of the crucial issues in automatic system design and has received great attention during last two decades. Due to manufacturing defects, erosion-corrosion and tear, and other kind of performance deteriorations in system's components, and in order to prevent major collapses in plant, system shutdowns, "early" diagnosis of faults is an important factor. Among different fault diagnosis approaches model based methods are still a wide open area of research. In order to make model-based fault detection(FD) algorithms more applicable to real industrial systems, neural networks, fuzzy sets or their combination (neuro-fuzzy) can be considered (Patan et al., 2008).

A FD method must be effectively developed to cope with un-wanted and uncontrolled effects such as disturbance, noise, uncertainty of the model, etc. which could dramatically decrease the reliability of fault detection. Robustness could be included in fault diagnosis procedure via active and passive approaches (Patan et al., 2008). Active methods usually leads to define suitable performance index and optimize it with the objective of achieving most sensitivity to fault and most robustness to disturbance, noise, etc. A variety of active robust fault diagnosis methods, of course, with application to linear/linearized systems, are proposed in the literature such as unknown input observer (Chen et al., 1996), robust parity equation (Gertler, 1998), H_∞ (Frank and Ding, 1994), H _ (Jaimoukha et al., 2006). The main drawback of above active methods is that they are not applicable in real industrial applications, because some realized hypothesis, which are not possible in practical environment, are taken in to account in enhancing of robustness to fault diagnosis such as: prior knowledge of disturbance and noise acting on the system is always available, and the model of the system is accurate enough to describe the plants dynamics. Another approach for RFD is passive approach that which is usually based on the adaptive threshold computed for the residual by propagation of uncertainty to residual. Passive approaches tackle to RFD problem despite of model uncertainty, and that is the main reason makes passive approach more suitable for experimental applications than active one. There are ideas which were proposed in order to drive adaptive threshold for nonlinear systems using soft computing techniques. Fuzzy logic was used to describe threshold changes (Sauter et al., 1993; Schneider, 1993). GMDH neural networks were also used for threshold adaptation by estimating of model uncertainty in order to perform robust fault diagnosis (Patan, 2008). Model error modeling (MEM) can be used as passive approach in RFD. Robust fault diagnosis using MEM was performed successfully on dynamic systems (Patan et al., 2008).

In addition to importance of robustness extension to fault detection procedure, FD method must also tackle to detection of incipient faults. Since in industrial applications, it is commonplace for most of the faults to develop slowly over a long period of time, these type of faults are hardly detectable immediately by a simple inspection of output signals, hence, a proposed FDI method must be developed effectively so that

timely detection of ramp faults as well as robustness to uncontrolled effects such as disturbance and noise, etc. achieved.

This study is concerned with the use of MLP neural networks applied to FD of nonlinear dynamic systems, based on multiple modeling, besides LLNF model trained with LOLIMOT algorithm is used for model error modeling in order to generate adaptive (thresholds) bands. Due to computationally complex of MEM in a data-driven identification task, ability of LOLIMOT algorithm in fast training of LLNF models with any given network topology leads to less computationally expense than exploiting classical or neural models, which will be very suitable for on-line RFD applications. The proposed robust FD scheme is validated with an industrial gas turbine engine (IGTE) developed at ABB-Alstom power, United Kingdom.

The rest of paper is organized as follows: In section 2 a brief overview of gas turbine under consideration and its possible faulty scenarios description is presented, section 3 introduces the proposed RFD method. Fault detection using simple thresholding and also MEM- based robust fault detection using LLNF models with LOLIMOT learning algorithm are included in sections 4, 4.1, respectively. Simulation results obtained by proposed techniques are included in section 5, moreover, brief comparison with other proposed fault diagnosis methods on the under consideration gas turbine benchmark is presented in the last part of this section, and finally main conclusions obtained are drawn in section 6.

2. SYSTEM AND FAULY SCENARIOS DESCRIPTION

In gas turbine engine of interest in this study air flows via an inlet duct to the compressor and the high pressure air from the compressor is heated in combustion chambers and expands through a single stage compressor turbine. A Butterfly valve provides a means of generating a back pressure on the compressor turbine (there is no power turbine present in the model). Cooling air is bled from the compressor outlet to cool the turbine stator and rotor, A Governor regulates the combustor fuel flow to maintain the compressor speed at a set-point value (Patton et al., 2000). For simulation purposes a full scale Simulink prototype model of such an industrial gas turbine Developed at ABB-Alstom Power, United Kingdom was used. The SIMULINK prototype simulates the real measurements taken from the gas turbine with a sampling rate of 0.08s. The model has two inputs and 28 output measurements which can be used for generating residuals The Simulink model where validated in steady state conditions against the real measurements and all the model variables were found to be within 5% accuracy. Four common faulty scenarios of an industrial gas turbine engine are proposed to be tackled in this paper:

1) Compressor contamination, [f_1], core engine performance deterioration.

2) Thermocouple sensor drift, [f_2], (Slowly increasing reading over the time).

3) High pressure Turbine seal damage, [f_3], (turbine efficiency gradual reduction).

4) Fuel actuator friction wear, [f_4].

Valve angle (a_v), and fuel flow (ff) that is also a control variable are the input measurements whereas The output sensors are those used for the measurement four output so-called compressor torque (Q_{oc}), Compressor outlet temperature (T_{oc}), combustion chamber outlet pressure (P_{occ}) and, combustion chamber outlet pressure back to compressor that for simplicity, we named this measurement as compressor inlet pressure (P_{ic}), are considered output measurements for intent of fault detection.

3. PROPOSED ROBUST FAULT DETECTION METHOD

The proposed RFD scheme which is shown in Fig.1 could be abbreviated in two parts: residual and uncertainty bands generation and decision making on fault occurrence time. Detection of faults is implemented by modeling of the normal behavior of monitored gas turbine closed loop system using MLP neural network, then, the residuals are generated by comparing the predicted output using MLP models and, system outputs. Since the MEM is quite a time consuming task and also high-order model is needed to create a faithful replica of modeling error, neuro-fuzzy (NF) models could have a better performance in modeling than other structures due to their ability in representing nonlinearity by several local linear models.

In order to perform robust fault detection, LLNF models trained with well-known fast training LOLIMOT incremental algorithm are used to identify error model. After selecting of proper inputs and output, an input-output model of modeling error (residual) is identified. Inputs of process and also the residual are used as the input of this model. The output of these modeling error models are added with correspondent nominal model's output in order to generate centre of uncertainty region (UC). Then in uncertainty bands generator block, upper and lower bands are built by some statistical extension to generated uncertainty centers. Decision on occurrence of fault is then made using these bands considering system outputs.

.

4. FAULT DETECTION

The residual signals are generated based on comparison between the measurements coming from plant full scale simulator and predicted signals given by the MLP models. The residual are calculated as follows:

$$R_i(k) = y_i(k) - \hat{y}_i(k) \quad (1)$$

where $y_i(k)$ and $\hat{y}_i(k)$ are process measurements and predictions, respectively

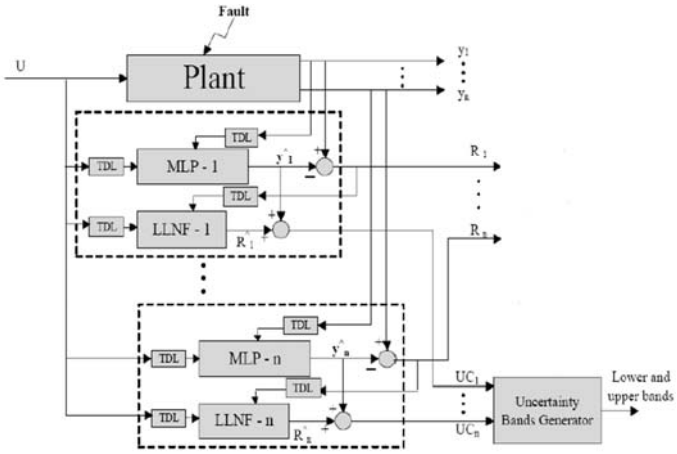


Fig.1. Neural and Neuro-Fuzzy based RFD scheme: residual generation, threshold adaptation.

. It is clear that the residual signals should have near zero behaviour in absence of faults otherwise meaningfully deviation from zero. The main objective of fault detection (FD) is timely detection of faults presented in the system to have enough time to take counteractions such as reconfiguration, maintenance, repair or other operations. The fact may vague fault detection by rising false alarms is that faults can commonly be described as inputs, and moreover, there is always modeling uncertainty due to an incorrect model structure or un-modeled disturbances, noise and dynamics. Hence, in a fault detection problem there is always a compromise between false alarms and revealing of fault detection, and when more delays in detection of faults is tolerable, false alarm rate reduction is preferable.
One common idea employed for fault detection is simple (constant) thresholding. In order to make a decision whether or not a fault occurred, fixed threshold can be applied. In decision making step, testing of threshold (λ) overpassing is applied to residual signal(R) as follows:

$$S(t) = \begin{cases} 0 & if \quad |R(t)| \le \lambda \\ 1 & if \quad |R(t)| \ge \lambda \end{cases} \qquad (2)$$

Where S is fault signature. In order to establish the identifiable fault signature the threshold value λ must be settled. There are a variety of statistical or experimental considerations based approaches to drive simple threshold. A frequently used constant thresholding method is η-standard deviation. Supposing that the residual as an $N(m,s)$ random variable, simple thresholds could be defied as:

$$\lambda = m \pm \eta s \qquad (3)$$

Where η is a tuning parameter that can be 1, 2 or 3.

4.1. *Model Error Modeling and robust fault detection*

Model Error Modeling (MEM) employs prediction error methods to identify a model from input-output data. After that, one can estimate the uncertainty of the model by analyzing residuals evaluated from the inputs. Uncertainty is a measure of unmodelled dynamics, noise and disturbances. The identification of residuals provides the so-called *error model* (Pattan et al., 2008).
Original frequency domain MEM algorithm (Reinelt *et al*., 2002) was mapped into time domain version one, and utilized in robust fault diagnosis of dynamic systems (Pattan et al., 2008).
Time domain MEM procedure for uncertainty interval generation aiming at robust fault diagnosis, can be sketched by the Algorithm 1 below.

Algorithm 1.

Step 1: Identify a nominal model of process, and then compute residual ($R = y - \hat{y}$).
Step 2: Identify an error model by collected $[U, R_i]$ data where U is system input and R_i is i-th residual.
Step 3: build the center of uncertainty region as $UC \approx \hat{y} + \hat{R}$, where $\hat{y}$ is the output of nominal model of the process and $\hat{R}$ is the output of error model.
Step 4: If the model error model is not falsified by the data, one can use statistical properties to calculate a confidence region. Uncertainty upper and lower bands (interval) can be generated using the response of error model and center of uncertainty region as $\lambda_{U/L} = \hat{y} + \hat{R} \pm t_\beta s$, where t_β is the N (0, 1) tabulated value assigned to a given confidence level, and s is the standard deviation of $\hat{R}$.

Soft computing technique can be employed to carry out process and model error modelling. When a nominal model of process is constructed then the most important phase is to create an accurate error model. How to construct and train an accurate error model by means of LLNF model is described in detail in the next section. It is worth noting that $\hat{R}$ represents a structured uncertainty, disturbances, etc.

4.1.1. *MEM-based threshold adaptation using LLNF model*

The more important use of model error models is to let it be different from the nominal model structure, e.g. non-linear and/or time varying and in most cases a linear model could not be a faithful replica of modeling error (Reinelt *et al*., 2002). But the golden rule in identification task is to try simple things first. If a linear model does a decent job, one should not bother wasting time on fancy nonlinear models such as neural networks. But the underlying residual being nonlinear, obviously linear error model doesn't work well. Hence one should start with fitting a linear model if the result weren't satisfactory try a nonlinear structured model. Additional to above remarks, faithful symmetric uncertainty bands generation and well-working of MEM technique in robust fault detection, critically depends on employing an effective identifier. Moreover, enhancing such MEM based

fault detection decision step to diagnostic system, especially in on-linefault diagnosis applications, leads to increasing in computationally expense of FD algorithm. In order to tackle such a problem, LLNF model training with a very fast training algorithm so-called LOLIMOT, in comparison to other time consuming conventional learning algorithm, is utilized to develop an accurate error model.

The network structure of LLNF with external dynamic is shown in Fig.2. Each neuron realizes a Local Linear Model (LLM) and an associated validity function that determines the region of validity of the LLM. The network output is calculated as a weighted sum of the outputs of the local linear models, where the validity function is interpreted as the operating point dependent weighting factors. The validity functions are typically chosen as normalized Gaussians.
The local linear modeling approach is based on a divided-and-conquer strategy. A complex error (residual) model divided into a number of smaller and thus simpler sub-problems, which are solved independently by identifying simple linear models (Nelles, 2001; Nelles and Iserman, 1996). The most important factor for the success of such an approach is the division strategy for the original complex problem this will be done by an algorithm named LOLIMOT (Locally Linear Model Tree). LOLIMOT is an incremental tree construction algorithm that partitions the input space by axis-orthogonal splits (Nelles, 2001).In each iteration, a new rule or local linear model is added to the model and the validity functions that correspond to the actual partitioning of the input space are computed, and the corresponding rule consequence are optimized by a local weighted least squares technique.

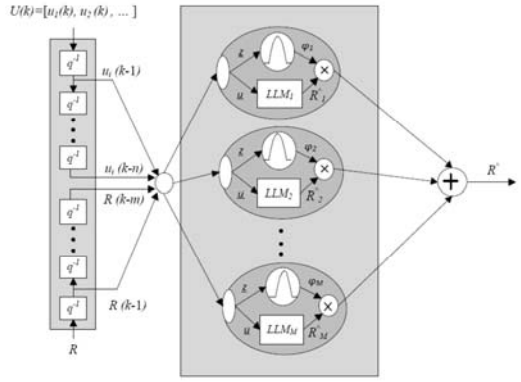


Fig.2. Network structure of a local linear neuro-fuzzy model with external dynamics.

Five step LOLIMOT algorithm is summarizes as (Nelles, 2001):

Algorithm 2.

*Step*1: Start with an initial model
*Step*2: Find worst locally linear model which has maximum local loss function.
*Step*3: Check all hyper-rectangles to split (through).
3a. Construction of the multi-dimensional Fuzzy membership Functions for both hyper rectangles.
3b. Construction of all validity functions.
3c. Local estimation of the rule consequent parameters for both newly generated LLMs.
3d.Calculation of the loss functions for the current overall model.
*Step*4: Find best division (the best of the alternatives checked in Step 3, and increment the number of LLMs: M → M+1).
*Step*5: Test for convergence.

The training procedure of error model using LLNF network is illustrated in Fig.3. Furthermore, in our case, MLP neural network is employed to extract the nominal model of process while LLNF network is used to model an "error" system with the input of the real system and the output R.

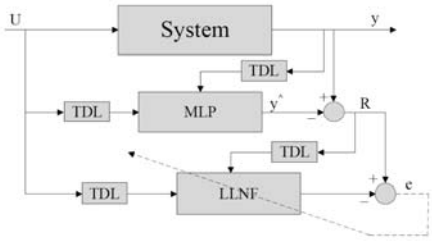


Fig.3. Error model training scheme using LLNF model

Choosing a proper structure for error model is very important in MEM technique, and starting with an *a priori* chosen flexible structure, e.g. the 10-th order FIR filter is recommended in (Reinelt et al., 2002).If this error model is not falsified by the data keep it. Otherwise increase the model error model complexity until it is un-falsified by the data (Patan et al., 2008).

5. SIMULATION RESULTS

In the case of NNs based system identification the important factor is number of neurons as well as epochs. Large number of neurons caused complexity in computations and also over parameterization problem. Thus, small and reasonable neuron number is preferable. Optimal neuron and epoch's numbers are determined by means of mean square error curves in logarithmic scaled plots due to sharp diminishing of mean square error values. As four outputs are taken into account for fault detection of underlying gas turbine, the relevant number of MLP models is also four MLP based models.

In order to obtain an accurate identification results for all of four cases a Levenberg–Marquardt training algorithm was employed and training was terminated, when a minimum in the mean-square error of the test data was achieved. Moreover, two time series of data from different turbine operation pints were used in order to evaluate the generalization ability of the networks. MISO model's output for the case of compressor torque is illustrated along with correspondent system outputs in Figs. 4. The number of neurons and their corresponding RMSE for each MLP model are listed in Table 1. Through trial and error during identification, the numbers of input and output dynamics are obtained.

Table 1. Root Mean Square Error and optimal neuron number

Model	Hidden Neuron number	RMSE		
		Train	Valid.1	Valid.2
MLP__Q_c	6	$2\ e^{-3}$	$2.05\ e^{-2}$	$2.55\ e^{-2}$
MLP__T_{oc}	7	$1.3\ e^{-3}$	$1.52\ e^{-2}$	$4.47\ e^{-2}$
MLP__P_{occ}	7	$8.94\ e^{-2}$	$1.84e^{-2}$	$1.1\ e^{-2}$
MLP__P_{ic}	6	$2.21\ e^{-3}$	$9.1\ e^{-3}$	$3.21\ e^{-2}$

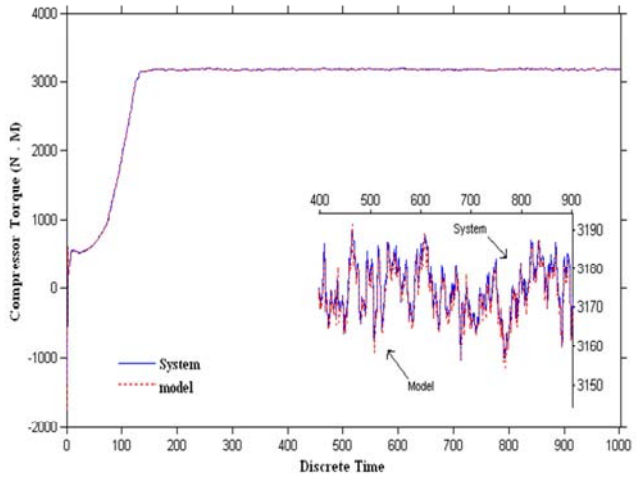


Fig.4. MLP model performance of compressor torque.

After training four accurate MLP models in fault free condition, these models are used to generate residuals by running the simulator with fault free and then all faulty cases operating one by one over the operating ranges. Using the scheme presented in Fig.1, in order to perform fault detection, both simple and adaptive thresholding methods presented in sections 4, 4.1 are employed. Constant thresholds were settled according to (3), and threshold adaptation was carried out using algorithm 1. Fig.5 shows the output of the error model along with the residual for the case of R_2 derived from T_{oc} output, exploiting both auto-regressive exogenous (ARX) and LLNF models. As can be seen high order ARX model(30-*th* order) is less effective than LLNF model and has severe problems, thus it could be concluded that this is due to that the underlying [U, R] system being nonlinear. Generated uncertainty bands for different faulty conditions as well as decision making on revealing of each faulty scenarios in terms of these uncertainty intervals pre-set on system output for the cases of f_1 and f_3 are also illustrated in Figs. 6-7. The results of the fault detection using both constant thresholding and proposed neuro-fuzzy based model error modelling technique are presented in Table 2.

