Performance Assessment for Process Monitoring and Fault Detection Methods

Kai Zhang

Performance Assessment for Process Monitoring and Fault Detection Methods

Kai Zhang
Duisburg, Germany

Dissertation, Duisburg-Essen University, 2016

ISBN 978-3-658-15970-2 ISBN 978-3-658-15971-9 (eBook)
DOI 10.1007/978-3-658-15971-9

Library of Congress Control Number: 2016954529

Springer Vieweg

© Springer Fachmedien Wiesbaden GmbH 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer Vieweg imprint is published by Springer Nature
The registered company is Springer Fachmedien Wiesbaden GmbH
The registered company address is: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

To my parents and Sissi

Preface

With the increasing demands on product quality and process operating safety, process monitoring and fault detection (PM-FD) has become an important area of research in recent decades. Numerous methods were developed in this area for different types of processes and applied to various industrial sectors. However, there is little work focusing on comparing and assessing their performance using a unified framework, and thus few suggestions and guidance for choosing an appropriate method can be provided to the practitioners. Therefore, the performance assessment study for PM-FD methods has become an area of interest in both academia and industry.

The first objective of this thesis is to assess the performance of basic FD statistics. The commonly used two statistics, namely, T^2 and Q are first examined. With the aid of χ^2 distribution, their differences to detect additive and multiplicative faults are revealed and compared under the statistical framework. Due to their low detectability to multiplicative faults, some alternative statistics are investigated.

Based on the basic FD statistics, different PM-FD methods have been proposed to monitor the key performance indicators (KPIs) of static processes, steady-state dynamic processes and dynamic processes including transient states. Thus, the second objective of this thesis is to assess the three classes of KPI-based PM-FD methods. Firstly, existing static methods are sorted into three categories based on the way to partition the KPI-correlated part from the KPI-uncorrelated part. A new EDD index is proposed to assess their performance to detect offsetting, drift and multiplicative faults. Secondly, two dynamic partial least squares (DPLS)-based methods for steady-state dynamic processes are compared, and their performance is assessed using EDD. Furthermore, the KPI-based PM-FD methods for general dynamic processes are introduced, some new developments are given.

Finally, to validate the theoretical developments, a case study on the Tennessee Eastman benchmark process that can be considered as a

steady-state dynamic process is performed to assess the two DPLS-based methods. In addition, a real large-scale hot strip rolling mill process is applied to assess the dynamic KPI-based PM-FD methods.

This work was done while I was with the Institute for Automatic Control and Complex Systems (AKS) at the University of Duisburg-Essen, Duisburg, Germany. I would like to express my deepest gratitude to my supervisor, Prof. Dr.-Ing. Steven X. Ding, for all the inspiration and help he provided during the course of the last three and a half years. I am sincerely grateful for his guidance and influence on my scientific research work. I would also like to thank Prof. Peng for his interest in my work. Without his valuable discussions and constructive comments, the thesis cannot have reached the current level.

Furthermore, I would like to express my appreciation to my colleagues, Zhiwen, Dr. Hao, Dr. Shardt, and Prof. Ge for all the impressive discussions and cooperations on my research topic as well as for their patience to go over the draft for this thesis. My special thanks should once again go to Dr. Shardt, who has shared his rich and valuable experiences on academic research and scientific writing.

In addition, I would like to thank Linlin, Changchen, Hao, Minjia, Sihan, Dongmei, Ying, and Yong for their support during my stay in AKS. My thanks also go to all my other AKS colleagues, Tim K., Chris, Shane, Tim D., Sabine, Dr. Köppen-Seliger, Klaus, Ulrich, Dr. Qiu, Dr. Li, and Dr. Jiang as well as my former colleagues, Prof. Lei, Prof. Shen, Prof. Dong, and Prof. Yang for their valuable discussions and helpful suggestions. Without them the completion of this thesis would not have been possible.

Finally, I would like to thank the China Scholar Council (CSC) for funding my stay in Germany.

