Table of contents

LIST OF SYMBOLS **1**

General symbols, vectors and matrices 1

Greek symbols 7

Acronyms 8

CHAPTER 1 - THESIS OVERVIEW AND LITERATURE SURVEY **11**

1.1 AIM OF THE PROJECT 11

1.2 INTRODUCTION TO QUALITY AND STATISTICAL QUALITY MONITORING 12

1.3 MULTIVARIATE STATISTICAL TECHNIQUES FOR PROCESS MONITORING 14

1.3.1 Multivariate statistical process control for batch processes *18*

1.3.1.1 Nonlinear multivariate models 19

1.3.1.2 Multiway multivariate models 19

1.3.1.3 Multiple multivariate models 21

1.3.1.4 Preliminary data treatment for multivariate statistical methods 22

1.3.2 Multivariate image analysis *23*

1.4 THESIS OVERVIEW 24

1.4.1 Realtime quality estimation and length prediction in batch processes *24*

1.4.2 Multivariate statistical quality monitoring through image analysis *27*

1.4.3 Thesis roadmap *29*

CHAPTER 2 - MATHEMATICAL AND STATISTICAL BACKGROUND **31**

2.1 MULTIVARIATE STATISTICAL TECHNIQUES 31

2.1.1 Principal component analysis (PCA) *31*

2.1.1.1 PCA algorithm 34

2.1.1.2 Data collection, variable selection and data pre-treatment 36

2.1.1.3 Selection of the principal component subspace dimension 37

2.1.2 Projection on latent structures (PLS; partial least squares regression) *38*

2.1.2.1 Non-iterative partial least squares algorithm 40

2.1.2.2 Variable selection in PLS models 41

2.1.3 Monitoring charts *42*

2.1.3.1 Contribution plots, limits on the contribution plots, and relative contributions 47

2.1.4 Enhancement for multivariate statistical methods *50*

2.1.4.1 Multi-way methods, data unfolding and data synchronization/alignment 51

2.2 MULTIRESOLUTION DECOMPOSITION METHODS 54

2.2.1 Continuous and discrete wavelet transform *55*
2.2.1.1 Bi-dimensional wavelet transform 59
CHAPTER 3 - INDUSTRIAL PROCESS FOR THE PRODUCTION OF RESINS BY BATCH POLYMERIZATION 63
3.1 THE INDUSTRIAL PRODUCTION PLANT AND THE OPERATING RECIPE 63
3.1.1 Resin A *64*
3.1.2 Resin B *66*
3.1.3 P&ID of the production facility *67*
3.2 DATA ACQUISITION 68
3.2.1 Monitoring of the process variables *68*
3.3 EMPIRICAL MONITORING OF THE PRODUCT QUALITY 70
3.4 CHALLENGES FOR THE STATISTICAL MONITORING OF PRODUCT QUALITY 72
3.5 AUTOMATED QUALITY MONITORING THROUGH SOFT-SENSORS 72
CHAPTER 4 - SOFT SENSORS FOR THE REALTIME QUALITY ESTIMATION IN BATCH PROCESSES 75
4.1 QUALITY ESTIMATION IN RESIN A USING PLS MODELS 75
4.1.1 Single-phase PLS model *77*
4.1.2 Multi-phase PLS model *78*
4.2 INCLUDING TIME INFORMATION TO IMPROVE THE ESTIMATION PERFORMANCE 82
4.2.1 Improving soft sensor performance through lagged process variables *82*
4.2.2 Improving soft sensor performance through moving-average process data *85*
4.3 COMPARISON OF THE ESTIMATION PERFORMANCES 87
4.3.1 Reliability of the estimations *89*
4.3.2 Diagnosis of the soft sensors faults *90*
4.4 SOFT SENSOR FOR ESTIMATION OF QUALITY IN RESIN B 91
4.4.1 Estimation of the quality indicators *92*
4.5 CONCLUDING REMARKS 95
CHAPTER 5 - REALTIME PREDICTION OF BATCH LENGTH 97
5.1 DESIGN OF AN EVOLVING PLS MODEL FOR THE PREDICTION OF BATCH LENGTH 97
5.2 PREDICTION OF BATCH LENGTH IN THE PRODUCTION OF RESIN B 100
5.2.1 Prediction of Stage 1 length *100*
5.2.2 Prediction of Stage 2 length *103*
5.3 PREDICTION OF BATCH LENGTH IN THE PRODUCTION OF RESIN A 105
5.4 CONCLUDING REMARKS 106
CHAPTER 6 - INDUSTRIAL IMPLEMENTATION OF A SOFT SENSOR PROTOTYPE 107