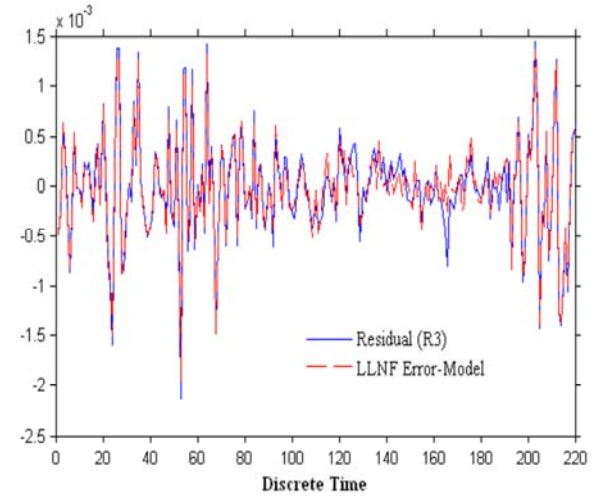


Fig.5. Residual and the error model output for the case of combustion chamber outlet pressure in nominal condition.

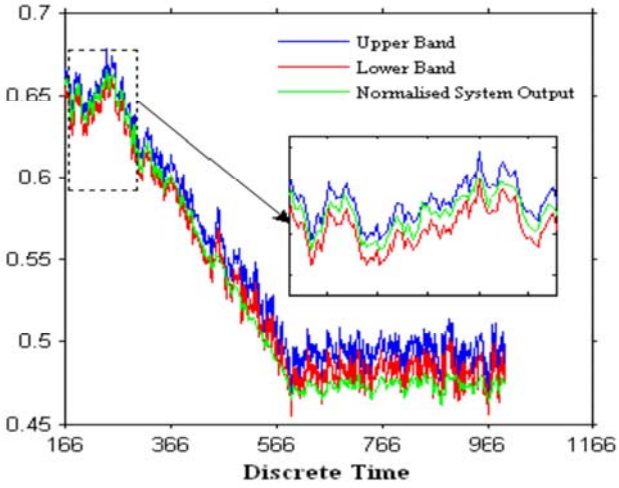


Fig.6. Fault detection using model error modelling for f_1 using compressor torque.

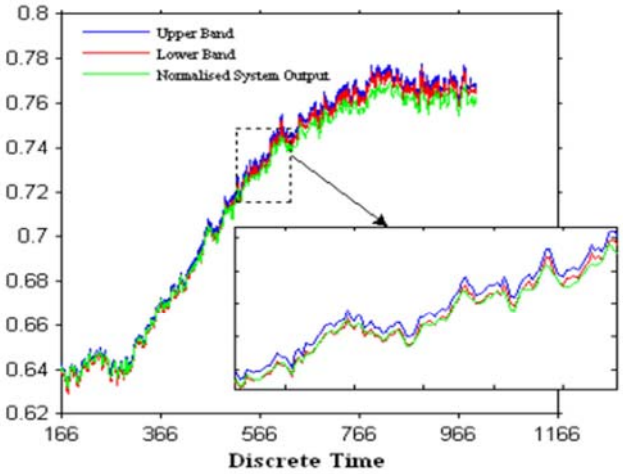


Fig.7. Fault detection using model error modelling for f_3 using combustion chamber outlet pressure.

Table.2. Results of fault detection using constant thresholds and Model Error Modelling.

	Fault	Fault Inception Time [Sample]	Detection Delay [Sample]	False Alarm Rate [%]	True Fault detection Rate [%]
Constant threshold	f_1	250	13	32.14	78.59
	f_2	250	12	46.43	95.35
	f_3	250	2	86.90	92.42
	f_4	250	0	92.86	77.26
MEM - LLNF	$f1$	250	13	2.38	66.49
	$f2$	250	12	4.76	90.69
	$f3$	250	14	1.19	69.15
	$f4$	250	5	8.33	48.94

By comparing achieved results from above fault detection methods in Table 2, and also comparing of the achieved results with the FD results presented in other related works on the literature on fault detection of this industrial gas turbine prototype confirm that the proposed neuro-fuzzy based RFD technique demonstrates more reliable behaviour and robustness than other FD methods. In addition to Timely fault detection in all of four ramp faulty scenarios, dramatic decrease in false alarm rates, which is an important factor in fault detection task, was also effectively achieved using MEM method.

6. CONCLUSIONS

A hybrid multiple model-based robust fault detection method was presented and early fault detection of an industrial gas turbine engine working on different operating points was achieved. MEM based method using LLNF techniques were also proposed for robust fault detection in comparison to simple thresholding methods.

The main contributions of this paper are:

1) Nonlinear dynamic identification of gas turbine engine was performed.
2) Moreover, robust identification of gas turbine engine using both neural and neuro-fuzzy techniques was carried out.
3) A straightforward robust fault detection method based on local linear neuro-fuzzy techniques was proposed and was applied effectively on an industrial gas turbine prototype.

Due to nonlinear aspect of both neural networks and LLNF models it could be concluded that proposed RFD approach can be easily exploited as a general approach to cope with fault diagnosis of other nonlinear dynamic systems.

Developing our RFD method in order to deal with detection of multiple faults could be a worthwhile future contribution and also proposed RFD method could be easily employed in on-line fault diagnosis applications.

REFERENCES

Basseville M, Nikiforov IV, (1993). Detection of abrupt changes: theory and application. Prentice-Hall Inc.

Chen J, Patton R J, (1999). Robust Model Based Fault Diagnosis for Dynamic Systems. Kluwer Academic Publishers ISBN 0-7923-8411-3.

Chen J, Patton RJ, Zhang HY, (1996). Design of unknown input observer and robust fault detection filters. Int J Control; 63(1):85–105.

Gertler J, (1998). *Fault detection and diagnosis in engineering systems*. New York: Marcel Dekker.

I. M. Jaimoukha, Z. Li, and V. Papakos (2006), A matrix factorization solution to the H−/H∞ fault detection problem, Automatic, vol. 42, pp. 1907–1912.

Kyriazis A., Aretakis N., Mathioudakis K, (2006). Gas Turbine Fault Diagnosis From Fast Response Data Using Probabilistic Methods And Information Fusion, Proceeding Of GT, ASME Turbo Expo2006, Power For Land ,Sea And Air., Barcellona, Spain.

M.Blanke, M.Kinnaert, J.Lunze, M.Staroswiecki, (2003). *Diagnosis and Fault Tolerant Control*. Springer-Verlag, Heidelberg.

Nelles O. (2001), *Nonlinear System Identification:From Classical Approaches to Neural Networks and Fuzzy Models*, Springer Press.

Onder U., Kyusung K., Emmanuel O.N. (2006). Synergistic Use of Soft Computing Technologies for Fault Detection in Gas Turbine Engines. *IEEE Transactions on Systems*, Vol. 36, No. 4.Man, And Cybernetics—Part C: Applications And Reviews.

Patan K. (2008). *Artificial Neural Networks for the Modeling and Fault Diagnosis of Technical Processes*. springer - Verlag Berlin Heidelberg.

Patan K., Witczak M., Korbicz J. (2008), Toward Robustness in Neural Network based Fault Diagnosis, *Int. J. Appl. Math. Comput. Sci.*, Vol. 18, No. 4, 443–454.

Patton, R. J., Simani S., Daley S., Pike A. (2000).Fault diagnosis of a simulated model of an industrial gas turbine prototype using identification techniques. In Proceedings of the 4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, pp 518-523, Budapest.

P.M. Frank and X. C. Ding (1994), Frequency-domain approach to optimally robust residual generation and evaluation for model-based fault-diagnosis, *Automatica*, vol. 30, no. 5, pp. 789–804.

Reinelt W., Garulli A. and Ljung L. (2002). Comparing different approaches to model error modeling in robust identification, *Automatica* 38(5): 787–803.

Sauter, D., Dubois, G., Levrat, E., Br´emont, J.(1993). Fault diagnosis in systems using fuzzy logic. In: Proc. First European Congress on Fuzzy and Intelligent Technologies, EUFIT'93, Aachen, Germany, 781–788.

Schneider, H. (1993), Implementation of a fuzzy concept for supervision and fault detection of robots. In: Proc. First European Congress on Fuzzy and Intelligent Technologies, EUFIT'93, Aachen, Germany, 775–780.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fault Detection and Isolation of Tennessee Eastman Process Using Improved RBF Network by Genetic Algorithm

Somayeh Hekmati Vahed*. Mohammad.Mokhtare .*Hassan Abbasi Nozari.*
Mahdi Aliyari Shoorehdeli. Silvio Simani.*****

** Faculty of Eng., Mechatronics Dept. Science and Research Branch, Islamic Azad University, Tehran, Iran, (e-mails: s.hekmati@ srbiau.ac.ir, m.mokhtare@ srbiau.ac.ir, h.abbasi@srbiau.ac.ir)*
*** Faculty of Electrical Engineering, Mechatronics Dept., K. N. Toosi University of Tech., Tehran, Iran, (e-mail: aliyari@eetd.tntu.ac.ir)*
**** Department of Engineering University of Ferrara,Via Sargat,IE-44122 Ferrara (FE),Italy (e-mail: silvio.simani@unife.it)*

Abstract: This paper presented Radial Basis Function (RBF) network as a classifier for Fault Detection and Isolation (FDI) and to enhance the model accuracy, a real-coded Genetic Algorithm (GA) is implemented to search for optimal model parameters. Comparison between the performance of FDI by using RBF Neural Network and RBF Neural Network which improved by Genetic algorithm is presented. These techniques are applied to simulated data collected from the Tennessee Eastman Process (TEP) chemical plant simulator that is designed to simulate a wide variety of faults occurring in a chemical plant based on a facility at Eastman chemical. In this study detection and isolation of all faults recorded in the process is investigated.

Keywords: Fault Detection, Fault isolation, Radial Basis Function (RBF) Network, Genetic Algorithm, Principle Component Analysis (PCA),Tennessee Eastman process (TEP).

1. INTRODUCTION

Modern chemical process is becoming more and more complex, usually including a large number of components (such as sensors, actuators, and computers, etc). They require more reliable operations since system malfunctions may cause serious safety problems. In order to maintain a high level of safety, quality and reliability in control systems, it is very important that abnormal system operations and component faults are detected promptly. Therefore, there has been a surge of interest in research and application on fault detection and isolation (FDI) techniques in the last three decades. Since early development in 1980's this area has matured with the conception of various FDI methods. The main purpose of all FDI methods is to monitor system operations in the case of faults, accommodate the source of the faults so that timely corrective actions are taken. System reconfiguration can be accomplished afterwards by human operators or automatic configuration to maintain nearly normal operation. The process simulator for the Tennessee Eastman industrial challenge problem was created by the Eastman chemical company to provide a realistic industrial process in order to evaluate process control and monitoring methods.

This paper presents fault detection and isolation using RBF network and increases its performance with applying genetic algorithm optimization on TEP.

We should note that in this paper detection and isolation of all faults recorded in the process has been discussed. While in most papers in this field, only some specific faults have been investigated.

1.1 Fault detection and isolation techniques on TEP

Fault is a non permitted deviation of a characteristic property which leads to the inability to fulfil the intended purpose and failure is complete breakdown of a system component or function. Fault detection is determination of the presence of a fault in a system and the time of its occurrence. Fault isolation is estimation of the type, magnitude and cause of the fault.

Signal based fault detection is the most frequently used diagnosis method in practice. The idea is to monitor the level of a particular signal and raise alarm when the signal reaches a certain threshold. Sumana C. suggests Dynamic Kernel Scatter-difference-based Discriminator Analysis (DKSDK) a novel method fault diagnosis of TEP (2009). Leo H. Chiang used Fisher discriminator analysis and Support Vector Machines (SVM) for fault diagnosis in TEP (2004). LI Gang applied PLS based contribution plots for fault diagnosis for this process (2009). Zhang Ying-Wei offers decentralized fault diagnosis of large-scale processes using Multi Block Kernel Principle Component Analysis (MBKPCA) (2010). S.Bahrampour applied modified Gath-Geva clustering technique for fault detection in TEP (2009). Manabu Kano used dissimilarity of process data for process monitoring (2000). Ashish Singhal evaluate a pattern matching method the TEP (2006). Sylvian Verron applied Bayesian Network to the TEP (2006). Noruzi M. applied fault detection of the

Tennessee Eastman process using improved PCA and neural classifier (2009).

Model based fault detection can be defined as detection and isolation of faults by comparing the system's available measurements with a priori information represented by a mathematical model of system. The difference between real measurements and estimates of these measurements are used to generate a residual quantity. Fault is then detected by setting a threshold on this residual quantity. Nan.Ye applied optimal averaging level control for the Tennessee Eastman problem (1995).

Expert system is based on heuristic knowledge as, if set of symptoms were occurred then fault occurred. Methods based on soft computing consist of artificial neural network (ANN), fuzzy logic, genetic algorithm and combination of those methods. Yong MAO applied fault diagnosis based on fuzzy support vector machine with parameter tuning and feature selection on TEP (2007). So many other researches and methods are proposed for fault detection and isolation in TEP as a challenge. This paper presents a combination of RBF network and genetic algorithm optimization for fault detection and isolation and introduces a novel method to get better results in this challenge.

2. RBF NETWORK

2.1 Network structure for FDI

If several RBF neurons are used in parallel and are connected to an output neuron the radial basis function network is obtained. Structure of RBF network is shown in (Fig. 1).

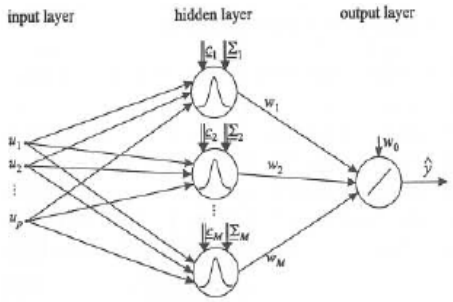


Fig. 1. Structure of RBF network

In basis function formulation can be written as

$$\hat{y} = \sum_{i=1}^{M} w_i \Phi_i(\left\| \underline{u}_i - \underline{c}_i \right\| \cdot \sum_i) \; with \; \Phi_0(.) = 1 \qquad (1)$$

With the output layer weights W_i. The hidden layer parameters contain the centre vector $\underline{C}_i$, which represents the position of the ith basis function, and the norm matrix $\sum_i$, which represents the widths and rotations of the ith basis function.

2.2 RBF Network Training

For training of RBF networks different strategies exist. Typically, they try to exploit the linearity of the output layer weights and the geometric interpretability of the hidden layer parameters. Thus, most strategies determine the hidden layer parameters first, and subsequently the output layer weights are estimated by least squares. There are two methods for training of hidden layer parameter such as centre of gravity ($\underline{C}_i$) and standard deviation ($\sum_i$); that consist of supervised and unsupervised methods. Due to used data for FDI system are selected from real data of TEP, thus supervised method is selected for training of gravity centres and standard deviation. Because of TEP is a nonlinear system, thus in designing RBF network, Gaussian function and standard deviation are used separately for each dimension. So the basis functions considering as

$$\Phi_i(.) = \exp(-\frac{1}{2}\sum_{J=1}^{n} \frac{(u_j - c_{ij})^2}{\sigma_{ij}^2}) \qquad (2)$$

The derivative of the RBF network output with respect to the jth coordinate of the centre ith neuron is (i=1,…M ,j=1,… p)

$$\frac{\partial \hat{y}}{\partial c_{ij}} = w_i \frac{u_i - c_{ij}}{\sigma_{ij}^2} \Phi_i(.) \qquad (3)$$

The derivative of the RBF network output with respect to the standard deviation in the jth dimension of the ith neuron is (i=1 , …, M , j= 1 , … , p)

$$\frac{\partial \hat{y}}{\partial \sigma_{ij}} = w_i \frac{(u_i - c_{ij})^2}{\sigma_{ij}^3} \Phi_i(.) \qquad (4)$$

2.3 Using RBF Network for FDI

First, inputs that consist of faulty and normal inputs enter to the $\Phi_i(.)$ that is a Gaussian function with adjustable mean and variance, then $w_i \times \Phi_i$ (i=1,…, M) are formed in the output. So network can detect and identify the fault.

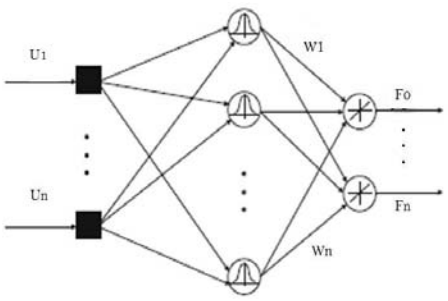


Fig. 2. RBF Network structure for FDI

As it is shown in (Fig. 2), it is supposed that F0 is representative of normal state and F1,…, Fn, are representative of faulty states. When the system works in

normal state, In this case, the network is trained only for F0. For this purpose we choose gradual training. E.g. in normal state, the network is trained as positive training for F0 and negative training for F1,…, Fn. In this case (W1,…, Wn) are trained, as outputs converge only to F0. This method repeated to all faulty cases.

3. GENETIC ALGORITHM

Genetic algorithm is a global optimization that its object is finding minimum of function.

3.1 Implementation of genetic algorithm for FDI

First the same as number of inputs, population of X is generated that X is a vector and is called chromosome then we put them in the fitness function. In this case the numbers of fitness function will be the same as the number of inputs. Then by using single parent combination (Mutation) and double parent combination (cross over) we produce so many new X population. Then we put this new population into fitness function and proportional to the fitness function we give them survival probability. The more output of fitness function is, the higher will be probability of selecting this population. In any case there is a probability of selection all individuals in the population. This approach of selection is called Roulette wheel. Criterion of impress in classification is a standard for fitness selecting so it is recommended that definition of fitness function is been as

$$Fitness \ = \#missclass \ (X) + \lambda \#(1) \qquad (5)$$

That # is symbol of number and λ is balance factor. Number of miss class shows that fault is occurred but network do not identify it or the network detects a fault but fault does not occurred. The number of ones in (5) is number of ones in selected chromosome. Advantage of using genetic algorithm in selecting of RBF network's inputs for FDI is that with using genetic algorithm we can classify sensors with depending of the importance in fault. So it will be better than PCA for selecting of network's inputs. Because the PCA is a linear transform and it can't illuminate importance of sensors in fault. In this paper it is supposed that length of chromosome is constant and probability of mutation is 0.1 and probability of cross over is 0.8.This supposition leads us to better result.