Kai Zhang

Contents

Preface	VII
List of Figures	XIII
List of Tables	XVII
List of Notations	XIX
1 Introduction	1
1.1 Background and basic concepts	1
1.2 Motivation for the work	4
1.2.1 Basic FD test statistics	4
1.2.2 KPI-based PM-FD methods for static process . . .	6
1.2.3 KPI-based PM-FD methods for steady-state dynamic processes	8
1.2.4 KPI-based PM-FD methods for dynamic process .	9
1.2.5 Performance evaluation	9
1.3 Objectives	10
1.4 Outline of the thesis	11
2 Basics of fault detection and performance evaluation techniques	15
2.1 Technical description of static processes	15
2.2 Technical description of dynamic processes	17
2.3 FD performance evaluation indices	18
2.3.1 FDR and FAR	18
2.3.2 Expected detection delay	21
2.4 Simulation results	25
2.5 Conclusions	27

3	Common test statistics for fault detection	29
3.1	Background	29
3.2	Statistical properties of the T^2 - and Q -statistics	30
3.3	Detecting additive faults	33
3.4	Detecting independent multiplicative faults	36
3.5	Alternative statistics for detecting multiplicative faults	41
3.5.1	The extension of traditional methods	41
3.5.2	Wishart distribution-based methods	42
3.5.3	Information theory-based methods	44
3.5.4	Theoretical comparisons	46
3.6	Simulation results	49
3.6.1	Additive faults	49
3.6.2	Multiplicative faults	53
3.7	Conclusions	59
4	KPI-based PM-FD methods for static processes	61
4.1	Background	61
4.2	Classification of existing approaches	63
4.2.1	A direct method	63
4.2.2	Linear regression-based methods	64
4.2.3	PLS-based methods	67
4.3	Theoretical comparisons	70
4.3.1	Interconnections among the approaches	70
4.3.2	Geometric properties and computations	73
4.3.3	Remarks for PM-FD	80
4.4	Performance evaluation	81
4.4.1	A unified form of KPI-related fault detection	82
4.4.2	Calculation of FDR for $J_{T^2, \mathbb{P}}$ and $J_{Q, \mathbb{P}}$	83
4.4.3	Simulation results	84
4.5	Conclusions	88
5	KPI-based PM-FD methods for steady-state dynamic processes	91
5.1	Background	92
5.2	A comparison of two DPLS models	93
5.2.1	Two DPLS methods	93
5.2.2	The NIPALS alternative	96
5.2.3	Deflations and the complete DPLS model	98

5.3	EDD-based performance evaluation	100
5.3.1	KPI-based monitoring using DPLS models	100
5.3.2	Performance evaluation with respect to EDD	101
5.4	Simulation results	102
5.5	Conclusions	107
6	KPI-based PM-FD methods for dynamic processes	109
6.1	Background	109
6.1.1	Parity-space-based fault detection	111
6.1.2	Data-driven diagnostic observer	112
6.2	KPI-based FD using DO-based method	113
6.3	KPI-based FD using subprocess-based method	115
6.4	Simulation results	116
6.5	Conclusions	118
7	Benchmark study and industrial application	121
7.1	Case studies on TE process	121
7.1.1	A brief introduction to TE process	121
7.1.2	Results and discussion	124
7.2	Application to an industrial HSMR process	128
7.2.1	An introduction to the HSMR process	128
7.2.2	Results and discussion	130
7.3	Conclusions	136
8	Conclusions and future work	137
8.1	Conclusions	137
8.2	Future work	139
	Bibliography	141

List of Figures

1.1	Schematic description of an industrial process	2
1.2	Schematic description of PM-FD methods	3
1.3	Basics of statistical fault detection methods	5
1.4	Structure of the thesis	13
2.1	Demonstration of additive and multiplicative faults	19
2.2	Demonstration of false alarm rate and fault detection rate	20
2.3	Schematic description of detection delay using FAR and FDR	23
2.4	An example with FDR for a drift fault	24
2.5	EDD performance for constant additive faults	26
2.6	EDD performance for drift faults	26
2.7	EDD performance for constant multiplicative faults	27
3.1	Demonstration of J_{T^2} for detecting additive faults	34
3.2	Demonstration of J_{T^2} for detecting multiplicative faults .	39
3.3	Comparison of FDR for additive and multiplicative faults	40
3.4	Different thresholds for J_{T^2} and J_Q	50
3.5	Demonstration of J_{T^2} and J_Q for detecting additive faults	52
3.6	Schematic description of J_{T^2} and J_Q for detecting multiplicative faults	54
3.7	Performance of ϑ with different g_f and h_f , $m = 10$, $n = 10$	55
3.8	Performance of ϑ with different n , $m = 10$	55
3.9	Performance of J_{T^2} , $J_{T_n^2}$, J_Q and J_{Q_n} for Scenario 1 . . .	57
3.10	Performance of J_γ , $J_{\mathcal{T}}$ and $J_{\mathcal{D}}$ for Scenario 1	57
3.11	Performance of $J_{\mathcal{L}}$ for Scenario 1	58
3.12	Performance of different statistics for Scenario 2	58
4.1	Demonstration of the projections of the direct method . .	77
4.2	Demonstration of the projections of LS and PCR	78