6.1 INDUSTRIAL SUPERVISION SYSTEM 107
6.2 IMPLEMENTATION OF THE SOFT SENSOR 108
6.2.1 Architecture of the soft sensor 109
6.2.2 Matlab™ codes of the soft sensors 110
6.2.2.1 Prototype A 111
6.2.2.2 Prototypes B1 and B2 113

CHAPTER 7 - SURFACE CHARACTERIZATION THROUGH MULTIRESOLUTION AND MULTIVARIATE IMAGE ANALYSIS 115

7.1 PHOTOLITHOGRAPHY PROCESS AND INSPECTION TOOLS 115
7.2 IMAGE ANALYSIS THROUGH MULTIRESOLUTION AND MULTIVARIATE STATISTICAL TECHNIQUES 118
7.2.1 Image multiresolution denoising 119
7.2.2 Multivariate statistical surface monitoring methods 123
7.2.2.1 LER monitoring 123
7.2.2.2 Surface roughness monitoring 124
7.2.2.3 Edge shape monitoring 126
7.3 CASE STUDY: MONITORING RESULTS 129
7.3.1 LER monitoring system 129
7.3.2 Surface roughness monitoring system 130
7.3.3 Edge shape monitoring system 133
7.4 THE EDGE³ MONITORING INTERFACE 135
7.5 CONCLUDING REMARKS 136

CONCLUSIONS AND PERSPECTIVES 139

REFERENCES 143

Web sites 157

ACKNOWLEDGEMENTS 159

List of symbols

General symbols, vectors and matrices

a	=	wavelet scale
A	=	total number of latent variables
$\overline{\mathrm{AAE}}$	=	overall average absolute error
b	=	wavelet location
b_r	=	regression coefficient of the r^{th} latent variable
$\mathbf{B}$	=	matrix of regression coefficients
$c_{i,j}^{T^2}$	=	contribution $c_{i,j}^{T^2}$ of the variable j to the T_i^2 of the i^{th} observation
$c_{i,j}^{\mathbf{t}}$	=	contribution of the variable j to the scores that compose the T_i^2 of the i^{th} observation
$c_{i,j}^{\mathbf{E}}$	=	contribution of the variable j to the square predicting error SPE_i of the i^{th} observation
$\bar{c}_j^{\mathbf{E}}$	=	average contributions of variable j over all the I observations of the reference for the SPE statistics
$\bar{c}_j^{T^2}$	=	average contributions of variable j over all the I observations of the reference for the Hotelling statistics
$c_{j,\mathrm{lim}}^{\mathbf{E}}(\alpha)$	=	the 100(1-α)% confidence intervals for the contributions $c_{i,j}^{\mathbf{E}}$
$c_{j,\mathrm{lim}}^{T^2}(\alpha)$	=	the 100(1-α)% confidence intervals for the contributions $c_{i,j}^{T^2}$
$\mathbf{C}^{\mathbf{E}}$	=	matrix of the contributions to SPE of all the J variables for all the I observations in of $\mathbf{X}$ matrix
$\mathbf{C}^{T^2}$	=	matrix of the contributions to T^2 of all the J variables for all the I observations in of $\mathbf{X}$ matrix
d_m	=	detail of the signal x at the m^{th} wavelet decomposition scale
$\mathbf{D}_m^{\mathrm{h}}$	=	reconstruction of the horizontal detail $\mathbf{T}_m^{\mathrm{h}}$
$\mathbf{D}_m^{\mathrm{v}}$	=	reconstruction of the vertical detail $\mathbf{T}_m^{\mathrm{v}}$
$\mathbf{D}_m^{\mathrm{d}}$	=	reconstruction of the diagonal detail $\mathbf{T}_m^{\mathrm{d}}$