4. EXPERIMENTAL SIMULATIONS OF THE PROPOSED FDI APPROACHES ON THE TEP PLANT

The process simulator for the Tennessee Eastman Industrial Challenge Problem was created by the Eastman Chemical Company to provide a realistic industrial process in order to evaluate process control and monitoring methods. Tennessee Eastman process is a chemical process. It is composed of five major operation units: a reactor, a condenser, a compressor, a stripper and a separator. This process has 12 input variables and 40 output variables. It has 21 types of identified faults.

Table 1. Process Fault for Tennessee Eastman Process

Fault ID	Description	Type
F1 (IDV1)	A/C feed ratio, B composition constant	step
F2 (IDV2)	B composition, A/C ratio constant	step
F3 (IDV3)	D feed temp	step
F4 (IDV4)	reactor cooling water inlet temp	step
F5 (IDV5)	condenser cooling water inlet temp	step
F6 (IDV6)	A feed loss	step
F7 (IDV7)	C header pressure loss-reduced	step
F8 (IDV8)	A, B, C feed composition	variation
F9 (IDV9)	D feed temp	variation
F10 (IDV10)	C feed temp	variation
F11 (IDV11)	reactor cooling water inlet temp	variation
F12 (IDV12)	condenser cooling water inlet temp	variation
F13 (IDV13)	reaction kinetics	slow drift
F14 (IDV14)	reactor cooling water valve	sticking
F15 (IDV15)	condenser cooling water valve	sticking
F16-20(IDV16-20)	Unknown	unknown
F21 (IDV21)	valve for stream 4 fixed at the Steady -state Position	constant position

Decentralized control system of the Tennessee Eastman process is shown in (Fig. 3).

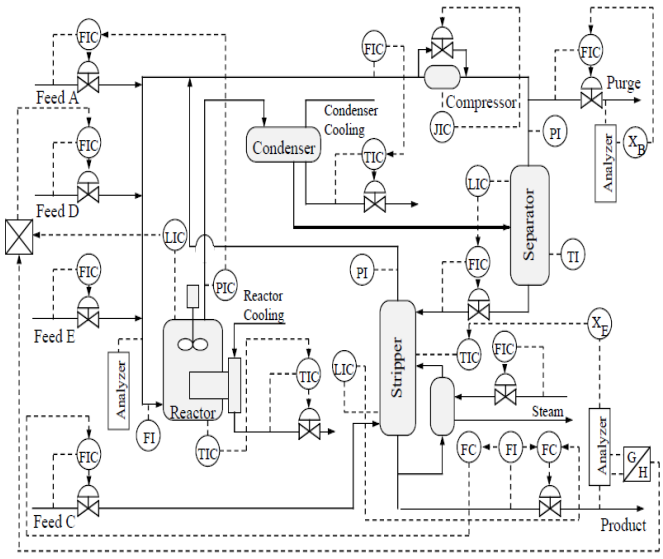


Fig. 3. Decentralized control system of the Tennessee Eastman process.

Authors implement FDI using RBF network and improved RBF network by genetic algorithm. Before simulation to get better result it is recommended that all data should be normalised so it is called data pre processing.

4.1 Results of simulation for fault detection using RBF network

After pre-processing of data, they are divided to two sections. For fault detection, 1004 numbers of data are selected for train and 1968 numbers of data are selected for test. The RBF network is trained with train data and it is evaluated with test data. Performance of mean square error (MSE) with increasing number of neuron for test and train data is shown in (Fig. 4).

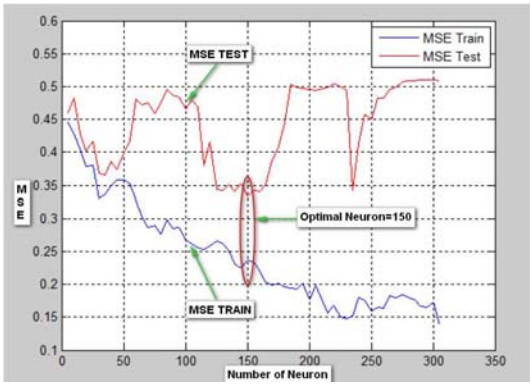


Fig. 4. Performance of MSE for test and train data with increasing number of neuron for fault detection using RBF network.

Beside of MSE there are two criteria for validation of fault detection that are specifity and sensitivity and defined as

$$Specifity = \frac{TN}{TN + FP} \quad (6)$$

Where (TN) means normal case correctly identify as normal and (FP) means normal case incorrectly identify as fault.

$$Sensitivit\ y = \frac{TP}{TP + FN} \quad (7)$$

Where (TP) means faulty case correctly diagnosis as fault and (FN) means faulty case incorrectly identify as normal.

As it is shown in (Fig. 4.), the number of optimum neuron for RBF network is 150. Specifity and especially Sensitivity is very important in FDI so with consideration of specifity and sensitivity as a criterion for selecting optimal neuron, it is illustrated that number of optimum neuron is 170. These two criteria for selecting optimal neuron are compared in table (2).

Table 2. Comparison of two methods for select of optimum neuron in fault detection

Number of neuron	TN	TP	FP	FN	Specifity %	Sensitivity %
150	914	398	46	610	95.2083	39.4841
170	617	592	343	416	64.2708	58.7302

Before entering the inputs to network we do principle component analysis (PCA) and reduce dimension of inputs then enter to the network. In this case we get better results.

Fig. 5 shows performance of the MSE for train and test data by increasing number of neurons.

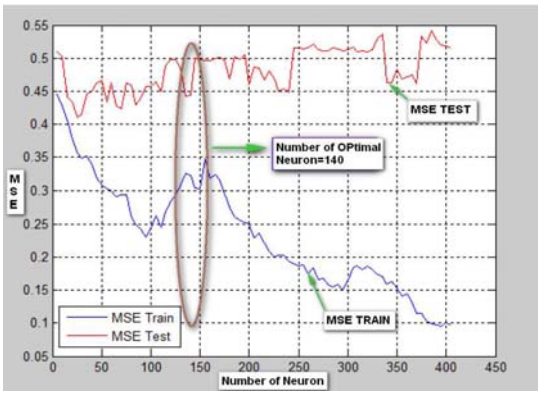


Fig. 5. Performance of MSE for test and train data with increasing number of neurons for fault detection with PCA for RBF network.

Fig. 5 shows the optimal neuron is equal to 140.
Specifity and sensitivity are calculated with 140 neurons and compared with170 neuron that is number of optimal neuron without PCA. The results show in table 3.

Table 3. Comparison between specifity and sensitivity with PCA and without PCA for fault detection using RBF network

	Number of Neuron	Specifity%	Sensitivity%
Without PCA	170	95	39.4
With PCA	140	61	50.5

Selecting of optimal neuron is a trade off between MSE, specifity and sensitivity. Because of sensitivity is more important in fault detection so according table (3) we select the input of network after PCA.

4.2 Result of simulation for improved fault detection using RBF network by genetic algorithm

Considering the number of input equal 1968, balance factor equal 4 in (5), so network has been trained and result will be as:
Number of miss class=722 and number of ones =31 so fitness will be 846. In this case, performance of network for fault detection is as:
performance = ((1968-722)/1968) =63.31%.
If GA is executed, because of number of sensors are 52, so, number of gene in optimal chromosome will be 52.
With using GA, the genes of chromosome that is one, they will be identify effective sensors.
If criteria will be MSE for selecting optimal neuron, the result will be shown in Fig. 6.

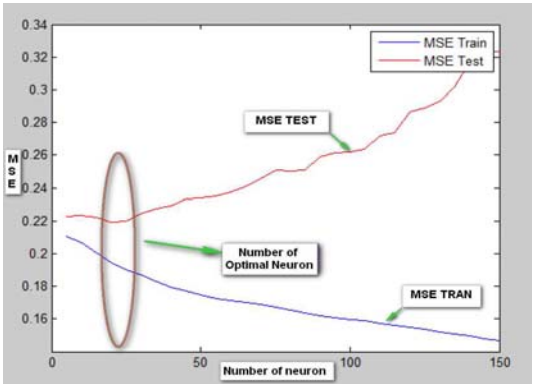


Fig. 6. Performance of MSE for test and train data with increasing number of neurons for fault detection with using GA for selecting RBF network's inputs

As it is shown in Fig. 6 number of optimal neuron is equal to 20.
If specifity and sensitivity will be criteria for selecting number of optimal neuron so it will be 115 neurons. Comparison of these two cases is shown in table 4.

Table 4. Comparison of different criteria for selecting optimal neuron using GA optimization for fault detection

	Number of Optimal Neuron	Specifity%	Sensitivity%
Fault detection with MSE criteria and using GA	20	100	6.7
Fault detection with Sensitivity criteria and using GA	115	92	36

In all cases, number of optimal neurons is less than without using GA

4.3 result of fault isolation using RBF network

For isolation of 21 recorded faults in TEP using RBF network, number of neurons are increased and MSE of train and test data is calculated as criteria for selecting number of optimal neurons in network.
Results are shown in Fig. 7.

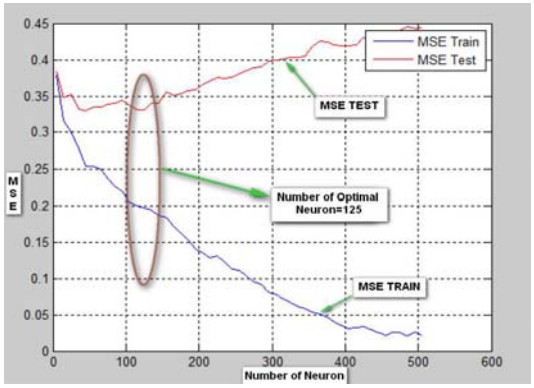


Fig. 7. Performance of MSE with increasing number of neuron for fault isolation with RBF network

As it is shown in Fig. 7 optimal neuron is 125.

4.4. Results of simulation for fault isolation using improved RBF network by genetic algorithm

With considering 528 data for training and 1056 data for test, GA is executed. For calculating fitness function with considering balance factor =1, fitness will be equal 775.
If MSE will be criteria for selecting number of optimal neuron,so the results are shown in fig. 8.

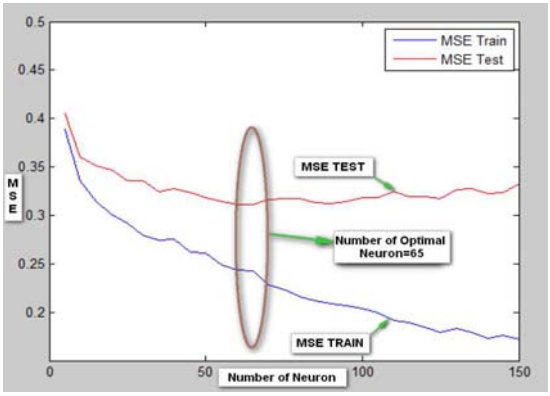


Fig. 8. performance of MSE for test and train data with increasing number of neurons for fault isolation by using GA for selecting RBF network's input

As it is shown in fig. 8 number of optimal neurons is equal to 65.
For more survey, performance of fault isolation using improved RBF network by GA is calculated separately and it is compared with other fault diagnosis methods such as RBF network, improved RBF network using PCA and it is shown in table 5.

Table 5. Performance of fault isolation using improved RBF network by GA comparison with other methods

Fault	Performance of fault diagnosis using RBF network %	Performance of fault diagnosis using improved RBF network by PCA %	Performance of fault diagnosis using improved RBF network by GA %
F1	50	54.16	82.91
F2	70.83	64.58	72.5
F3	8.33	12.5	20.41
F4	18.75	20.83	26.66
F5	60.41	56.25	80.83
F6	70.83	81.25	89.16
F7	66.66	64.5	87.08
F8	16.66	18.75	18.33
F9	4.16	6.25	10
F10	6.25	6.25	12.08
F11	20.83	22.91	24.58
F12	12.5	12.5	22.5
F13	10.4	10.41	24.58
F14	2.08	2.08	22.5
F15	14.5	14.58	26.66
F16	12.5	8.33	41.25
F17	47.91	45.83	53.75
F18	56.25	56.25	55.83
F19	10.41	10.41	16.25
F20	43.75	47.91	43.33
F21	0	0	2.08

In average, the performance of fault diagnosis using improved RBF network by GA is 39.6%. Considering that all 21 faults have been investigated so this method in compare with other methods for fault isolation has better performance.

5. CONCLUSION

The main interest of this paper is the application of a new procedure for FDI on TEP. RBF network is a classifier that used for fault detection and isolation. If genetic algorithm is used for selecting input variables in RBF networks, Results of this method comparison other methods will be improved. So we have a high performance comparison with other FDI methods.

6. REFERENCES

Chiang, L.H., Kotanchek, M.E., Kordon, A.K. (2004). Fault diagnosis based on Fisher discriminant analysis and support vector machines. *Computers and Chemical Engineering,* 28, 1389–1401

Coley, D.A.,An (1999). *Introduction to Genetic Algorithm for Scientists and Engineers.* World Scientific Publishing, London.

Gang, L., Si-Zhao, Q., and other (2009). Total PLS Based Contribution Plots for Fault Diagnosis., *ACTA AUTOMATICA SINICA*, 35, No. 6.

Kano, M., Nagao, K., and other (2000). Issimilarity of process data for statistical process monitoring, *Proceeding of IFAC Symposium on Advanced Control of Chemical Processes,* 1 , 231-236

Mao, Y., Xia, Z., Yin, Z., Sun, Y., Zheng, W. Z. (2007). Fault Diagnosis Based on Fuzzy Support Vector Machine with Parameter Tuning and Feature Selection. *Chinese Journal of Chemical Engineering,* Vol15, Issue2, pp 233-239

Nan Ye, Thomas J. Mcavoy (1995). Optimal averaging level control for the Tennessee Eastman problem*, Canadian Journal of Chemical Engineering,* vol. 73, Issue2, pages 234-240

Nelles, O. (2000). *Nonlinear System Identification*, Springer

Noruzi,M., Aliyari,M. (2009), fault detection of the Tennessee Eastman process using improved PCA and neural classifier, International journal of Electrical & computer science IJECS vol:9 N:9

Singhal A. , Seborg D. E. (2006). Evaluation of a pattern matching method for the Tennessee Eastman challenge process. *elsevier. Journal of Process Control,* 16, 601–613

Sumana,C. , Venkateswarlu,Ch. and other (2009). Dynamic Kernel Scatter-difference-based Discriminant Analysis for Diagnosis of Tennessee Eastman Process, *American control conference*.ThC06.2

Verron, S., Tiplica, T., Kobi, A., (2006). Fault Diagnosis with Bayesian Networks, Application to the Tennessee Eastman Process. published in "IEEE International Conference on Industrial Technology

Zhang, Y.W., Zhou, H., Qin, S. J. (2010). Decentralized Fault Diagnosis of Large-scale Processes Using Multiblock Kernel Principal Component Analysis. *Acta Automatica Sinica.* Vol. 36, No. 4

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Heating ventilation and air conditioning system energy consumption dependency on control set-point selection

Ivan Zajic * **Tomasz Larkowski** * **Dean Hill** **
Keith J. Burnham *

* *Control Theory and Applications Centre, Coventry University, Coventry, UK (e-mail: zajici@uni.coventry.ac.uk or ctac@coventry.ac.uk)*
** *Abbott Diabetes Care Ltd, Witney, Oxfordshire, UK*

Abstract: Nowadays, the general aim is to increase the energy efficiency of heating ventilation and air conditioning (HVAC) systems by achieving highly stable control performance allowing for operation close to the specification limits, where the highest profitability can be obtained. Considering the nonlinear behaviour of HVAC systems the use of advanced control techniques is a necessity in achieving this goal. One of the key questions arises here: are the economical benefits of implementing advanced control techniques higher than the installation costs? In this paper, the steady state characteristics between control set-points and HVAC system energy usage are estimated, in order to help to answer this question.

Keywords: HVAC system, energy consumption analysis, system identification

1. INTRODUCTION

In this paper heating ventilation and air conditioning (HVAC) systems dedicated for clean room production areas are under consideration. These systems provide manufacturing areas, i.e. controlled zones, with conditioned air such that the temperature and the humidity are regulated within predefined limits. Considering the complexity, nonlinear character issues and non-stationary operational conditions of HVAC systems the control of such systems is a non-trivial task. In practice, this commonly results in poor control performance and increased energy consumption, see Underwood (1999). Whilst the suboptimal control performance on such plants is tolerable, the increased energy consumption is becoming increasingly problematic nowadays with concern for the effect on the natural environment. It is estimated in Levenmore (2000) that 15% of a typical HVAC overall energy usage is avoidable via an improved control. Consequently, this research focuses on increasing the energy efficiency of HVAC systems through the analysis and optimisation of control.

Abbott Diabetes Care (ADC) UK, an industrial collaborator of the Control Theory and Applications Centre, develops and manufactures the blood glucose and ketones test strips, which are designed to assist people with diabetes, see Hill et al. (2009). One of the manufacturing requirements is that the environmental conditions during production are highly stable and within defined limits. To achieve this goal ADC UK utilises HVAC systems for clean room production. It has been found that the HVAC systems used on site are rather poorly tuned, hence an analysis of the control system with a view to subsequent optimisation is considered as a high priority.

Recent work with ADC UK has, firstly, focused on the humidity control analysis based upon a developed zone humidity model, see Larkowski et al. (2009). It has been found that the thermodynamical process of dehumidification exhibits bilinear characteristics. A linear in the parameters black-box modelling approach has been chosen to replicate such a behaviour. The derived model has subsequently been used for parameter optimisation of a currently utilised proportional and integral (PI) controller, see Hill et al. (2009) and references therein. Results of this work has yielded a stable and tight control performance, which has subsequently allowed adjustment of the control set-point closer to the specification limit; as a consequence a reduction in gas consumption of approximately 20% has been achieved.

In a similar manner to the work involved in developing zone humidity model, a zone temperature model has been identified, with the aim being to utilise this for subsequent control optimisation, see Zajic et al. (2010). The approach undertaken was to use data from existing sensors installed as a part of the building management system, hence avoiding additional costs and issues connected with installation of new instrumentation. Subsequently, the parameters of the PI controller have been appropriately tuned, which yielded a stable control performance leading to a decrease in the wear and tear of control valves.

Further improvement would require an order of magnitude increase in complexity of the control scheme, leading to the use of more suitable advanced control techniques, such as optimal gain scheduling or multi-variable model-based controllers. Similar trends can be also seen in e.g. chemical process control, see Kano and Ogawa (2009), where the maximal utilisation of conventional advanced control tech-

niques is of a high priority. However, there is the question of weather the use of the more advanced control techniques will yield higher economical benefits, than the use of well tuned, currently utilised, PI controllers. It is considered, that the use of advanced control will enable closer operation to the specification limits, hence potentially increasing the overall HVAC energy efficiency and product quality. Nevertheless the implementation and development costs of such a control scheme might be even higher than the actual energy savings over a realistic payback period.