4.3	Demonstration of the projection relationship between PLS and T-PLS	78
4.4	Demonstration of the projection relationship between PLS and C-PLS	79
4.5	Flops costed by the examined methods	80
5.1	Cross-validation results in the numerical example	103
5.2	The mixture of AIC and cross-validation result in the numerical example	104
5.3	Comparison of the original method to the alterative NIAPLS method	104
5.4	AIC results of the VAR model in the numerical example .	105
5.5	Residuals obtained by performing VAR model on $\mathbf{t}$	105
5.6	Comparison of EDD in the numerical example	106
6.1	Detection performance for Scenario 1	117
6.2	Profile of variables in Scenario 2	118
6.3	Fault detection performance for fault Scenario 2	118
6.4	Profile of variables in Scenario 3	119
6.5	Fault detection performance for fault Scenario 3	119
7.1	Schematic description of the TE process	122
7.2	Detection performance for fault-free case using two DPLS methods	125
7.3	Detection of fault 1 using two DPLS methods	126
7.4	Probability distribution of DD for fault 7 using two DPLS methods	127
7.5	Probability distribution of DD for fault 4 using two DPLS methods.	127
7.6	Schematic description of a large-scale FMP	129
7.7	Schematic description of the stand in FMP	130
7.8	Normal distribution plot of residual signals using DO-based method	131
7.9	Normal distribution plot of residual signals for subprocess 1 and 2	132
7.10	Monitoring result for Scenario 1	133
7.11	Monitoring result for Scenario 2	133
7.12	Monitoring result for Scenario 2 using DO-based method .	134
7.13	Monitoring result for Scenario 3	134

7.14 Monitoring result for Scenario 3 using DO-based method . 135

7.15 Monitoring result for Scenario 4 135

7.16 Monitoring result for Scenario 4 using DO-based method . 136

List of Tables

3.1	Comparison of different test statistics for multiplicative faults	46
3.2	FDR for different additive faults (J_{T^2}/J_Q)	52
4.1	Summary of projectors	76
4.2	Information about KPI-correlated subspaces	77
4.3	Summary of the computational complexity and parameter	80
4.4	EDD for different KPI-related faults for the numerical example	85
4.5	EDD for KPI-unrelated faults in numerical example . . .	86
4.6	EDD for different multiplicative faults	87
4.7	EDD for different drift faults	88
5.1	Original algorithm for the DDPLS method	94
5.2	Original algorithm for the IDPLS method	96
5.3	NIPALS algorithm for the DDPLS method	96
5.4	NIPALS algorithm for the IDPLS method	98
5.5	Comparison of the average EDD given by two DPLS methods	106
7.1	Process and manipulated variables of TE process	123
7.2	EDD of two DPLS methods for additive faults in TE process	126

Abbreviations and notations

Abbreviations

Abbreviation	Expansion
AIC	Akaike Information Criterion
ARMA	Auto-Regressive Moving Average
CCA	Canonical Correlation Analysis
CDF	Cumulative Distribution Function
C-PLS	Concurrent Partial Least Squares
DD	Detection Delay
DDPLS	Direct Dynamic Partial Least Squares
DO	Diagnostic Observer
DPLS	Dynamic Partial Least Squares
EDD	Expected Detection Delay
FAR	False Alarm Rate
FD	Fault Detection
FDR	Fault Detection Rate
FIR	Finite Impulse Response
FMP	Finishing Mill Process
HSMR	Hot Strip Mill Rolling
IDPLS	Indirect Dynamic Partial Least Squares
KPI	Key Performance Indicator
LS	Least Squares
LTi	Linear Time-Invariant
MSPM	Multivariate Statistical Process Monitoring
NIPALS	Nonlinear Iterative PARTial Least Squares
PCA	Principal Component Analysis
PCR	Principal Component Regression
PDF	Probability Distribution Function
PLS	Partial Least Squares
PM	Process Monitoring
PM-FD	Process Monitoring and Fault Detection
PRESS	Predicted REsidual Sum of Squares