ΔK	=	lag on the process variables in the TP-PLS models
$\Delta K'$	=	length of the moving window in the MATP-PLS models
Δn_{segm}	=	edge segment width (pixel)
Δn_{mw}	=	size of the moving window (pixel)
$e_{i,j}$	=	element of row *i* and column *j* of the residual matrix **E**
$\mathbf{e}$	=	residual of a test sample
$\mathbf{e}_{I+1}$	=	error of reconstruction for the projection of $\mathbf{x}_{I+1}$ onto the latent variable space
$\mathbf{e}_{I+1}^{\mathrm{T}}$	=	transpose of $\mathbf{e}_{I+1}$
$\mathrm{E}(\mathbf{x}_{I+1})$	=	expected value of $\mathbf{x}_{I+1}$
$\mathbf{E}$	=	2D residual matrix of **X**
$\underline{\mathbf{E}}$	=	2D residual matrix of $\underline{\mathbf{X}}$
$F_{J,I-J,\alpha}$	=	upper $100\alpha^{\mathrm{th}}$ percentile of the F-distribution with *I* and *I-J* degree of freedom
$\mathbf{F}$	=	residual matrix of **Y**
h	=	sampling instant of the quality variables in the 3D data matrix of regular shape
h_i	=	sampling instant of the quality variables in the 3D data matrix of irregular shape
h_0	=	parameter of the Jackson-Mudholkar equation
H_i	=	total number of quality samples for the observation *i*
H_0	=	null hypothesis
H_1	=	alternative hypothesis
i	=	observation of the reference dataset
i_0	=	image (pixel)
i_M	=	filtered image at the M^{th} scale of wavelet decomposition (pixel)
I	=	total number of observations in the reference dataset
$\mathbf{I}$	=	identity matrix
$L^2(\Re)$	=	Hilbert space of square integrable functions in $\Re$
j	=	variable of the reference dataset
J	=	total number of variables of the reference dataset
k	=	sample for the process variables in the 3D data matrix of regular shape

k_i	=	sample for the process variables in the 3D data matrix of irregular shape
K_i	=	total number of samples of observation i
m	=	decomposition scale
M	=	selected decomposition level
M_1	=	decomposition level selected for image denoising
$\mathbf{M}_r$	=	matrix of rank 1 of the r^{th} latent variable
$\mathrm{MRPE}_{i,q}$	=	mean relative prediction error for quality variable q in batch i during a single estimation phase
n	=	counter
n_{el}	=	size of the edge length (pixel)
n_{iw}^{E}	=	size of the image width for the edges
n_{iw}^{V}	=	size of the image width for the valleys
n_{levels}	=	number of selected topological levels
n_{sample}	=	total number of quality samples in an estimation phase
n_{tsw}	=	size of the trans-section width (pixel)
N_x	=	length of the signal x
N_A	=	acidity number $\left(\frac{\mathrm{mg_{KOH}}}{\mathrm{g_{resin}}}\right)$
N_{image}	=	number of images of edge segments
$p_{i,j}$	=	element of the i^{th} row and j^{th} column of the matrix $\mathbf{P}$
$\mathbf{p}_j$	=	row vector referring to the j^{th} variable of the loading matrix $\mathbf{P}$
$\mathbf{p}_r$	=	loading of the r^{th} latent variable of $\mathbf{X}$
$\mathbf{p}_r^{\mathrm{T}}$	=	transpose of the loading of the r^{th} latent variable of $\mathbf{X}$
P	=	probability function
$\mathbf{P}$	=	loading matrix of $\mathbf{X}$
$\mathbf{P}^{\mathrm{T}}$	=	transpose of $\mathbf{P}$
$\mathbf{P}_r$	=	matrix of the loadings for all the J variables and all the K samples
$PRESS$	=	prediction error sum of squares
q	=	quality variable
$\mathbf{q}_r$	=	loading of the r^{th} latent variable of $\mathbf{Y}$
Q	=	total number of quality variables