In order to help to answer the above question, the steady state characteristics between control set-points and HVAC system energy usage are estimated. In this regard, it is possible to evaluate the economical benefits of adjustment of the control set-points. Moreover, such an energy characteristics may be used to assist the decision of choosing appropriate safety margins, i.e. distance between set-points and specification limits, where there is a trade off between manufactured product safety and the economical benefits. The designed method, could then be used in future controller tuning, e.g. an online control set-point optimiser. For example the zone humidity model has been used for the control synthesis and tuning. The most promising results to date have been obtained by applying a nonlinear compensator, see Zajic et al. (2009), but there is now a need to take a decision via quantification of benefits prior to implementing on a real plant.

The paper is structured as follows: The plant details are given in Section 2. The zone temperature and humidity models are given in Section 3 together with the control valve characteristics. These are subsequently used for the energy consumption analysis given in Section 4. Further work and conclusions are given in Section 5.

2. PLANT DETAILS

The manufacturing requirements in ADC UK are that the environmental conditions during the production must be highly stable, where the air dry-bulb temperature has to be lower than 24°C and the air relative humidity lower than 20% (corresponding dew-point temperature −0.28°C). There are over 70 separate HVAC systems used to condition the air within the manufacturing areas and each has its own dedicated control system. A typical HVAC system set-up is shown in Figure 1.

Consider Figure 1, at the point where the return air is extracted by suction from the controlled zone and passed through the main duct to the mixing section, in which the return air is mixed with the supply air from the fresh air plant (FAP) in the ratio of 17:3. The air mixture then progresses through to the dehumidification unit (DU), where it is dehumidified. The dehumidified air is progressed to the air handling unit (AHU), where the air is heated or cooled depending on the operating requirements. Note, that the controlled zone is a clean room production area, where, in order to avoid any environmental contamination by dust and other air pollutant, a higher air pressure is maintained when compared to the atmospheric pressure. This is achieved by having lower air outflow from the controlled zone than the air inflow, hence part of the air ventilates through the gaps around the doors and windows. A portion of the return air (approximately 30%) is led through the bypass directly to the AHU. For the comfort of personnel the designed air flow through the controlled zone is $2\text{m}^3\text{s}^{-1}$, however this would require a larger size of the DU than that currently installed. Therefore, a small amount ($0.61\text{m}^3\text{s}^{-1}$) of the return air is led through the bypass directly to the AHU allowing for the use of an undersized DU, in this case it is the Munster MX5000. The volumetric flow in the bypass is controlled by dampers, whose position is currently fixed.

2.1 Dehumidification unit

The DU comprises of a large wheel with a honeycomb structure coated with a moisture absorbent desiccant, in this case, silica gel. The functionality of the DU is described as follows. The wheel rotates with a constant angular velocity of approximately 1/10rpm. The processed air is driven through the lower part of the wheel, which is approximately 3/4 of its overall surface. The silica gel absorbs the moisture from the air, however, this process is exothermal and the air is additionally warmed. Consequently, to remove the absorbed water, from the silica gel, the inverse process has to be applied, hence heat needs to be provided. The hot reactivation air is blown through the upper part of the dehumidification wheel, where external air is used for the reactivation, which is heated with a gas burner and after use is then exhausted to atmosphere.

The gas burner is controlled by the central control system, where a PI controller is utilised. The controller drives the electronic actuator mounted to the gas valve, by applying the control action, denoted u_1. The gas valve itself has a lower safety limit of 26% (low fire), hence is never switched off due to the safety issues at the ignition stage. The feedback signal for the controller is the return air dew-point temperature and the current designed set-point is −11°C. The dew-point temperature sensor is denoted D in Figure 1.

2.2 Air handling unit

The air handling unit has three main separate components the cooling coil unit (CCU), the heating coil unit (HCU) and the main fan. The purpose of the AHU is to maintain the conditioned air dry-bulb temperature at a given set-point. Both, the CCU and HCU, are in fact water to air heat exchangers. The CCU is supplied with chilled water, denoted CHW in Figure 1, from the chilled water plant. The HCU is supplied with low temperature hot water (LTHW) from the LTHW plant, where the water is heated in gas driven boilers. Since the DU heats the processed air a normal operation condition of the AHU is that the HCU is disabled by the control system and hence is not considered hereafter, in order to simplify the energy consumption analysis.

The cooling capacity of the CCU is regulated by means of the chilled water mass-flow rate, i.e. by the cooling valve. The valve is controlled by a PI controller and the control action signal is denoted u_2. The feedback signal for the controller is the return air dry-bulb temperature and the designed set-point is 20 ± 2°C. The decision logic of the local controller decides whether to use the CCU or HCU,

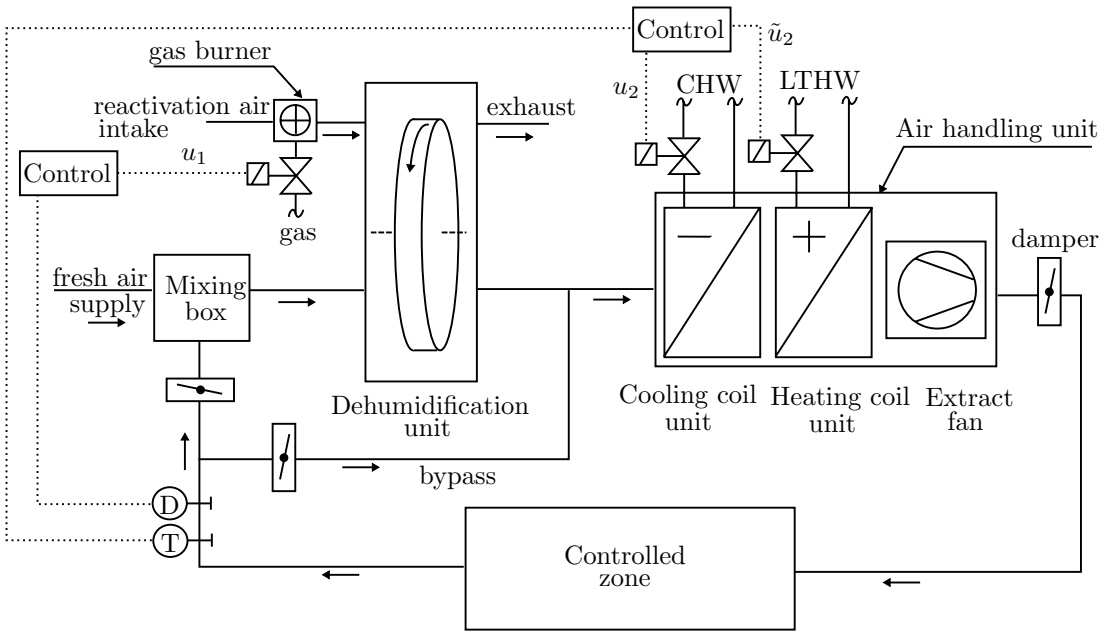


Fig. 1. The schematic diagram of the investigated HVAC system.

hence both units do not operate at the same time. The dry-bulb temperature sensor, denoted T, is installed in the return duct, see Figure 1.

2.3 Fresh air plant

The main purpose of the FAP is to pre-dehumidify and pre-cool the outdoor air. The main component of the considered FAP is another AHU. The dehumidification is performed by cooling the air below the dew-point temperature, i.e. latent cooling. The dew condensates on the body of the cooling element of the CCU, hence the coil temperature has to be lower than the outdoor dew-point temperature. For this reason the CHW temperature set-point is 5°C.

The FAP maintains the fresh air supply temperature, to the downstream HVAC, system between 5°C and 10°C. In the case, when the outdoor air temperature is above 10°C the CCU is switched on and the air is cooled, i.e. sensible cooling. In the case, when the outdoor air dew-point temperature is above the cooling element temperature, latent cooling occurs. The lower limit of 5°C is a part of the frost protection scheme for use in the winter season.

3. HVAC SYSTEM MODELLING

3.1 Zone temperature and humidity models

The zone humidity model, denoted $\mathcal{M}_1$, and zone temperature model, denoted $\mathcal{M}_2$, are modelled as a first order, nonlinear, multi-input single-output (MISO) systems, see Larkowski et al. (2009) and Zajic et al. (2010), respectively. The models are given by

$$\begin{aligned} y_{D,t} =& \theta_1 y_{D,t-1} + \theta_2 u_{1,t-d1} + \theta_3 u_{1,t-d1} u_{3D,t-d1} \\ &+ \theta_4 y_{D,t-1} u_{1,t-d1} + \theta_5 \end{aligned} \quad (1)$$

and

$$\begin{aligned} y_{T,t} =& \theta_1 y_{T,t-1} + \theta_2 u_{1,t-d1} + \theta_3 u_{2,t-d2} + \theta_4 u_{3T,t-d1} \\ &+ \theta_5 u_{4T,t-d1} + \theta_6 u_{1,t-d1} y_{T,t-1} + \theta_7 u_{2,t-d2} y_{T,t-1} \\ &+ \theta_8 u_{1,t-d1} u_{2,t-d2} y_{T,t-1} + \theta_9. \end{aligned} \quad (2)$$

Here, $y_{D,t}$ and $y_{T,t}$ are outputs, namely, the air dew-point temperature and dry-bulb temperature, respectively, at the discrete time index t, measured by sensors located in the main return duct. The inputs are: scaled gas valve position $u_{1,t}$ within 0 and 1, scaled cooling valve position $u_{2,t}$ within 0 and 1, fresh air supply dew-point temperature $u_{3D,t}$, fresh air supply dry-bulb temperature $u_{3T,t}$ and outdoor air dry-bulb temperature $u_{4T,t}$. The quantities $d1 \geq 1$ and $d2 \geq 1$ denote the normalised integer valued time delays which relate to the system time delay expressed as an integer multiple of the sampling interval, while $\theta_{1,\ldots,9}$ denote the parameters of the models. The system input and output signals are given in [°C], except the scaled valve positions, which are without units.

The input and output signals were acquired on 16th June and 17th June 2010, respectively, where the first data set is exclusively used for parameter estimation and the second data set is used for the model validation. The sampling interval is 5 seconds, $d1 = 26$ and $d2 = 16$ samples. The parameters of model $\mathcal{M}_1$ and $\mathcal{M}_2$ are given in Table 1.

Table 1. The model parameters of models $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$. (For the space reasons the parameter $\theta_9 = -0.15$ of model $\mathcal{M}_2$ is given in this caption.)

Model	θ_1	θ_2	θ_3	θ_4
$\mathcal{M}_1$	1.00	-0.37	1.26×10^{-2}	-1.34×10^{-2}
$\mathcal{M}_2$	1.00	3.34×10^{-2}	0.13	0.02
$\mathcal{M}_3$	0.99	0.86	0.26	-3.86×10^{-2}
	θ_5	θ_6	θ_7	θ_8
$\mathcal{M}_1$	7.05×10^{-2}	-	-	-
$\mathcal{M}_2$	6.82×10^{-5}	-7.00×10^{-4}	-0.01	-7.66×10^{-4}
$\mathcal{M}_3$	1.76×10^{-2}	-	-	-

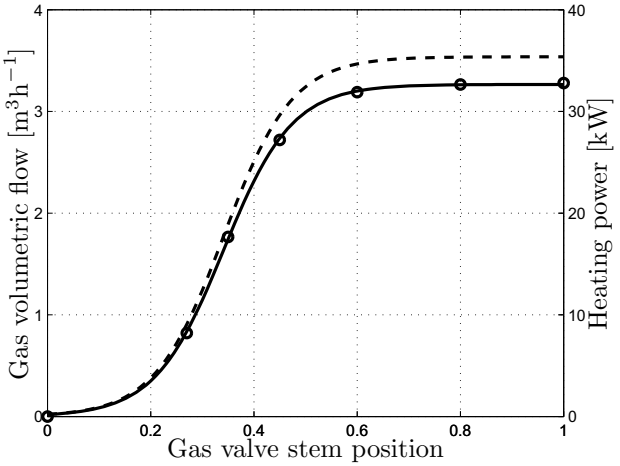


Fig. 2. Estimated steady state characteristics between gas valve position, $u_{1,\infty}$, and gas volumetric flow, $y_{G,\infty}$, (solid line) and heating power, $y_{H,\infty}$, (dashed line). The measured data are depicted by circles.

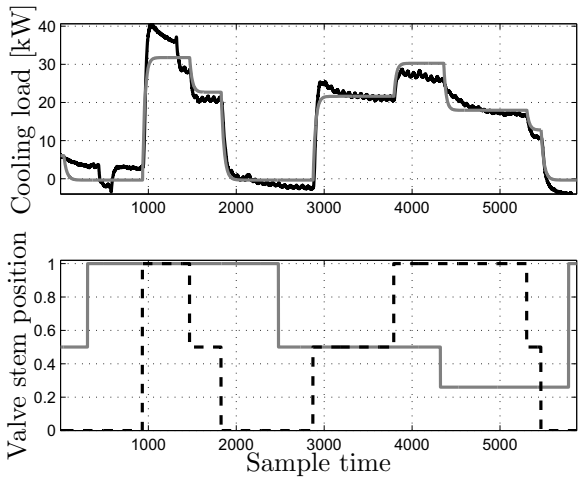


Fig. 3. The upper plot shows measured (black solid line) and simulated (grey solid line) cooling load. The lower plot shows gas valve position (grey solid line) and cooling valve position (black dashed line). Sampling interval 5 seconds.

3.2 Heating power

The gas valve characteristics takes the form of a S-shaped power curve, which can be expressed by a relationship based on a logistic growth function, see Taylor et al. (2007), i.e.

$$y_{G,\infty} = \frac{y_{G,max}}{1 + exp[-a(u_{1,\infty} - b)^{1/c}]}, \tag{3}$$

where $y_{G,\infty}$ denotes the steady state gas volumetric flow [$\mathrm{m^3h^{-1}}$], $u_{1,\infty}$ is a constant gas valve position, and $y_{G,max}$, a, b and c are constant coefficients. Utilising the Matlab function *fminsearch* the coefficients are found to be

$$a = 4.56,\ b = 0.34,\ c = 0.30,\ y_{G,max} = 3.27.$$

For the purpose of the energy consumption analysis the steady state relationship between gas valve position $u_{1,\infty}$ and heating power is required, i.e.

$$y_{H,\infty} = CV \times y_{G,\infty} \times \frac{1000}{3600}. \tag{4}$$

Here, $y_{H,\infty}$ is the heating power [kW] in steady state, CV is the gas calorific value [$\mathrm{m^{-3}MJ}$] and 1000/3600 is a constant (conversion of units). The gas calorific value was $CV = 39$ $\mathrm{m^{-3}MJ}$ on 16th June 2010. The measured and estimated gas volumetric flow and heating power steady state characteristics are given in Figure 2.

3.3 Cooling load

For the purpose of energy consumption analysis the steady state relationship between the cooling valve position $u_{2,\infty}$ and the cooling load, denoted $y_{C,\infty}$ [kW], is required to be known, i.e. the amount of heat removed from the processed air by the CCU such that the desired zone temperature set-point is achieved. Note, that the HCU is not considered here. In order to simplify the modelling task, sensible cooling only is considered. Subsequently, knowing the on coil air dry-bulb temperature, denoted $T_{in,t}$ [°C], and the discharge air dry-bulb temperature, denoted $T_{out,t}$ [°C], the cooling load is

$$y_{C,t} = \frac{\dot{m}c_a}{1000}\left(T_{in,t} - T_{out,t}\right), \tag{5}$$

where $\dot{m} = 2.38$ $\mathrm{s^{-1}kg}$ is air mass flow rate, $c_a = 1005$ $\mathrm{kg^{-1}K^{-1}J}$ is air specific heat capacity and 1/1000 is a constant (conversion of units).

The cooling load depends on many factors. The first main factor is the temperature difference between the processed air and the cooling coil body, caused mainly by the heating influence of the DU, hence the gas valve position is considered in the modelling stage. Assuming constant CHW temperature, the second main factor is the CHW mass flow rate controlled by the cooling valve position. There is a static nonlinear relationship between the scaled cooling valve position and scaled CHW mass flow rate, denoted $\tilde{u}_{2,t}$, which is given by the valve geometry and the CHW pressure. Assuming a linear type valve, this static nonlinearity can be expressed by, see Underwood (1999),

$$\tilde{u}_{2,t} = \frac{u_{2,t}}{\sqrt{u_{2,t}^2\left(1-\gamma\right)+\gamma}}, \tag{6}$$

where γ is the valve authority. The cooling load model, denoted $\mathcal{M}_3$, is modelled as

$$\begin{aligned} y_{C,t} =& \theta_1 y_{C,t-1} + \theta_2 \tilde{u}_{2,t-d4} + \theta_3 u_{1,t-d3}\tilde{u}_{2,t-d4} \\ &+ \theta_4 y_{C,t-1}\tilde{u}_{2,t-d4} + \theta_5. \end{aligned} \tag{7}$$

The parameters of the cooling load model (7) together with the valve model (6) were identified such that the simulation error is minimised utilising the Matlab function *fminsearch*. The cooling load model parameters are given in Table 1. The identified cooling valve authority is $\gamma = 0.4$, $d3 = 20$ and $d4 = 3$ samples. The measured cooling load, given by (5), and simulated cooling load are given in Figure 3.

Subsequently, the steady state relationship between the cooling valve position $u_{2,\infty}$ and the cooling load $y_{C,\infty}$, is given by

$$y_{C,\infty} = \frac{\theta_2\tilde{u}_{2,\infty} + \theta_3 u_{1,\infty}\tilde{u}_{2,\infty} + \theta_5}{1 - \theta_1 - \theta_4\tilde{u}_{2,\infty}} \tag{8}$$

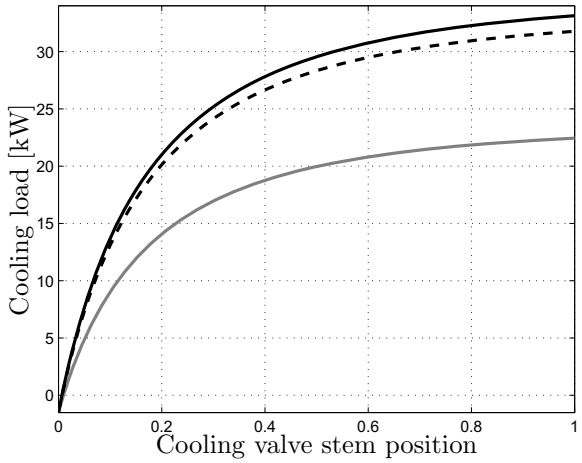


Fig. 4. The steady state cooling load characteristics for gas valve positions 0.3 (grey thick line), 0.5 (black dashed line) and 0.8 (thick solid line).