PS	Parity Space
SVD	Singular Vector Decomposition
TE	Tennessee Eastman
T-PLS	Total Partial Least Squares
VAR	Vector Auto-Regression

Mathematical notations

Notation	Description
$\forall$	For all
$\sim$	Follow
$\approx$	Approximately equal
$\triangleq$	Defined as
$\gg$	Much larger than
$\rightarrow$	From...to
$\mathbb{R}^m$	Set of m -dimensional real vectors
$\mathbb{R}^{m \times n}$	Set of $m \times n$ -dimensional real matrices
$\mathbf{I}_m$	m -dimensional identify matrix
$\ \cdot\ _{\mathbf{E}}$	Euclidean norm of a vector
$\ \cdot\ _{\mathbf{F}}$	Frobenius norm of a matrix
$\ \cdot\ _2$	2-norm of a matrix
$ \cdot $	Determinant of a matrix or absolute value
c	A real constant
$\mathbf{y}$	A vector
$\mathbf{y}^{(i)}$	The i^{th} element of $\mathbf{y}$ or the i^{th} sample of $\mathbf{y}$
$\mathbf{y}^i$	i^{th} iteration of $\mathbf{y}$
$\mathbf{Y}$	A matrix
$\hat{\mathbf{y}}$	Estimate of $\mathbf{y}$ or KPI-related part in $\mathbf{y}$
$\tilde{\mathbf{y}}$	Residual of $\mathbf{y}$, or KPI-unrelated part in $\mathbf{y}$
$\mathbf{Y}^T$	Transpose of $\mathbf{Y}$
$\mathbf{Y}^{-1}$	Inverse of a square matrix $\mathbf{Y}$
$\mathbf{Y}^\dagger$	Pseudoinverse of $\mathbf{Y}$
$\mathbf{Y}^\perp$	Orthogonal complement of $\mathbf{Y}$
$\mathbf{f}$	Fault vector
f	Fault magnitude
$\mathbf{\Xi}$	Fault direction
λ	Eigenvalue
σ	Singular value or standard derivation

$\text{tr}(\mathbf{Y})$	Trace of $\mathbf{Y}$
$\text{diag}(\mathbf{y})$	A diagonal matrix with non-zeros elements $\mathbf{y}$
$\text{rank}(\mathbf{Y})$	Rank of matrix $\mathbf{Y}$
$\text{dim}\{\cdot\}$	Dimension of a space
$\text{span}\{\mathbf{y}\}$	Space spanned by $\mathbf{y}$
$\oplus$	Direct sum of two vector-spaced spaces
$\otimes$	Kronecker product
$\propto$	Proportional
$\text{E}(\cdot)$	Mean value/vector
$\text{Var}(\cdot)$	Variance value/vector
$\text{Cov}(\cdot)$	Covariance matrix
$\text{prob}(x)$	Probability of x
$\mathcal{N}_m(\mu, \Sigma)$	m -dimensioned Normal/Gaussian distribution with mean μ and covariance matrix Σ
J	Test statistic
J_{th}	Threshold
χ_m^2	Chi-squared distribution with m degrees of freedom
$\chi_m^2(\delta)$	Noncentral χ^2 distribution with m degrees of freedom and noncentrality parameter δ
$\mathcal{F}(a, b)$	$\mathcal{F}$ distribution with a and b degrees of freedom
α	Significance level
$\chi_{m,\alpha}^2$	Confidence value corresponding to α
$\mathcal{F}_\alpha(a, b)$	Confidence value corresponding to α
$\mathcal{W}_m(\Sigma, n)$	Wishart distribution with n degrees of freedom based on m -dimensional covariance matrix Σ
$e, \exp(\cdot)$	Base of natural logarithm, natural exponential function