$\mathbf{Q}$	=	loading matrix of $\mathbf{Y}$
$\mathbf{Q}^{\mathrm{T}}$	=	transpose of $\mathbf{Q}$
r	=	generic counter
R	=	rank of the $\mathbf{X}$ matrix
$\Re$	=	space of the real numbers
RMSECV	=	root-mean-square error of cross-validation
$\mathrm{R}(\mathbf{X})$	=	100(1-α) % confidence region of the likely value of a population containing $\mathbf{X}$
s	=	generic counter and spatial coordinate
$s_{\bar{c}_j^{T^2}}$	=	standard deviation of the contributions of variable j over all the I observations of the reference for the Hotelling statistics
$s_{\bar{c}_j^{\mathbf{E}}}$	=	standard deviation of the contributions of variable j over all the I observations of the reference for the *SPE* statistics
s_r	=	semi-axis of the confidence ellipse for the r^{th} latent variable
$\mathbf{s}$	=	coordinate in the domain of pixel space (squared pixel)
$S_{m,n}$	=	approximation coefficient at the m^{th} scale of wavelet decomposition
$S_{m+1,(n_1,n_2)}$	=	approximation coefficients of the multiresolution decomposition of an image
SPE_i	=	squared predicting error for the observation i
SPE_{I+1}	=	squared predicting error for the validation observation $\mathbf{x}_{I+1}$
$SPE_{\lim}(\alpha)$	=	upper limit of SPE_i at a confidence level 100(1-α)%
$\mathbf{S}$	=	estimated value of the covariance matrix $\mathbf{\Sigma}$
$\mathbf{S}_m$	=	approximation matrix of an image at the m^{th} wavelet decomposition scale
t^*	=	maximum time horizon for the prediction of the batch length (h)
$t_{I-1,\frac{\alpha}{2}}$	=	Student t-distribution for *I-1* and $\frac{\alpha}{2}$ degrees of freedom
$t_{\lim}(r,\alpha)$	=	univariate limit at $100(1-\alpha)\%$ confidence level for score $\mathbf{t}_r$
$\mathbf{t}_1$	=	score vector of the first principal component
$\mathbf{t}_i$	=	row vector referring to the i^{th} observation of the score matrix $\mathbf{T}$
$\hat{\mathbf{t}}_{I+1}$	=	projection of the validation observation $\mathbf{x}_{I+1}$ onto the latent subspace
$\mathbf{t}_r$	=	score vector of the r^{th} principal component of $\mathbf{X}$
$\mathbf{t}_r^{\mathrm{T}}$	=	transpose of $\mathbf{t}_r$

T^2	=	Hotelling statistics
TAAE_i	=	time-averaged absolute error of the batch i (h)
$T(a,b)$	=	continuous approximation of the signal x for the wavelet decomposition by means of ψ at location b and scale a
T_i^2	=	value of the Hotelling statistics for the i^{th} observation
T_{I+1}^2	=	value of the Hotelling statistics for a validation observation $\mathbf{x}_{I+1}$
$T_{\mathrm{lim}}^2(A,I,\alpha)$	=	confidence limit of the Hotelling statistics at the 100(1-α) % level of confidence for a system of A latent variables and I samples
$T_{m,n}$	=	detail coefficient at the m^{th} scale of wavelet decomposition
$T_{m,n}^{denoised}$	=	denoised approximation of the wavelet decomposition of i_0 at M_1 scale
$T_{m+1,(n_1,n_2)}^{\mathrm{h}}$	=	horizontal detail coefficient of an image
$T_{m+1,(n_1,n_2)}^{\mathrm{v}}$	=	vertical detail coefficient of an image
$T_{m+1,(n_1,n_2)}^{\mathrm{d}}$	=	diagonal detail coefficient of an image
$\mathbf{T}$	=	score matrix of $\mathbf{X}$
$\mathbf{T}_m^{\mathrm{h}}$	=	horizontal detail matrix of an image at the m^{th} wavelet decomposition scale
$\mathbf{T}_m^{\mathrm{v}}$	=	vertical detail matrix of an image at the m^{th} wavelet decomposition scale
$\mathbf{T}_m^{\mathrm{d}}$	=	diagonal detail matrix of an image at the m^{th} wavelet decomposition scale
$\mathbf{u}_r$	=	score vector of the r^{th} latent variable of $\mathbf{Y}$
$\mathbf{u}_r^{\mathrm{T}}$	=	transpose of $\mathbf{u}_r$
$\mathbf{U}$	=	score matrix of $\mathbf{Y}$
VIP_j	=	importance of the variable j in the projection methods
$\mathbf{w}_r$	=	weight of the r^{th} latent variable
$\mathbf{w}_r^{\mathrm{T}}$	=	transpose of $\mathbf{w}_r$
$\mathbf{W}$	=	matrix of the weights
x	=	generic signal
$x_{i,j}$	=	element of row i and column j of the $\mathbf{X}$ matrix
$x_{i,j,k}$	=	element of the $\underline{\mathbf{X}}$ matrix
$\bar{x}_{i,j,k}$	=	moving average of the variable j on batch i in the k-th time instant, element of the $\bar{\mathbf{X}}_i$ matrix