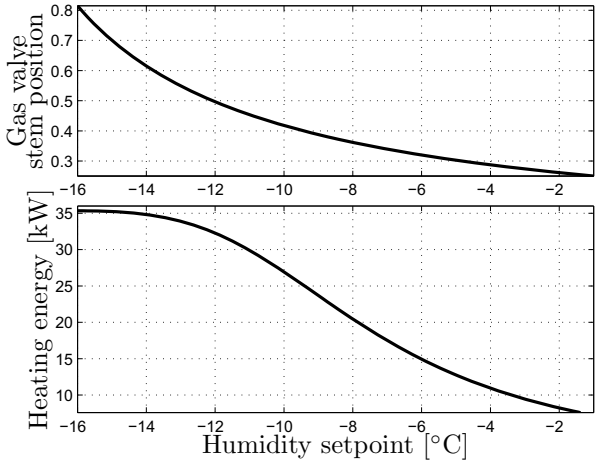


Fig. 5. The upper plot shows the steady state gas valve position plotted against the zone dew-point temperature set-point. The lower plot shows the steady state heating power consumption plotted against the zone dew-point temperature set-point.

together with the cooling valve static nonlinearity model (6). Finally, the steady state cooling load characteristics is given in Figure 4 for three different gas valve positions $u_{1,\infty} = 0.3$,0.5 and 0.8. It can be observed, that the cooling load increases as the gas valve is opened. Also, at $u_{2,\infty} = 0$ the cooling load is not null, as might have been expected. It is considered, however, that this has been caused by modelling errors and also by heat loss through the walls of the AHU.

4. ENERGY CONSUMPTION ANALYSIS

The aim is now to estimate the steady state characteristics between zone temperature and humidity control set-points and the HVAC system energy usage. This will allow the potential benefits of implementation of advanced control techniques to be evaluated. The overall task is, however, of rather a complex character and therefore the energy consumption analysis focuses only on power consumption of the DU and CCU, since it is believed that these are the main contributors to the overall energy usage on the site. The other contributors to the HVAC overall energy usage are the variable speed pumps for distribution of the CHW to the CCU and the chilled water plants, i.e. chillers, which also have nonlinear characteristics.

Firstly, attention is focused on the heating power consumption. The zone humidity model $\mathcal{M}_1$ can be expressed in steady state with respect to the gas valve position, i.e. control action, as follows

$$u_{1,\infty} = \frac{y_{D,\infty} - \theta_1 y_{D,\infty} - \theta_5}{\theta_2 + \theta_3 u_{3D,\infty} + \theta_4 y_{D,\infty}}. \tag{9}$$

Assuming, that in steady state the system output will be equal to the set-point, i.e. $y_{D,\infty} = r_{D,\infty}$, then (9) relates the set-point to the control action. The upper plot of Figure 5 shows the steady state relationship between the zone dew-point temperature set-point and the gas valve position, where $r_{D,\infty} = \langle -16, -1 \rangle$ °C and $u_{3D,\infty} = 6$ °C.

Subsequently, the computed gas valve position $u_{1,\infty}$ can be used to compute the heating power consumption by utilisation of equations (3) and (4), respectively. The lower plot of Figure 5 shows the steady state relationship between the zone dew-point temperature set-point and the heating power consumption.

Secondly, consider the cooling load steady state dependency on the zone dry-bulb temperature set-point, denoted $r_{T,\infty}$. The zone temperature model $\mathcal{M}_2$ can be expressed in steady state with respect to the cooling valve position as

$$u_{2,\infty} = \frac{y_{T,\infty} - p_1 - p_2 u_{1,\infty}}{p_3 + p_4 u_{1,\infty}}, \tag{10}$$

where

$$\begin{aligned} p_1 =& \theta_1 y_{T,\infty} + \theta_4 u_{3T,\infty} + \theta_5 u_{4T,\infty} + \theta_9, \\ p_2 =& \theta_2 + \theta_6 y_{T,\infty}, \\ p_3 =& \theta_3 + \theta_7 y_{T,\infty}, \\ p_4 =& \theta_8 y_{T,\infty}. \end{aligned} \tag{11}$$

Assuming, that in steady state $y_{T,\infty} = r_{T,\infty}$, then (10) relates the zone dry-bulb temperature set-point to the cooling valve position. This relationship is shown in the upper plot of Figure 6. The set-point is chosen to be $r_{R,\infty} = \langle 18, 24 \rangle$ °C, $u_{3T,\infty} = 10$ °C, $u_{4T,\infty} = 20$ °C and $u_{1,infty} = 0.3$, 0.5 and 0.8.

Finally, the steady state cooling valve position given by (10), together with the valve static nonlinearity (6) and the static cooling load model (8), can be use to compute the steady state relationship between the zone dry-bulb temperature set-point and the cooling load, which is shown in the lower plot of Figure 6.

4.1 Observations and final remarks

Originally, the dew-point temperature set-point was set to $r_{D,\infty} = -16$ °C. The parameter optimisation of the currently utilised PI controller allowed adjustment of the original control set-point closer to the specification limit, where the current set-point is $r_{D,\infty} = -11$ °C. As can bee seen in Figure 5 this yielded energy savings of approximately 5 kW, i.e. the relative power savings were 1kW per °C. It is anticipated here, that any further safe adjustment

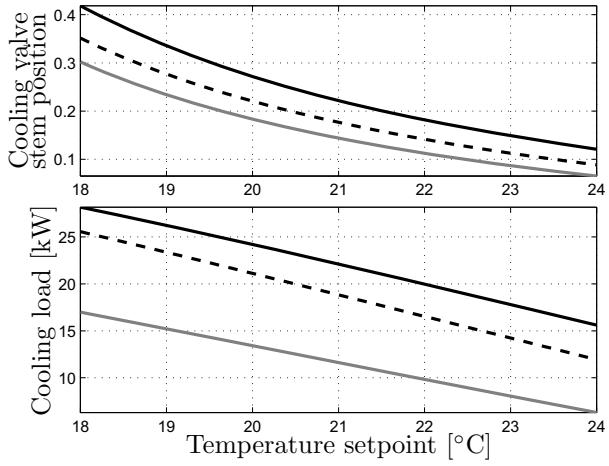


Fig. 6. The upper plot shows the steady state cooling valve position plotted against the zone dry-bulb temperature set-point. The lower plot shows the steady state cooling load plotted against the zone dry-bulb temperature set-point. Both, the upper and lower plot shows three different cases, where gas valve position is 0.3 (grey thick line), 0.5 (black dashed line) and 0.8 (thick solid line).

of the set-point is possible via utilisation of advance control techniques. For an example, adjusting the set-point further from -11 °C to -8 °C, would yield relative power savings of 3.3kW per °C, due to the S-shaped steady state power curve characteristics.

Noting the relationship between the gas valve position and the cooling load in Figure 6, it is possible to observe, that by adjustment of the zone dew-point temperature set-point, not only does the heating power reduce, but also does the cooling load.

Both, the heating power and the cooling load characteristics, were derived assuming certain outdoor weather conditions and consequently fresh air supply conditions, which enters the steady state characteristics as additional inputs $u_{3T,\infty}$, $u_{3D,\infty}$ and $u_{4T,\infty}$, respectively. Since the black-box, i.e. data driven, modelling approach has been chosen, these inputs must be carefully chosen. In particular, these inputs must be in the range of air temperatures present in the estimation data set. Whilst this is a limitation of the approach to energy consumption analysis (since the average temperatures would be probably more suitable) the results are indicative and would appear to justify the use of advanced control.

5. CONCLUSIONS AND FURTHER WORK

This paper has investigated the potential energy savings of a heating ventilation and air conditioning (HVAC) system based on the steady state relationship between the control set-points and the energy usage. It is anticipated that by utilisation of advanced control techniques, compared to the commonly used proportional and integral (PI) controller, a highly stable control performance can be achieved, which would subsequently allow adjustment of the control set-points. By operating closer to the specification limits, the highest profitability in terms of energy efficiency can be obtained.

The results presented for the specific period in question would suggest that given a realistic payback period, significant economical benefits can be achieved. Further research focuses on the use of simplified first principles models, which can also be used for energy analysis and control synthesis.

REFERENCES

Hill, D., Danne, T., and Burnham, K.J. (2009). Modelling and control optimistion of desiccant rotor dehumidification plant within the heating ventilation and air conditioning systems of a medical deivce manufacturer. In *Proceedings of the International Conference on Systems Engineering ICSE*, 207–212. Coventry, UK.

Kano, M. and Ogawa, M. (2009). The state of the art in advanced chemical process control in japan. In *IFAC Symposium on Advanced Control of Chemical Processes*, 11–26. Istanbul, Turkey.

Larkowski, T., Zajic, I., Linden, J.G., Burnham, K.J., and Hill, D. (2009). Identification of a dehumidification process of the heating, ventilation and air conditioning system using bias compensated least squares approach. In *Proceedings of the International Conference on Systems Engineering ICSE*, 296–305. Coventry, UK.

Levenmore, G. (2000). *Building control systems - CIBSE guide H.* Butterworth-Heinemann, Oxford.

Taylor, C.J., Shaban, E.M., Stables, M.A., and Ako, S. (2007). Proportional-integral-plus control applications of state-dependent parameter models. *J. of Systems and Control Engineering*, 221(7), 1019–1032.

Underwood, C. (1999). *HVAC control systems: Modelling, analysis and design.* E.&F.N. Spon, London.

Zajic, I., Larkowski, T., Hill, D., and Burnham, K.J. (2009). Nonlinear compensator design for HVAC systems: PI control strategy. In *Proceedings of the International Conference on Systems Engineering ICSE*, 580–584. Coventry, UK.

Zajic, I., Larkowski, T., Hill, D., and Burnham, K.J. (2010). Temperature model of clean room manufacturing area for control analysis. In *UKACC Int. Control Conf.*, 1251–1256. Coventry, UK.

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis Regular Paper
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fuel moisture content analysis as a basis for process monitoring of a BioGrate boiler

Alexandre Boriouchkine *. Alexey Zakharov.**
Sirkka-Liisa Jämsä-Jounela ***

* Aalto University, School of Science and Technology, Department of Biotechnology and Chemical Technology, Process Automation Research Group, 00076 Aalto, Finland. e-mail: aboriouc@cc.hut.fi (corresponding author, phone: +358-9-470 23178).
** Aalto University, School of Science and Technology, Department of Biotechnology and Chemical Technology, Process Automation Research Group, 00076 Aalto, Finland. e-mail: zakharov@cc.hut.fi
*** Aalto University, School of Science and Technology, Department of Biotechnology and Chemical Technology, Process Automation Research Group, 00076 Aalto, Finland. e-mail: sirkka-l@tkk.fi

Abstract: This paper considers the utilization of first principle models of a BioGrate boiler in a disturbance analysis study. The study focuses on the effect of fuel moisture content on the fuel combustion, since it is the most significant disturbance source in the boiler operation. The dynamic model of a BioGrate boiler, upon which the study is based, is heterogeneous, including solid and gas phases. Furthermore, the model considers chemical reactions in both gas and solid phases. In addition, fuel movement on the grate is included into the model. The energy required by the process is employed through a radiation function validated by industrial data. The model is implemented in a MATLAB environment and tested with industrial data. The results are presented and discussed.

Keywords: Power generation, Dynamic modelling, Biotechnology, Finite difference method, Renewable energy systems, System Identification

1. INTRODUCTION

The increasing utilization of renewable energy has created new energy efficiency challenges for industry. As biomass is one of the most important raw materials for renewable energy, all available biomass sources must be considered for energy production. However, the fuel properties of biomass tend to vary significantly depending, for example, on its origin, fuel processing and handling. Variable properties cause large fluctuations in combustion and thus, set challenges for an existing control strategy to keep the process within its constraints.

One of the latest successful processes developed, which uses wood waste as a fuel, is BioGrate-boiler technology, developed by MW Biopower. The combustion of wood waste is a very complex process involving several highly coupled chemical reactions. Furthermore, the operational conditions of the furnace greatly affect the yields of chemicals produced during the combustion process, i.e., fractions of tars, gases and char. Moreover, not only do the yields of chemicals differ under various combustion conditions, their reactivity in succeeding reactions also differ. In addition, significantly varying moisture content causes significant disturbance in the boiler operation. Fuel containing high amounts of moisture can occasionally cause a dramatic drop in a power production. These moisture-induced drops in power production are sometimes confused with a decrease in power production caused by a shortage of fuel in the boiler furnace. Since these two cases cannot be distinguished from each other quick enough, drops in power production are usually treated by adding more fuel. In the case of a fuel shortage, the added fuel would bring the process back to the specified operation conditions; nevertheless, in the case of increased fuel moisture, fuel addition causes an unexpected effect. The added fresh fuel will first continue to decrease the amount of produced heat and electricity, since fuel drying requires a significant amount of energy. After the fuel has dried it ignites and, consequently, raises the temperature of the flue gases causing an uncontrolled increase in steam production as a result. Finally, the unstable steam production leads to a turbine trip and thus financial losses due to unmet power production targets. Furthermore, the disturbances caused by the variation of the moisture content have a delayed impact on the operation of the boiler. Detecting an early disturbance can thus significantly improve the operation of the boiler.

This paper studies the detection of disturbances in boiler operation and is organized as follows: Section 2 describes the structure of a BioGrate boiler process; Section 3 presents the model and its aspects; and Section 4 studies the effect of moisture content on the combustion process. Section 5 presents a possible approach for the moisture content estimation. Section 6 summarizes the results.

2. PROCESS DESCRIPTION OF A BIOGRATE BOILER

A BioGrate consists of the following functional parts: a water filled ash space below the grate which, in turn, is located above the reservoir. The BioGrate is covered with a heat insulating refractory walls (combustion chamber) which reflects the heat radiation back to the grate Anon (2009).

The grate consists of several ring zones, which are further divided into two types of rings: rotating and fixed. Half of the grate rings rotate while the rest are fixed. Every second

rotating ring rotates clockwise and the others anticlockwise. This structure helps spreading fuel evenly upon the surface of the conical grate Anon (2009).

Fuel is fed into the centre of the grate from below. The fuel dries in the centre of the cone as a result of heat radiation, which is emitted by the combusting flue gas and reflected back to the grate by the grate walls. The dry fuel then proceeds to the outer shell of the grate where pyrolysis char gasification and combustion occur. The ash and carbon residues fall off the edge of the grate into the water-filled ash pit Anon (2009).

The air required in gasification and combustion is fed into the grate through the grate nozzles from the bottom of the grate (primary air) and through the nozzles of the combustion chamber (secondary air). In addition, in order to ensure clean combustion, additional air can be fed through the nozzles of the top of the combustion chamber (tertiary air) and the boiler walls. Burning produces heat that is absorbed in several steps. First, the evaporator absorbs the energy of the flue gases. Next, part of the energy of the flue gases is transferred to superheaters. In the third phase, the heat is transferred to the convective evaporator. Finally, economizers remove the remaining flue-gas energy Anon (2009).

The operation principle of a power plant is based on steam generation. As with any other bio power plant, a BioGrate power plant comprises a boiler, a turbine generator, a feed-water tank, a water treatment plant and a flue gas-cleaning system. Solid fuel is fed into the furnace of the boiler where it is combusted to generate heat and flue gases. As the flue gases contain fly ash which contains several harmful components, they are purified of the fly ash before being released into the atmosphere. The heat acquired from the fuel is then used for steam production.

The steam produced in the boiler is led to a generator turbine, which converts its mechanical energy into electricity. The steam pressure decreases as it performs the mechanical work; steam with decreased pressure, is then used to heat utility streams such as water Kiameh (2002). After the steam has released enough energy it condenses, condensed steam called a condensate which along with the pre-treated feed water, is fed into a feed-water tank. Inside the tank, liquid is heated with the bled steam from the turbine. This procedure increases the energetic efficiency of the process Kiameh (2002).

3. THE DYNAMIC MODEL OF A BIOGRATE BOILER AND ITS IMPLEMENTATION IN THE SIMULATION ENVIRONMENT

The current model of a BioGrate, which is utilized in the study, uses a walking grate concept modified for a BioGrate furnace. In addition, the chemical reaction kinetics were specially selected to fit the operational conditions of the BioGrate. Furthermore, an experimental model was used to model the radiation distribution inside the furnace.

The biomass bed reacts in a series of four different chemical reactions: drying, pyrolysis, char gasification and char combustion Peters and Bruch (2001). Active drying starts when the temperature of a particle reaches the boiling point of water. The high temperature of a furnace then initiates a pyrolysis reaction; which, in turn, produces three products: gases, char and tar. The gases are mainly composed of CO, CO_2, H_2, H_2O and C_1-C_3 hydrocarbons. Tar contains many organic components such as levoglucosan, furfural, furan derivatives and phenolic compounds Di Blasi (1996). Next, the model will be discussed in detail.

3.1 Assumptions

Several assumptions were made to simplify the modelling work and are listed in descending order of importance:

1. The system is one dimensional because the length of the grate is significantly longer than its height. Therefore, the temperature gradient in the horizontal direction is insignificant compared to that in the vertical direction.
2. Plug-flow gas assumption Zhou et al. (2005). The gas phase is assumed to be ideal Zhou et al. (2005), Kær (2005).
3. The solid is assumed to be a porous material Yang et al. (2003).
4. Diffusion in the gas phase is neglected, since the effect of convection on the transportation of the gas is significantly greater Peters and Bruch (2001).
5. Pressure dynamics are ignored because the release of gaseous species is negligible compared to the primary air flow; as a result, pressure evolution can be neglected Zhou et al. (2005).
6. Heat produced in char combustion is assumed to be retained in the solid phase Zhou et al. (2005).
7. No volume reduction (shrinkage) occurs during drying, pyrolysis and combustion Di Blasi (2009).
8. The temperature of the gas released from the solids is the same as that of the solids Zhou et al. (2005)
9. The temperature of the solids in a discretized block is uniform Zhou et al. (2005).
10. The heat capacity of the wood is assumed to be constant Kær (2005).
11. No heat loss.

Next, the simplified continuity equations are presented.

3.2 Solid phase continuity equation

The solid phase reacts through drying, pyrolysis and char combustion reactions:

$$\frac{\partial \rho_s}{\partial t} = -R_s \tag{1}$$

where ρ_s is the density of the solid phase, and R_s the overall reaction rate of the solid.

3.3 Energy continuity equation of the solid phase

The energy equation for the solid phase considers heat conduction; heat exchange between the phases; energy lost in the drying and pyrolysis reactions; and energy gained in char combustion:

$$\frac{\partial T_s}{\partial t} C_s \rho_s = \frac{\partial}{\partial x}\left(k_{cond} \frac{\partial T_s}{\partial x} \right) + k_{conv} v_p (T_f - T_s) - \ldots$$
$$R_{evap} \Delta H_{evap} - R_{pyr} \Delta H_{pyr} + R_{comb,C} \Delta H_{comb,C} - \ldots \quad (2)$$
$$R_{gasi,CO2} \Delta H_{gasi,CO2} - R_{gasi,H2O} \Delta H_{gasi,H2O}$$

where T_s is the temperature of the solid phase; C_s the heat capacity of the solid phase; ρ_s the density of the solid phase; x the vertical coordinate; k_{cond} the heat conduction coefficient of the solid phase; k_{conv} the heat convection the coefficient between the gas and solid phases; v_p the density number; T_f the temperature of the gas phase; and R_{evap} and R_{pyr} the reaction rates of the drying and pyrolysis. The reaction rates $R_{comb,C}$, $R_{gasi,CO2}$ and $R_{gasi,H2O}$ correspond to the reaction rates of the char combustion, gasification with carbon dioxide and gasification with water steam, respectively. ΔH_{evap} and ΔH_{pyr} are the reaction enthalpies of drying, and pyrolysis. The reaction enthalpies $\Delta H_{comb,C}$, $\Delta H_{gasi,CO2}$ and $\Delta H_{gasi,H2O}$ correspond to the reaction enthalpies of char combustion, gasification with carbon dioxide and gasification with water steam, respectively.