x_m	=	approximation of the signal x at the m^{th} wavelet decomposition scale
$\mathbf{x}_i$	=	row vector of the i^{th} observation of the $\mathbf{X}$ matrix
$\mathbf{x}_{I+1}$	=	vector of a validation observation
$\hat{\mathbf{x}}_{I+1}$	=	projection of $\mathbf{x}_{I+1}$ onto a latent space
$\mathbf{x}_{i,j}$	=	j^{th} variable time profile in batch i in form of column array of the $\mathbf{X}_i$ matrix
$\mathbf{x}_j$	=	j^{th} variable column vector of the $\mathbf{X}$ matrix
$\bar{\mathbf{x}}_j$	=	average value of the j^{th} variable (column) of $\mathbf{X}$
$\mathbf{x}_{i,j}^{-\Delta K}$	=	vector of the j^{th} variable time trajectory for the i^{th} batch lagged of $-\Delta K$ time instants
$\mathbf{X}$	=	reference data matrix of the process
$\overline{\mathbf{X}}$	=	array of the mean values of the variables of $\mathbf{X}$
$\hat{\mathbf{X}}$	=	projection of the $\mathbf{X}$ matrix onto the space of the latent variables
$\underline{\mathbf{X}}$	=	3Dreference data matrix of the process variables
$\mathbf{X}_0$	=	2D data matrix at the zero decomposition scale
$\mathbf{X}^{\text{BWU}}$	=	2D data matrix derived form $\underline{\mathbf{X}}$ form batch-wise unfolding
$\overline{\mathbf{X}}^{\text{BWU}}$	=	input matrix of the moving averages for the MATP-PLS model
$\mathbf{X}^{\text{D}}$	=	matrix of lagged variables
$\mathbf{X}_i$	=	i^{th} horizontal slice of $\underline{\mathbf{X}}$, i.e. matrix of the trajectories of all the J variables in all the K_i samples in time or space for the observation i
$\overline{\mathbf{X}}_i$	=	matrix of the moving average data of the i^{th} batch
$\mathbf{X}_i^{\text{D}}$	=	matrix of lagged variables for the i^{th} batch
$\mathbf{X}_j$	=	j^{th} vertical slice of $\underline{\mathbf{X}}$, i.e. matrix of the time/space evolution of the variable j for all the samples K and all the observations I
$\mathbf{X}_k$	=	k^{th} vertical slice of $\underline{\mathbf{X}}$, i.e. matrix of the time/space sample k for all the J variables and all the I observations
$\mathbf{X}^{\text{L}}$	=	augmented matrix with lagged variables for the LTP-PLS model
$\mathbf{X}_M$	=	2D data matrix at the M^{th} decomposition scale
$\mathbf{X}^{\text{T}}$	=	transpose of $\mathbf{X}$
$\mathbf{X}^{\text{VWU}}$	=	bi-dimensional data matrix derived by variable-wise unfolding $\underline{\mathbf{X}}$
$y_{i,q,h}$	=	element of the $\underline{\mathbf{Y}}$ matrix
$\hat{y}_{i,q,h}$	=	estimated value of $y_{i,q,h}$