The radiation reflected from the grate walls to the fuel bed is described through boundary conditions. To describe the energy flux of the radiation energy, an experimental model was used. The model was defined from the experimental data of a BioGrate boiler located in Trolhättan, Sweden.

The heat conduction coefficient from Yagi and Kunii (1957) was used to describe heat conduction in the bed, while the heat conduction coefficient for wood particles was based on Janssens and Douglas (2004).

3.4 Gas phase continuity equation

The reacted solid components of wood are transferred to the gas phase; in addition, the the gas phase continuity equation considers gas flow:

$$\frac{\partial}{\partial t}(\rho_f \varepsilon_b Y_i) - \frac{\partial}{\partial x}(v_f \rho_f \varepsilon_b Y_i) = R_i \quad (3)$$

where ρ_f is the density of gas phase; ε_b the bed porosity; Y_i the mass fraction of the gaseous component i; v_f the gas flow velocity; and R_i the rate of formation of gaseous component i.

3.5 Energy continuity equation of the gas phase

Assuming no heat loss will occur, the energy continuity equation can be denoted as follows:

$$\frac{\partial h_f}{\partial t} \rho_f = -\frac{\partial}{\partial x}\left(\varepsilon_b v_f h_f \right) - k_{conv} v_p (T_f - T_s) + \ldots \quad (4)$$
$$R_{comb,CO} \Delta H_{comb,CO} + R_{comb,H2} \Delta H_{comb,H2}$$

where h_f is an enthalpy of the gas phase; ρ_f the density of the gas phase; ε_b the bed porosity; v_f the gas flow velocity; R_i the rate of formation of gaseous component i; k_{conv} is the heat convection coefficient between the gas and solid phases; v_p the density number; T_f the temperature of the gas phase; and T_s the temperature of the solid phase.

3.6 Chemical reactions of the model

The thermal decomposition of wood comprises three main chemical reactions: drying, pyrolysis and char gasification with char combustion. In general, the chemical reactions can be depicted using experimental or semi-experimental models. However, since Arrhenius dependence equations are simple to use and are also accurate, they have been used in this work.

3.7 Moisture evaporation

Typically, solid fuels used in power production contain moisture. Depending on the type of fuel, a fuel particle can contain various amounts of moisture. According to Thunmann et al. (2004), fuel particles can contain up to 60 wt. % of moisture while char residue can be as low as 10 wt. % of the wet wood. Water can be bound to the structure of a wood particle or reside in its pores. The drying model used in the current model is after Di Blasi et al. (2003).

3.8 Pyrolysis

After a particle has dried, the next reaction to occur is pyrolysis. In the pyrolysis reaction, a dry wood particle is decomposed into tar, volatile organic components and char. However, fractions of tar, gas and char in the product yield are strongly dependent on the reaction conditions of the combustion process. The current pyrolysis model is based on a study by Alves and Figueiredo (1989).

3.9 Combustion of pyrolysis gases

The yield of pyrolytic gases is around 85 wt. % under the operation conditions of a BioGrate boiler, since under these conditions the gasifying pyrolysis is the dominant pyrolysis mode. Therefore, a significant amount of energy used by the boiler comes from the combustion of gases; this fact indicates that the combustion of pyrolytic gases is the most important energy source. However, the composition of the gaseous products of pyrolysis, reported in Dupont et al (2009), suggests that carbon monoxide has the highest concentration in the pyrolytic gas, while the fraction of other combustible gases remains under 10 wt. %. Therefore, in order to ensure the acceptable accuracy of the model, while keeping the model simple, only the oxidation of carbon monoxide to carbon dioxide is considered. In addition to the oxidation of carbon monoxide, the combustion of hydrogen is also included in the model. Reaction rates of gas combustion reactions used in the model are presented in the study of Babushik and Dakdancha (1993)

3.10 Char conversion reactions

Char combustion in the model is based on that presented in Senneca (2007). This model was chosen because it is valid over the temperature range 440-800°C, which corresponds to the temperature of char combustion in a BioGrate. The

gasification reaction of char with carbon dioxide is based on kinetics reported by Senneca (2007).

3.11 Implementation of the dynamic model

The model was implemented in the MATLAB environment, in which a set of finite difference methods was used to solve the continuity equations. The overall solving algorithm is presented in Fig. 1.

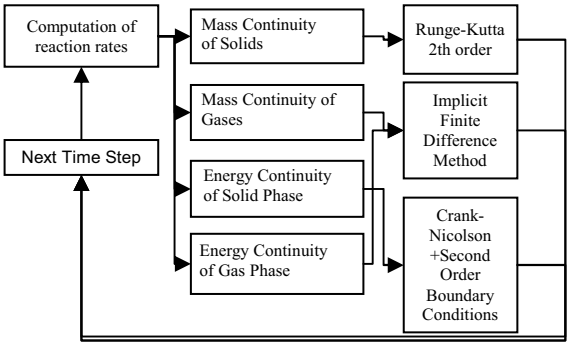


Fig. 1. Model solving algorithm.

4. THE EFFECT OF FUEL MOISTURE CONTENT ON THE FUEL COMBUSTION

In order to evaluate the effect of moisture on the overall combustion process, a simulation with varying moisture content is conducted. The simulation considers a moisture content variation in a form of step changes. The moisture content function included a step of 200 samples describing the moisture content of 45 wt. %, while the second step was 200 samples long representing a moisture content increase to 70 wt. %, and the nominal value of the function remains at 60 wt. %. The dynamic model was simulated for 2000 samples, during the simulation the first step started at the time period of 200 samples and returned to nominal value at the time period of 400 samples. The second step was introduced at the time interval of 500 through 700 samples.

The simulation shows that the decrease in the moisture content increased the surface temperature of the fuel layer almost immediately. However, the introduced moisture content increase started to decrease the surface temperature around 550 samples, nevertheless the ignition of drier fuel increased the mean surface temperature by 10 K despite the increased moisture content in fuel. Nevertheless, at the time point of 950 samples the increased moisture content decreased the surface temperature by 10 K. The mean gas temperature resembles the behaviour of the mean surface temperature, however, after the ignition moist fuel started to increase the mean temperature of the gas. This is a result of water steam reaction with the char which produces hydrogen. Increased moisture content increased the production of steam, which in turn, increased the reaction rate of char with the steam and, consequently, the production of hydrogen. While combusted, the hydrogen releases significant amounts of energy, thus, increasing the temperature of flue gases. However, despite the increased amount of reacting char the production of carbon monoxide decreases. This is the consequence of delayed pyrolysis, since the drying of moist fuel requires a longer time and delays the initiation of the pyrolysis. Furthermore, the increased moisture content decreases maximum gas temperature significantly along with the amount of produced flue gases. This decrease causes large fluctuations in the power production of the boiler. Fig. 2 presents the form of moisture content step function along with gas flows and maximum gas and solid temperatures while Fig 3 depicts mean surface and gas temperatures.

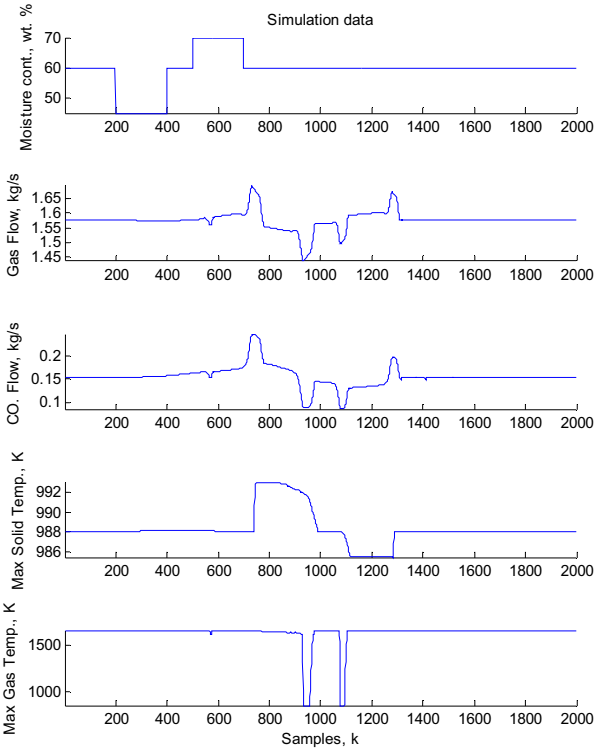


Fig. 2. The simulation with step changes in the fuel moisture content

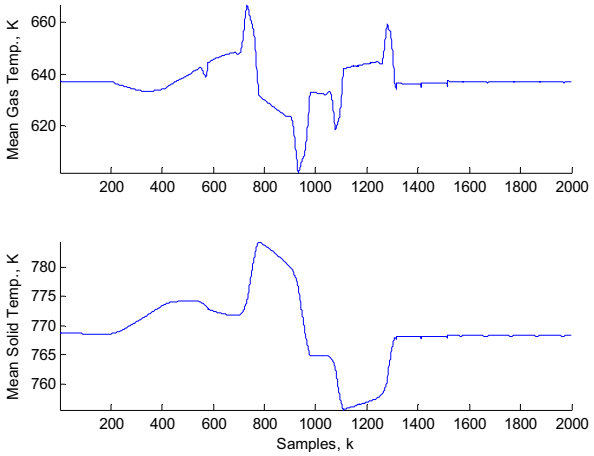


Fig. 3. The mean surface and gas temperatures

5. AN EXAMPLE FOR THE ESTIMATION OF THE MOISTURE CONTENT

This section proposes an example and a preliminary study on how the moisture content of the fuel can be estimated. This approach is based on system identification utilizing both a linear and a nonlinear model. The models are identified from three variables, which are usually measured at the power

plant: oxygen content, flue gas flow and flue gas temperature. These variables are then simulated with a dynamic model for 2,000 samples with varying moisture content at a point 1.5 m from the center of the grate. This point corresponds to 0.5 m^2 of the overall grate area. The first 700 samples are used for the model identification of both an ARMAX and a nonlinear ARX model. The identified models are then validated using the whole dataset. In addition, another 2,000 samples are generated to validate the models with an altered fuel moisture content; in addition, a random variation of moisture content is introduced at the period of 700-1,100 samples. The moisture content of the fuel is generated with a sum of sinusoidal signals. The data set used for the model identification is presented in Fig. 4., while Fig. 5 presents the second data set used for model validation. The data was normalized prior to the identification procedure. Both models are identified using the MATLAB system identification toolbox.

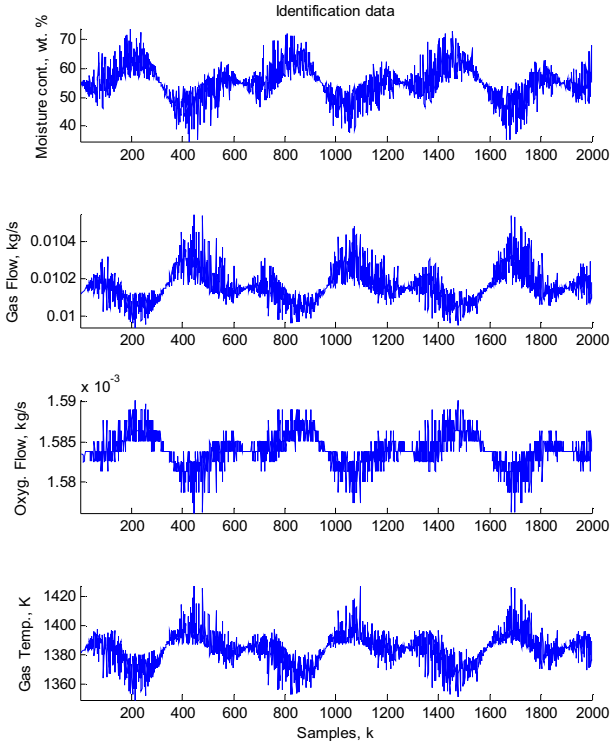


Fig. 4. Data used for model identification and validation.

The identified ARMAX model has the following structure:

$$A(q)y(k) = B(q)(u(k) - n_k) + C(q)e(k)$$

where the order of polynomial *A(q)* is 2, *B(q)* is 4 and *C(q)* is 5 while n_k = 100 samples.

The nonlinear ARX model utilizes a wavelet network with one unit to describe the nonlinear terms of the model while the current model output is a function of two previous inputs and outputs

The simulation results show that both models accurately predict the moisture content of the first dataset based on the flue gas flow, the temperature of the flue gas and the oxygen content. Although the ARMAX model is accurate, the nonlinear model exhibits better accuracy of the prediction, since not all relationships between inputs and outputs can be captured from the dynamic model. The simulation results of the first dataset for the identified ARMAX and nonlinear ARX models are presented in Fig. 6. For convenience, the figures present every fifth data sample instead of every sample.

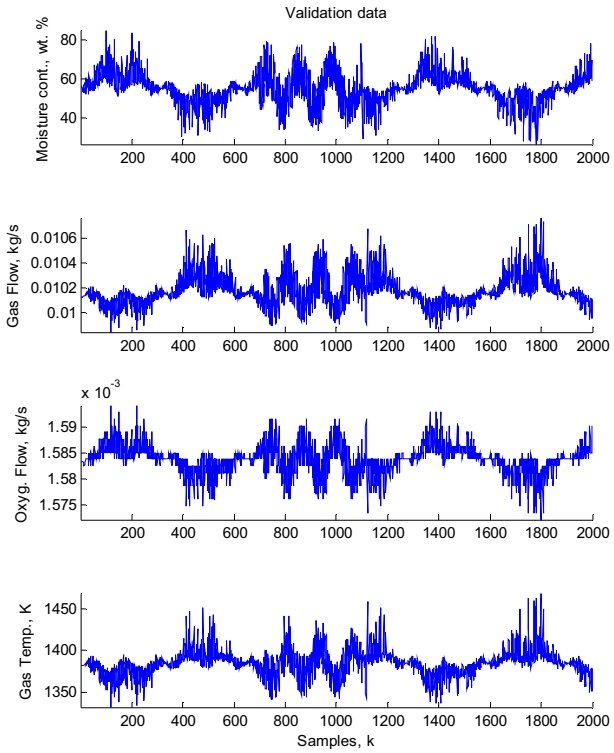


Fig. 5. Data set used for model validation.

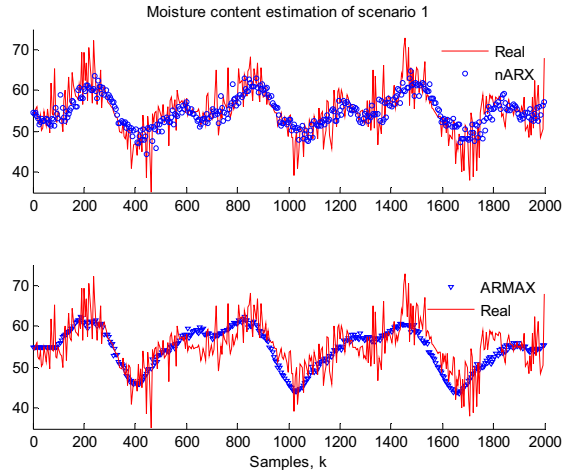


Fig. 6. Simulation results of the ARMAX and the nARX models with the data set presented in Fig. 2.

In order to evaluate the accuracy of the identified model, a different pattern of moisture variation with another 2,000 samples was generated with the dynamic model. The results show that with the second data set, the linear model is significantly less accurate than the nonlinear one. This is especially so at the time period between 700-1,100 samples, at which an additional disturbance is introduced and the linear model fails to predict the moisture content of the fuel. These results suggest that moisture content estimation requires a nonlinear model, since the linear model is not able to capture all the relations between the input and output

variables. The simulation results for the ARMAX and the nARX models are given in Fig. 7.

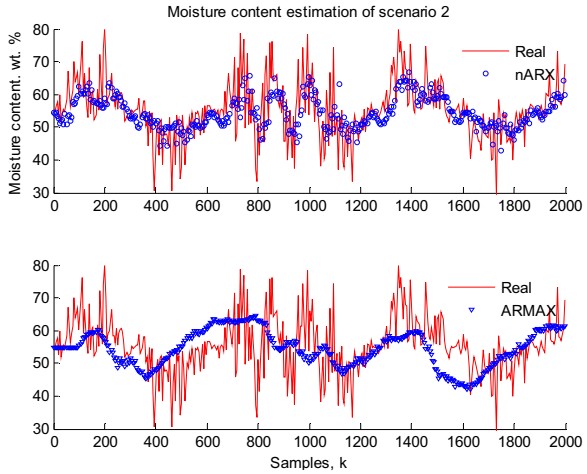


Fig. 7. Simulation results of the ARMAX and the nARX models with the data set generated with the second data set.

6. CONCLUSIONS

To summarize, the moisture content affects the combustion process significantly; therefore, it is important to detect the changes in moisture content as early as possible. The paper presented a possible approach for the moisture content estimation using linear and nonlinear models identified from the process data. The linear model failed to predict the moisture content when a significant disturbance was introduced, mainly due to the nonlinearity of the reactions and the heat transfer involved in wood combustion. In contrast to the linear model, the nonlinear model exhibited a better accuracy. Nevertheless, the study showed that moisture content could easily be estimated from the measurements available at the plant. However, further study requires evaluating the accuracy of the black box models under the changed operation conditions of the boiler, since in the presented simulations operation point of the boiler was assumed to be constant. Particularly, the combustion air feed was kept constant, however, in practice the volumetric air feed is not a constant. In addition to the air feed, many other variables, assumed in this study as constants, can, in practice, exhibit significant variations. Therefore, since the mechanistic models are typically valid under the changed operation point, it is highly motivated to simplify the dynamic model of a BioGrate boiler for the moisture content estimation.

REFERENCES

Alves, S., S., Figueiredo, J., L., A model for pyrolysis of wet wood, Chemical Engineering Science 44 (1989), pp. 2861-2869

Babushik, V., I., Dakdancha, A., N., Global kinetic parameters for high-temperature gas-phase reactions, *Combustion, Explosion, and Shock Waves* 29 (1993), pp. 464-489

Di Blasi, C., Heat, momentum and mass transport through a shrinking biomass particle exposed to thermal radiation, *Chemical Engineering Science* 5 (1996), pp. 1121-1132

Di Blasi, C., Branca, C., Sparano, La Mantia, B., Drying Characteristics of wood cylinders for conditions pertinent to fixed-bed countercurrent gasification, *Biomass and Bioenergy* 25 (2003), pp. 45-58

Di Blasi, C., Combustion and gasification rates of lignocellulosic chars, *Progress in Energy and Combustion Science* 35 (2009), pp. 121 – 140

Dupont, C., Chen, Li., Cances, J., Commandre, J.-M., Cuoci, A., Pierucci, S., Ranzi, E., Biomass pyrolysis: Kinetic modeling and expiremental validation under high temperature and flash heating rate conditions, *Journal of Analytical and Applied Pyrolysis* 85 (2009), pp. 260-267

Janssens, M., Douglas, B., *Wood and wood products*, Handbook of Building Materials for Fire Protection, Edited by Harper, C., A., McGraw-Hill 2004, 542 p.

Kær S., K., Straw combustion on slow moving grates-a comparison of model predictions with experimental data, *Biomass and Bioenergy* 28 (2005), pp. 307-320

Kiameh, P., *Power generation handbook*, McGraw-Hill, 2002, 557 p.

Peters, B., Bruch, C., A flexible and stable numerical method for simulating the thermal decomposition of wood particles, *Chemosphere* 42 (2001), pp. 481 – 490

Senneca, O., Kinetics of pyrolysis, combustion and gasification of three biomass fuels, Fuel Processing Technologies 88 (2007), pp. 87 - 97

Shin, D, Choi, S, The combustion of simulated waste particles in a fixed bed, *Combustion and Flame* 121 (2000), pp. 167-180

Thunman, H., Davidsson, K., Leckner, B., Separation of drying and devolatilization during conversion of solid fuels, *Combustion and Flame* 137 (2004), pp. 242 - 250

WWW, Anon, Wärtsilä Oyj, Brochure, http://service.wartsila.com/Wartsila/global/docs/en/power/media_publications/brochures/bioenergy_fi.pdf, 03.08.2009

Y. R. Goh, Y. B. Yang, R. Zakaria, R. G. Siddall, V. Nasserzadeh, J. Swithenbank, Development of an incinerator bed model for municipal solid waste incineration, *Combustion Science and Technology* 162 (2001), pp. 37-58

Yagi, S., Kunii, D., Studies on effective thermal conductivities in packed beds, *A.I.Ch.E. Journal* 3 (1957), pp. 373-381

Yang, Y., B., Yamauchi, H., Nasserzadeh, V., Swithenbank, J., Effects of fuel devolatilization on the combustion of wood chips and incineration of simulation municipal solid wastes in a packed bed, *Fuel* 82 (2003), pp. 2205-2221

Zhou H, Jensen, A. D., Glaborg, P., Jensen, P., A., Kavaliauskas, A., Numerical modeling of straw combustion in a fixed bed, *Fuel* 84 (2005), pp. 389-403

Proceedings of the 8th ACD 2010 European Workshop on Advanced Control and Diagnosis
Department of Engineering, University of Ferrara, Ferrara, Italy
18-19 November, 2010

Fault detection and accommodation of the boiler unit using state space neural networks

A. Czajkowski * **K. Patan** *,1

* *Institute of Control and Computation Engineering, University of Zielona Góra, ul. Podgórna 50, 65-246 Zielona Góra,*
e-mail: {*A.Czajkowski, K.Patan*}*@issi.uz.zgora.pl*

Abstract:
This paper deals with the application of state space neural network models to fault detection and accommodation of the boiler unit. The work describes approach based on the so-called instantaneous linearization of the already trained nonlinear state space model of the system. With obtained linear model it is possible to derive a new control law of the boiler unit in order to eliminate the fault effect in the case of faults. All data used in experiments are collected from the simulator of the boiler unit implemented in Matlab/Simulink.

Keywords: state space model, dynamic system, neural network, fault tolerant control.

1. INTRODUCTION

Recently, it has been observed an increasing development of the Fault Tolerant Control (FTC) systems. The two main advantages of the FTC systems which attract researchers, are maintaining the current performance as close to the desirable one, and preserve stability conditions in the presence of faults. Faults and equipment failures directly affect the performance of the control system and can result in large economic losses and even violation of the safety regulations. The existing FTC approaches can be split in two groups: passive and active ones. The second group can deal with both anticipated and unanticipated fault. Therefore, the application of these techniques seems to be more promising and effective than passive ones. In this paper an active fault tolerant control scheme is proposed. This work describes the approach based on the so-called instantaneous linearization of the already trained nonlinear state space model of the system. With obtained linear model it is possible to calculate a new control value for the boiler unit in the case of a fault. The goal of experiments is to minimize influence of the fault effect on working conditions of the system.

The paper is organized as follows. Section 2 presents a general description of the boiler unit and provides information about faulty scenarios considered. The state space neural networks are described in Section 3. Section 4 presents a fault detection and accommodation, while experimental results are included in Section 5.

2. BOILER UNIT

The object considered in this work is the laboratory installation developed at the Institute of Automatic Control and Robotics of the Warsaw University of Technology. The installation is dedicated for the investigation of diagnostic methods of industrial actuators and sensors Koj et al. (2005). The whole system consists of the boiler, storage tank, control valve with positioner, pump and transducers to measure process variables. The boiler is realized in the form of a horizontally placed cylinder, which introduces a strong nonlinearity into the static characteristic of the system. The scheme of the boiler unit with measurably available process variables marked is presented in Fig. 1.

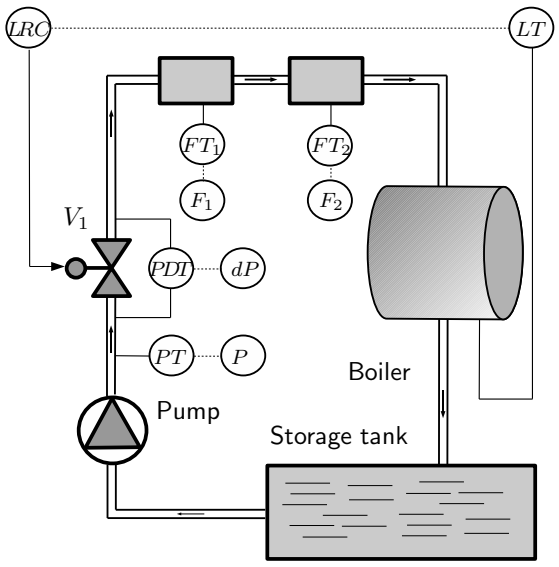


Fig. 1. The boiler unit.

In turn, the specification of process variables is shown in Table 1.The objective of the control system is to keep a required level of the water in the boiler. The control system uses the classical PID controller. The boiler unit together

[1] This work was supported in part by the Ministry of Science and Higher Education in Poland under the grant N N514 1219 33.

with control system was implemented in Matlab/Simulink. Simulations are performed with sample time equal to 0.05. The simulation model was validated using data acquired from the physical laboratory installation.
The model of the boiler unit makes it possible to generate a number of faulty situations. The specification of faults considered is included in Table 2. Considered faults are of different nature. As one can see in Table 2, there are multiplicative as well as additive faults. It is possible to set also the intensity of a fault.

Table 1. Specification of process variables.

Variable	Specification	Range
CV	control value	0-100 %
dP	pressure difference on the valve V_1	0-275 kPa
P	pressure before the valve V_1	0-500 kPa
F_1	flow (electromagnetic flowmeter)	0-5 m^3/h
F_2	flow (Vortex flowmeter)	0-5 m^3/h
L	water level in the boiler	0-0.5 m

Table 2. Specification of faulty scenarios considered.

Fault	Description	Type
f_1	fluid choking	partly closed (0.5)
f_2	level transducer failure	additive (-0.05)
f_3	positioner failure	multiplicative (0.7)
f_4	valve head or servo-motor fault	multiplicative (0.8)

3. STATE SPACE NEURAL MODEL

A fault can be represented by some nonlinear function f, which acts on the state equation of the system changing its characteristic and behaviour. From such a point of view state space models became an important class of models. Moreover such definition of a fault makes it possible to consider wide class of faults not only additive ones.
A very important class of dynamic neural networks is the State Space Neural Network (SSNN). Let $\boldsymbol{u}(k) \in \mathbb{R}^n$ be the input vector, $\boldsymbol{x}(k) \in \mathbb{R}^q$ - the output of the hidden layer at time k, and $\boldsymbol{y}(k) \in \mathbb{R}^m$ - the output vector. Then the state space representation of the neural model is described by the equations

$$\begin{aligned} \boldsymbol{x}(k+1) &= g(\boldsymbol{x}(k), \boldsymbol{u}(k)), \\ \boldsymbol{y}(k) &= \mathbf{C}\boldsymbol{x}(k) \end{aligned} \tag{1}$$

where $g(\cdot)$ is a nonlinear function characterizing the hidden layer, and $\mathbf{C}$ represents synaptic weights between hidden and output neurons. This equation can be shown in mathematical form:

$$\begin{aligned} \boldsymbol{x}(k+1) &= g(\mathbf{W}^x\boldsymbol{x}(k) + \mathbf{W}^u\boldsymbol{u}(k)), \\ \boldsymbol{y}(k) &= \mathbf{C}\boldsymbol{x}(k) \end{aligned} \tag{2}$$

where $\boldsymbol{W}^x$ and $\boldsymbol{W}^u$ are weights in neural connections of model. For the space state model the outputs which are fed back are unknown during training. As a result, state space models can be trained only by minimizing the simulation error. In spite of the fact that state space neural networks seem to be more promising that fully or partially neural networks, in practice a lot of difficulties can be encountered :

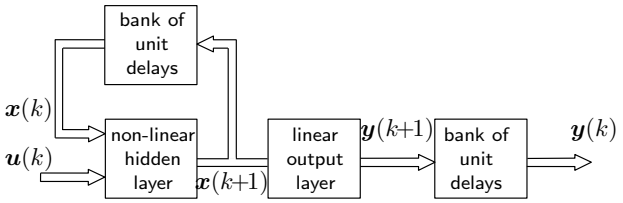


Fig. 2. Space state model, block scheme.

- model states do not approach true process states;
- wrong initial conditions can deteriorate the performance, especially when short data sets are used for training;
- training can become unstable;
- the model after training can be unstable.

In spite of that a very important property of the state space neural network is that it can approximate a wide class of nonlinear dynamic systems.

3.1 State space innovation form

In order to carry out the compensation of the fault effect, there is a need to design a state observer of the system. In this paper it is proposed to use the model in the so-called State Space Innovation Form (SSIF) represented as follows:

$$\begin{cases} \boldsymbol{x}(k+1) = g(\boldsymbol{x}(k), \boldsymbol{u}(k), \boldsymbol{e}(k)) \\ \boldsymbol{y}(k) = \mathbf{C}\boldsymbol{x}(k) \end{cases}, \tag{3}$$

where $e(k)$ is the error between the model output $\hat{y}(k)$ and measured system output $y(k)$. This equation can also be shown in mathematical form:

$$\begin{cases} \boldsymbol{x}(k+1) = g(\mathbf{W}^x\boldsymbol{x}(k) + \mathbf{W}^u\boldsymbol{u}(k) + \mathbf{W}^e\boldsymbol{e}(k)) \\ \boldsymbol{y}(k) = \mathbf{C}\boldsymbol{x}(k) \end{cases}, \tag{4}$$

where $\boldsymbol{W}^e$ are also weights in neural connections of model. The block-scheme of SSIF model shows Fig. 3.

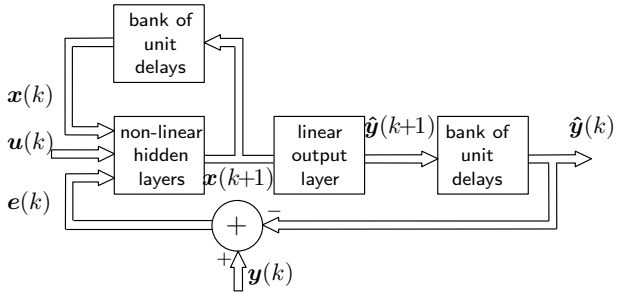


Fig. 3. SSIF model, block scheme.

The identified SSIF model can be regarded as an extended

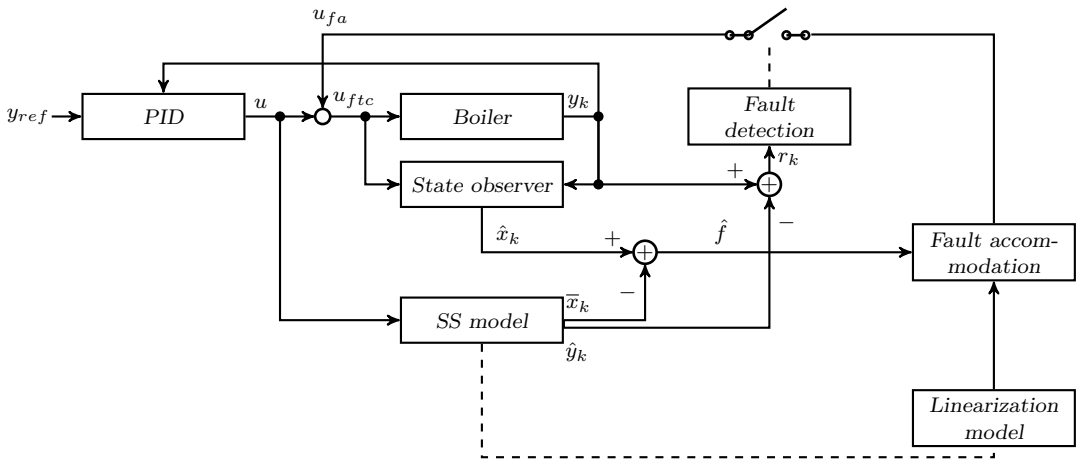


Fig. 4. Block scheme of presented FTC strategy.

Kalman filter for unknown nonlinear systems (Norgaard et al. (2000)). In the framework of fault tolerant control such kind of neural networks can be used as a state observer.

4. FAULT TOLERANT CONTROL

In the nominal condition system can be described as state equation:

$$\boldsymbol{x}(k+1) = h(\boldsymbol{x}(k), \boldsymbol{u}(k)). \tag{5}$$

When a fault occurs in the system above equation is extended by introduction of a fault function f and described as:

$$\boldsymbol{x}(k+1) = h(\boldsymbol{x}(k), \boldsymbol{u}(k)) + f(\boldsymbol{x}(k), \boldsymbol{u}(k)). \tag{6}$$

To model behavior of the system in nominal conditions, one can use the SSNN model with state described as:

$$\overline{\boldsymbol{x}}(k+1) = \hat{h}(\boldsymbol{x}(k), \boldsymbol{u}(k)), \tag{7}$$

where $\hat{h}$ is the estimation of the function h. As a state observer the SSIF model is used in the form

$$\hat{\boldsymbol{x}}(k+1) = h_f(\hat{\boldsymbol{x}}(k), \boldsymbol{u}(k), \boldsymbol{e}(k)) \tag{8}$$

Then the unknown fault function f can be approximated as:

$$\begin{aligned} \hat{f}(\hat{\boldsymbol{x}}(k), \boldsymbol{u}(k)) &= \hat{\boldsymbol{x}}(k) - \overline{\boldsymbol{x}}(k) \\ &= h_f(\hat{\boldsymbol{x}}(k), \boldsymbol{u}(k), \boldsymbol{e}(k)) - \hat{h}(\boldsymbol{x}(k), \boldsymbol{u}(k)). \end{aligned} \tag{9}$$

The fault effect can be eliminated/compensated by a proper definition of the augmented control u_{ftc}

$$\boldsymbol{u}_{ftc}(k) = \boldsymbol{u}(k) + \boldsymbol{u}_{fa}(k), \tag{10}$$

In this paper, the control $\boldsymbol{u}_{ftc}$ is determined via the instantaneous linearization of the state-space model. Linearization can be carried out by expanding function approximating system into Taylor series. The state-space model expanded into the Taylor series of first-order about the point $(\boldsymbol{x}, \boldsymbol{u}) = (\boldsymbol{x}(\tau), \boldsymbol{u}(\tau))$ have the form:

$$\begin{aligned} \hat{h}(\boldsymbol{x}(k), \boldsymbol{u}(k)) &= \hat{h}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)) + \frac{\partial \hat{h}}{\partial \boldsymbol{x}}\Big|_{(x,u)} \Delta x \\ &\quad + \frac{\partial h}{\partial \boldsymbol{u}}\Big|_{(x,u)} \Delta u \\ &= \hat{h}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)) + \hat{h}' \boldsymbol{W}^x (\boldsymbol{x}(k) - \boldsymbol{x}(\tau)) \\ &\quad + \hat{h}' \boldsymbol{W}^u (\boldsymbol{u}(k) - \boldsymbol{u}(\tau)) \end{aligned} \tag{11}$$

Now linearized state-space model can be presented in the form:

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k) + \boldsymbol{F} \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}(k) \end{cases} \tag{12}$$

where $\boldsymbol{A} = f'\boldsymbol{W}^x$, $\boldsymbol{B} = f'\boldsymbol{W}^u$, $\boldsymbol{F} = \hat{h}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau)) - \boldsymbol{A}x(\tau) - \boldsymbol{B}u(\tau)$.

With linearized state-space model in (12) and augmented control law in (10), the fault equations in (9) can be described as follows:

$$\begin{aligned} \hat{f} &= \hat{\boldsymbol{x}}(k) - \overline{\boldsymbol{x}}(k) = \\ &= \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}_{ftc}(k) + \boldsymbol{F} - (\boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}\boldsymbol{u}(k) + \boldsymbol{F}) \\ &= \boldsymbol{B}\boldsymbol{u}_{fa}(k) \end{aligned} \tag{13}$$

Therefore, the final conclusion can be made:

$$\boldsymbol{u}_{fa}(k) = \boldsymbol{B}^{-1}\hat{f}. \tag{14}$$

By adding $\boldsymbol{u}_{fa}$ to the control value one can achieve compensation of the fault effect occurring in the system.

4.1 FTC application

For the purpose of application of the presented FTC strategy, it is proposed to use the NNSSIF function from NNSYSID toolbox for building models (Norgaard et al.

(2000)). Main advantages of this function is mutual work of model as SSNN with $e(k) = 0$ and as SSIF in the case when output of the system is observed.
Through series of experiments (described in detail by Czajkowski and Patan (2009)), the network structure of the second order with 4 hidden neurons with hyperbolic tangent activation function was selected as the best performing one. The training set was generated using the control signal in the form of random steps with the values from the interval [0;0.5] and consists of 1500 samples. The training was carried out off-line for 100 epochs with the Levenberg-Marquardt algorithm. To verify the networks performance Sum of Squared Errors (SSE) was used. The neural connections in model structure are presented in Fig. 5.
Figure 6 shows the validation of the model with the use of input data which are generated by the random sinusoidal signal.