$\hat{\mathbf{y}}_{I+1}$	=	estimated value of a quality index for the $(I+1)^{\text{th}}$ observation
$\mathbf{Y}$	=	matrix of the quality variables
$\underline{\mathbf{Y}}$	=	three dimensional reference matrix of the quality variables
$\mathbf{Y}_i$	=	i^{th} horizontal slice of $\underline{\mathbf{Y}}$, i.e. matrix of the trajectories of all the Q quality variables in all the H_i samples in time or space for the observation i
z_α	=	normal standard deviate corresponding to the upper 100(1-α)% percentile

Greek symbols

α	=	percentile of the confidence limits
$\delta_{r,s}$	=	Kronecker delta
ε_i	=	instantaneous error of estimation of stage length in batch i
θ	=	generic parameter
θ_n	=	parameter of the Jackson-Mudholkar equation
Θ	=	space of all the possible parameters θ
λ	=	forgetting factor
$\mathbf{\Lambda}$	=	diagonal matrix of the eigenvalues λ_r
λ_r	=	eigenvalue of the r^{th} latent variable
μ	=	viscosity
$\boldsymbol{\mu}_0$	=	vector of the expected values of the J variables of the matrix $\mathbf{X}$
$\varphi_{m,n}$	=	discretized father wavelet
$\varphi(\mathbf{s})$	=	bidimensional wavelet function
$\mathbf{\Sigma}$	=	covariance matrix
τ	=	batch length (h)
τ_i	=	actual length of the stage in the same batch
τ^*	=	number of samples corresponding to the time horizon t^*
$\hat{\tau}_i(t)$	=	prediction at time t of the stage length in batch i
ψ	=	mother wavelet function

$\psi_{a,b}$	=	mother wavelet function for a dilation parameter *a* and a location parameter *b*
$\psi^{*}_{a,b}$	=	complex conjugate of a "mother" wavelet function $\psi_{a,b}$
$\psi_{m,n}$	=	discretization of the mother wavelet function
$\psi^{h}(\mathbf{s})$	=	bidimensional horizontal wavelet
$\psi^{v}(\mathbf{s})$	=	bidimensional vertical wavelet
$\psi^{d}(\mathbf{s})$		bidimensional diagonal wavelet
$\chi^{2}_{v,\alpha}$	=	χ^2-distribution with v and α degrees of freedom

Acronyms

2D	=	bi-dimesional
3D	=	three-dimensional
AR	=	autoregressive
ARMA	=	auto-regressive moving average
BWU	=	batch-wise unfolding
CA1	=	carboxylic acid 1
CA2	=	carboxylic acid 2
CD	=	critical dimension
CD-SEM	=	tool for the measurement of the CD through a SEM
D1	=	diol 1
D2	=	diol 2
DA1	=	dioic acid
DPCA	=	dynamic PCA
DPLS	=	dynamic PLS
IC	=	integrated circuit
IID	=	independent identically distributed
LAN	=	local area network
LER	=	line edge roughness
LTP-PLS	=	lagged three-phase PLS

LV	=	latent variable
LV1	=	first latent variable
LV2	=	second latent variable
MATP-PLS	=	moving-average three-phase PLS
MIA	=	multivariate image analysis
MPCA	=	multiway PCA
MPLS	=	multiway PLS
NIPALS	=	non-iterative partial least squares algorithm
NOC	=	normal operating conditions
OLE	=	object linking and embedding
OPC	=	OLE for process control
PC	=	principal component
PC1	=	first principal component
PC2	=	second principal component
PCA	=	principal component analysis
PLC	=	programmable logic controllers
PLS	=	partial least squares method (projection on latent structures)
PV	=	process value
P&ID	=	pipelines and instrumentation diagram
RGB	=	red, green, blue
RTU	=	remote terminal units
SCADA	=	supervisory control and data acquisition
SEM	=	scanning electron microscopy (or microscope)
SIMPLS	=	straightforward implementation of modified PLS
SP	=	setpoint
SPC	=	statistical process control
SQC	=	statistical quality control
SQL	=	structured query language
SWA	=	side wall angle
TP-PLS	=	three-phase PLS method
UV	=	ultra violet
VIP	=	variable importance in the projection methods

VO	=	valve opening
VWU	=	variable-wise unfolding