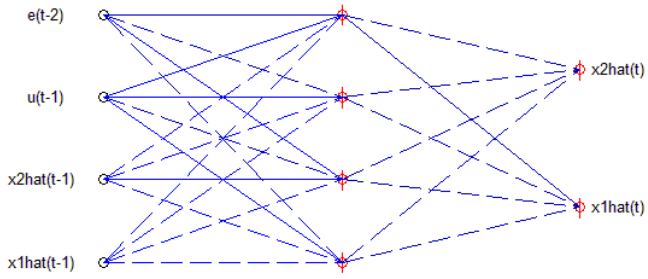


Fig. 5. Optimal structure of the neural connections in model.

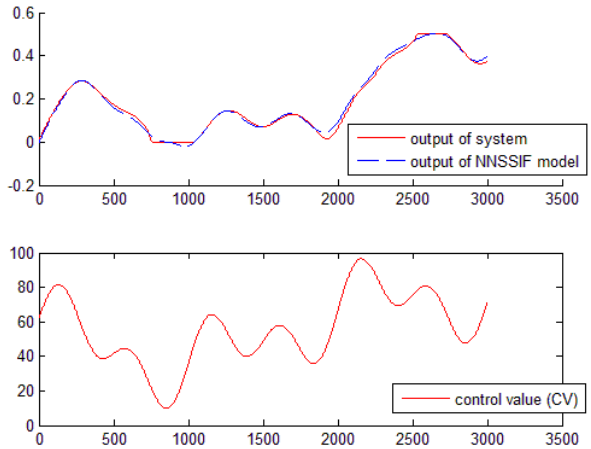


Fig. 6. Validation of model.

4.2 Fault detection

To introduce compensation to the system in case of a fault, the fault detection needs to be carried out. To evaluate residuals and to obtain information about faults, simple thresholding can be applied. If residuals are smaller than the threshold value, a process is considered to be healthy, otherwise it is faulty. In such a case the fault accommodation procedure starts to work determining additional input $\boldsymbol{u}_{fa}$ (see Fig. 4). For fault detection, the residual must meet the ideal condition being zero in the fault-free case and different from zero in the case of a fault. In practice, due to modelling uncertainty and measurement noise, it is necessary to assign thresholds larger than zero in order to avoid false alarms (Patan (2008), Basseville and Nikiforov (1993)). This operation causes a reduction in fault detection sensitivity. Therefore, the choice of the threshold is only a compromise between fault decision sensitivity and false alarm rate. In order to select the threshold, in this paper method of σ-standard deviation is used. Assuming that the residual is an $\mathcal{N}(m, v)$ random variable, thresholds are assigned to the values:

$$T = m \pm \sigma v \tag{15}$$

where m is mean value and v is standard deviation of residuals values and σ, in the most cases, is equal to 1, 2 or 3. The probability that a sample exceeds the threshold is equal to 0.15866 for $\sigma = 1$, 0.02275 for $\sigma = 2$ and 0.00135 for $\sigma = 3$, respectively. Based on residuals (set consisted of the 1500 samples) collected from the boiler simulator (Fig. 7), calculated as follows:

$$\boldsymbol{r}(k) = \boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k), \tag{16}$$

mean value m and standard deviation v are calculated as follows:

$$m = \frac{1}{N}\sum_{k=1}^{N} r_k = 0.0016 \tag{17}$$

$$v = \frac{1}{N-1}\sum_{k=1}^{N} (r_k - m)^2 = 0.0093 \tag{18}$$

In Table 3 thresholds values generated according to (15) with different σ levels are presented. Exceeding threshold calculated with the value of σ equal to 1, one can achieve the fastest fault detection, but with high probability of false alarm and for σ equal to 3, probability of false alarm is very low but detection time is relatively long.

Table 3. Thresholds for different σ values.

σ	**threshold range**
1	-0.0077 ÷ 0.0109
2	-0.0170 ÷ 0.0202
3	-0.0263 ÷ 0.0295

5. EXPERIMENTAL RESULTS

To validate above method a number of experiments with different fault scenarios has been carried out. Described method was tested in closed loop with PID controller, which was setup to keep value of 0.25 on the output of the boiler. Each fault was introduced to the system at 500 time instant. Fault detection was carried out with simple thresholding with $\sigma = 3$. Using $\sigma = 2$ or 1 would result in faster fault detection but also in a lot of false alarms. Smaller thresholds require more exact model which will be the subject of further researches. In the nominal condition, the fault value is in threshold range and $\boldsymbol{u}_{fa}$ is set to 0 which do not affect work of system, but in the case of a fault its value is changing and in this way it compensate

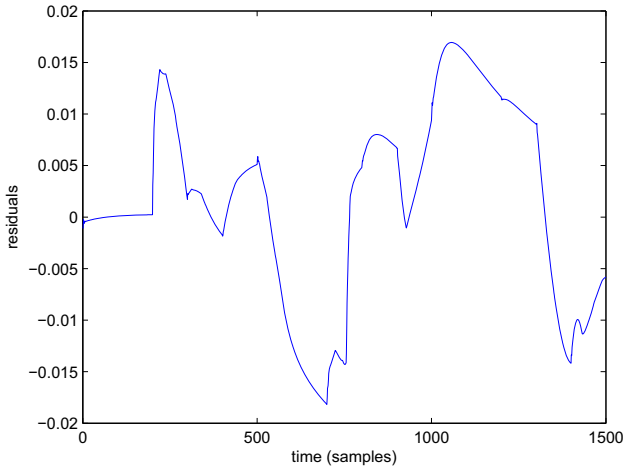


Fig. 7. Simulation residuals used for selecting of threshold.

the fault effect. In Figs 8-11 are presented charts of behaviour of the system in the case of faults listed in the Table 2. Each chart presents the output of the healthy system (solid line), output of the faulty system (dashed line) and output of the compensated system (dash-dotted line). In turn, efficiency indexes are presented in Table 4 and the detection times which are a periods of time needed for the detection of a fault measured from the time of the fault start-up, to a permanent, true decision about a fault is presented in Table 5.

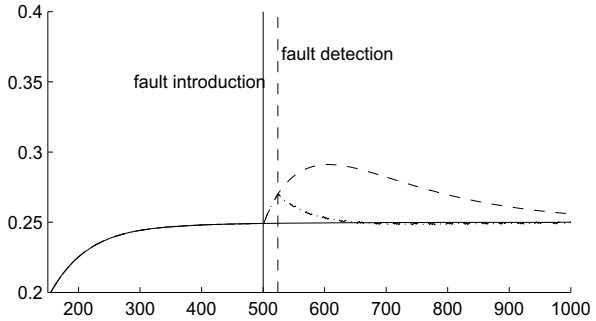


Fig. 8. Comparison of system work without fault, with fault f1 but without FTC and with fault f1 and with FTC.

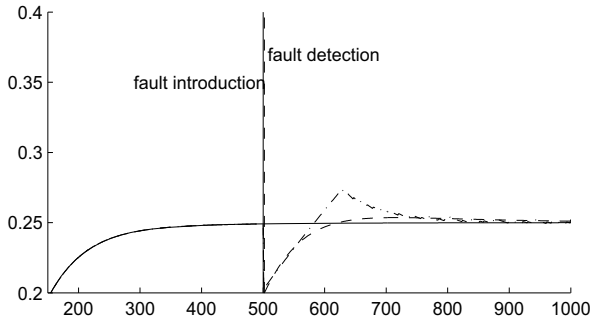


Fig. 9. Comparison of system work without fault, with fault f2 but without FTC and with fault f2 and with FTC.

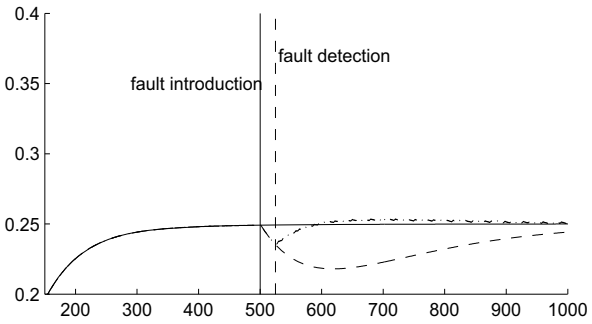


Fig. 10. Comparison of system work without fault, with fault f3 but without FTC and with fault f3 and with FTC.

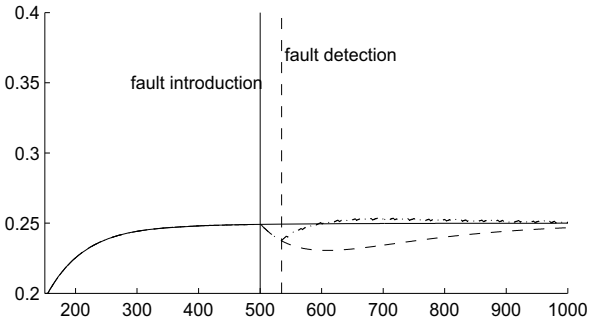


Fig. 11. Comparison of system work without fault, with fault f4 but without FTC and with fault f4 and with FTC.

Table 4. Results of experiments in form of SSE and percentage index.

	f_1	f_2	f_3	f_4
SSE without FTC	0.3447	0.0626	0.2187	0.0756
SSE with FTC	0.0139	0.0862	0.0069	0.0058
Improvement (%)	95.9815	-37.8339	96.8487	92.2801

Table 5. The detection times of the faults.

fault	t_{dt}
f_1	25
f_2	3
f_3	26
f_4	36

6. CONCLUSION

In the paper it was shown, that the proposed method makes it possible to improve the work of the boiler unit in the case of faults. As was shown through the experiments, the model in the form of the State Space Neural Network can be effectively and easily used to minimize the residuum defined as a difference between outputs of nominal and faulty system. Linearization of model in that form is very simple and efficient. As shown in Table 4, in three out of

four cases the fault was compensated in about 95%, which are very satisfactory results, but in the case of fault f2 the new augmented control law resulted even in setback due to specific nature of fault. This specific behavior needs to be taken in considiration in the further work on improving proposed method and model.
In spite of the fact that results are quite satisfactory we believe that they can be improved by improving the model of the system and this improvements are the subject of the future work. Improving the model would lead to faster fault detection without false alarms.

REFERENCES

Basseville, M. and Nikiforov, I.V. (1993). *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, Englewood Cliffs.

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. (2006). *Diagnosis and Fault-Tolerant Control*. Springer, Berlin.

Bouthiba, T. (2004). Fault location in ehv transmission lines using artificial neural networks. *International Journal of Applied Mathematics and Computer Science*, vol. 14, pp. 69–78.

Czajkowski, A. and Patan, K. (2009). Fault detection of the boiler unit using state space neural networks. In *Advanced Control and Diagnosis - ACD 2009 : 7th Workshop. Zielona Góra, Polska,*, CD–ROM.

Koj, J., Żelazny, M., and Kościelny, J. (2005). Laboratory stands for research and didactic purposes in the area of automatic control and diagnosis. *Measurements, Automatic Control, Monitoring*, 50(9), 261–264. Special issue, in Polish.

Korbicz, J., Koscielny, J.M., Kowalczuk, Z., and (eds) Cholewa, W. (2004). *Fault Diagnosis. Models, Artificial Intelligence, Applications*. Springer-Verlag, Berlin.

Leszczyński, M. and Syfert, M. (2005). Application of fault tolerant control system for the boiler laboratory set-up. In *Proc. XV National Control Conference, KKA'05, Warsaw, Poland*, volume II, 175–178. In Polish.

Ljung, L. (1999). *System Identification - Theory for the User*. Prentice Hall, Englewood Cliffs.

Norgaard, M., Ravn, O., Poulsen, N.K., and Hansen, L.K. (2000). *Neural Networks for Modelling and Control of Dynamic Systems*. Springer-Verlag, London.

Patan, K. (2007). Stability analysis and the stabilization of a class of discrete-time dynamic neural networks. *IEEE Trans. Neural Networks*, 18, 660–673.

Patan, K. (2008). *Artificial Neural Networks for the Modelling and Fault Diagnosis of Technical Processes*. Springer-Verlag, Berlin.

Patan, K., Witczak, M., and Korbicz, J. (2008). Towards robustness in neural network based fault diagnosis. *International Journal of Applied Mathematics and Computer Science*, 18(4), 443–454.

Patton, R.J. (1997). Fault-tolerant control: the 1997 situation (survey). In *Proc. IFAC Symp. on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS'97, Hull, U.K.*, 1029–1052.

Polycarpou, M. and Vemuri, A.T. (1995). Learning methodology for failure detection and accommodation. *IEEE Control Systems Magazine*, 15, 16–24.

Theillol, D., Cédric, J., and Zhang, Y. (2008). Actuator fault tolerant control design based on reconfigurable reference input. *International Journal of Applied Mathematics and Computer Science*, 18(4), 553–560.

Witczak, M. (2006). Advances in model-based fault diagnosis with evolutionary algorithms and neural networks. *International Journal of Applied Mathematics and Computer Science*, vol. 16, pp. 85–99.

Zamarreno, J.M. and Vega, P. (1998). Fault location in ehv transmission lines using artificial neural networks. *Neural Networks*, vol. 11, pp. 1099–1112.

Zhang, Y. (2007). Active fault-tolerant control systems: integration of fault diagnosis and reconfigurable control. In J. Korbicz, K. Patan, and M. Kowal (eds.), *Fault Diagnosis and Fault Tolerant Control, ISBN: 978-83-60434-32-1*, Challenging Problems of Science - Theory and Applications : Automatic Control and Robotics, 21–41. Academic Publishing House EXIT, Warsaw.

Index of Authors

Abdelkrim, N…….. *190*
Ammannito, M……. *64*
Aubrun, C……. *190*
Balas, G…… *223*
Barbato, S……. *241*
Bartoszewicz, B….. *73*
Belter, D……... *235*
Benmohammed, M... *326*
Bennouna, O……. *68*
Blasi, L……. *241*
Bokor, J…… *223*
Bonfè, M…….. *332*
Boriouchkine, A….. *374*
Botia, J. F……. *196*
Boussaid, B…….. *190*
Burnham, K. J…….. *207, 213, 339, 345, 368*
Calabrò, C…… *320*
Caliskan, F……... *46*
Casavola, A…….. *247*
Caselli, M……. *64*
Castaldi, P…… *217, 332*
Chafouk, H……... *68, 174*
Chiesa, S……... *229*
Cocconcelli, M……. *350*
Corpino, S…… *229*
Corradini, M. L…… *40*
Corraro, F……. *320*
Cristofaro, A. ……... *40*
Czajkowski, A…….. *380*
D'Elia, G…….. *350*
Dabo, M……... *174*
Dabrowska, A……... *146*
Dalpiaz, G…… *350*
Delvecchio, S……... *350*
Djurovic, Z……... *130*
Dunik, J…… *79*
Dziekan, L…… *85*
Fairusz, M…… *101*
Famularo, D……. *247, 257*
Ferracuti, F……... *290*
Filasova, A……... *184*
Formularo, D…… *247, 257*
Foszczynski, M…… *164*
Franzè, G…….. *247*
Freddi, A…….. *119*
Frédérik, T…… *141*
Friebel, T…….. *253*
Gagliardi, G…… *247*
Ghenai, A……... *326*
Giambò, R…….. *40*
Giantomassi, A……... *290*
Golonka, K……. *158*
Haber, R……. *135, 180, 253*
Hans, M…….. *170*
Hashemi-Nejad, H…….. *314*
Hekmati Vahed, S…….. *356*
Henry, D……. *4, 308*
Hill, D……. *368*
Hitzemann, U……. *207*
Hoblos, G……... *290*
Ignaciuk, P…….. *73*
Ilic, N…….. *91*
Incardona, M…….. *64*
Ippoliti, G……... *290*
Isaza, C……... *196*
Iyibakanlar, G……. *52*
Jalal, M. F. A…….. *101*
Jämsä-Jounela, S. L…… *374*
Jelicic, Z……. *271*
Kasprzak, A…… *152, 158, 164, 170*
Kiyak, E…….. *46, 52*
Kleinmann, S…….. *101, 146*
Kmiecik, W…… *152, 170*
Koechl, F…… *257*
Koller-Hodac, A……. *146*
Kopka, R……. *113*
Koszalka, L……. *152, 158, 164, 339*
Kret, P……. *339*
Krokavec, D……... *184*
Krokowicz, J……... *95*
Kvascev, G……. *130*
Labate, C. V……... *257*
Labecki, P……... *235*
Langlois, N……. *174, 290*
Larkowski, T. M……. *213, 368*
Leonardo, D…… *146*
Longhi, S…… *119, 290*
Luzar, M……. *202*
Malgorzata, S. M…… *213*
Martinez-Martinez, S……. *314*
Mattei, M…… *241, 257, 320*
Medici, G…… *229*
Mimmo, N…….. *217*

Miozza, L…….. *119*
Mokhtare, M…….. *362*
Monteriù, A…… *119*
Mouzakitis, A……. *339*
Negre, P. L……. *263*
N'Goran, Y……. *141*
Nozari, H. A…….. *356*
Orlowski, P……. *125*
Paczynski, A…….. *95, 107*
Parail, V……. *257*
Patan, K…….. *95, 380*
Patton, R. J……. *36*
Peñarrocha, I…….. *58*
Pettinari, S…….. *40*
Pieczyński, A……. *107*
Pillosu, S…… *278*
Pineda, I……. *263*
Pisano, A…… *271, 278, 284*
Pozniak-Koszalka, I…….. *152, 164, 170*
Puig, V…….. *263*
Puncochar, I…….. *79*
Raissi, T……. *290*
Rapaic, M…….. *271*
Robert, H…… *135, 180, 253*
Rouissi, F…….. *308*
Roux, J…….. *68*
Ruta, M…….. *345*
Sanchis, R…….. *58*
Sarmiento, H. O……. *196*
Sauter, D……. *314*
Schmitz, U…….. *135*
Scodina, S…….. *284*
Scordamaglia, V……. *320*
Seiler, P…….. *223*
Seybold, L…….. *95, 107*
Seydou, R…….. *308*
Shooredeli, M. A…… *356*
Simandl, S…….. *79*
Simani, S…… *217, 332, 356, 362*
Simon, C……. *302*
Skrzypczynski, P…… *235*
Sollazzo, A……. *320*
Stankovic, S…….. *91*
Staroswiecki, M……. *17*
Stetter, R……. *95, 101, 107, 146*
Stockmann, M…… *135*
Sumislawska, M…….. *213*
Tadic, P……. *130*
Theilliol, D…….. *302*
Thiery F…… *141*
Traore, A…….. *141*
Usai, E…….. *271, 278, 284*
Vanek, B…….. *223*
Visioli, A…….. *64*
Weber, P…….. *302*
Witczak, M…….. *85, 202*
Zabet, K…… *180*
Zaher, A…… *141*
Zajic, I…….. *368*
Zakharov, A…….. *374*
Zolghadri, A……. *308*