

University of Alberta

Control and monitoring of sheet and film forming processes

by

Jaganath Ramarathnam

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Process Control

Department of Chemical and Materials Engineering

©Jaganath Ramarathnam

Fall 2009

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms. The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Examining Committee

Sirish L. Shah, Chemical and Materials Engineering

Biao Huang, Chemical and Materials Engineering

Tongwen Chen, Electrical and Computer Engineering

Vinay Prasad, Chemical and Materials Engineering

To

*my mom, dad, brother,
sister-in-law and
the love of my life,
Ashwini Srinivas*

Abstract

Sheet and film forming processes refer to a set of processes that have a 2-dimensional sheet or film as their output. They are typically characterized by a scanning sensor which moves between the edges of the sheet in a periodic manner. This work is concerned with control and monitoring of such processes. There are three main contributions in this study. The first and foremost contribution is the reformulation of the Linear Quadratic Gaussian (LQG) objective function to give one the ability to tradeoff between control over either dimension of the sheet. The second contribution is a method to derive the LQG trade-off surface between output variability in either dimension and input variance. The third contribution is a set of data driven techniques for performance assessment of these processes. Simulation results using the model used by Bergh and MacGregor (1987) are provided to support the proposed methods.

Contents

Abstract	iv
1 Introduction	1
1.1 Overview of sheet and film forming processes	1
1.2 Motivation and contributions	3
1.3 Thesis outline	6
2 Literary review of sheet and film forming processes	8
2.1 Paper making process	8
2.2 Plastics extrusion	11
2.3 Metal rolling	12
2.4 Coating	13
2.5 Scanning versus stationary sensors	13
2.6 Modelling	14
2.6.1 Mapping	18
2.6.2 Constraints	19
2.6.3 Identification	21
2.7 Control	22
2.7.1 Separate MD and CD control	23
2.7.2 Coupled MD and CD control	30
2.8 Performance Assessment	31
3 Trade-off in coupled control of sheet and film forming processes	33
3.1 Introduction	33
3.2 Decoupling of MD and CD control objectives	35
3.2.1 When Q is a multiple of the identity matrix	35
3.2.2 When Q is a diagonal matrix	36
3.2.3 When Q is any symmetric, positive definite matrix	37

3.3	Decoupled objective function and its implementation	37
3.4	Sample model	40
3.5	Analysis of interaction parameters	44
3.6	Results	45
3.7	Conclusion	49
4	LQG benchmark for sheet and film forming processes	52
4.1	Introduction	52
4.2	LQG trade-off curve for time-varying MIMO processes	54
4.2.1	Time-varying state space model of the closed loop system	54
4.2.2	Time-varying impulse response coefficients	57
4.2.3	H_2 norm of the closed-loop system	58
4.3	Performance evaluation using full profile information	59
4.3.1	Evaluating the performance by location	61
4.3.2	Evaluating the performance by direction (machine/cross-direction)	62
4.4	Coupled control of sheet and film forming processes	62
4.4.1	Model	63
4.4.2	State estimation and control	64
4.5	Simulation results	66
4.5.1	Comparison between PID and LQG controllers as an example	73
4.6	Conclusion	78
5	Data driven techniques for performance assessment of sheet and film forming processes	80
5.1	Introduction	80
5.2	ACF based performance index	81
5.3	ACF of the scanning sensor output	82
5.4	ACF of estimated MD and CD variations	86
5.4.1	Comparison of estimation methods	86
5.4.2	Assessing MD and CD control	90
5.5	Conclusion	90
6	Concluding Remarks	93
6.1	Suggestions for Future Research	95
	Bibliography	97

List of Tables

4.1	Controller settings for the PID controllers	76
4.2	Settings for the LQG controllers	76

List of Figures

1.1	Typical sheet or film forming process	2
2.1	Industrial paper making machine (adapted from Flickr (2009))	9
2.2	Schematic of a paper making process	9
2.3	Schematic of a plastic blown film extrusion process	12
2.4	Schematic of a Sendzimir mill	13
2.5	Shrinking of paper along the CD	19
3.1	Open-loop simulation showing the output of the new disturbance model	41
3.2	Open (dash-dot line) and closed loop (starred line) response under LQG control for a 5-input, 4-output system.	43
3.3	Control actions for closed loop response shown in Fig. 3.2	43
3.4	Output profiles at various time points for closed loop response in Fig. 3.2 (— shows average profile).	44
3.5	Open (dash-dot line) and closed loop (starred line) response under LQG control with heavy weighting in the machine direction	46
3.6	Input for closed loop response under LQG control with heavy weighting in the machine direction	47
3.7	Output profile at various time points for closed loop response under LQG control with heavy weighting in the machine direction (— line shows average profile).	47
3.8	Open (dash-dot line) and closed loop (starred line) response under LQG control with heavy weighting in the cross direction	48
3.9	Input for closed loop response under LQG control with heavy weighting in the cross direction	49
3.10	Output profile for closed loop response under LQG control with heavy weighting in the cross direction (— shows average profile).	50

3.11	Comparison of zero-mean cross-directional profiles for different MD and CD weightings. The following offsets are used: +1.5 for $t = 100$, +0.5 for $t = 200$, -0.5 for $t = 300$, -1 for $t = 400$	51
4.1	LQG tradeoff curve	53
4.2	Measurements times for section of sheet used for performance assessment	60
4.3	Open-loop simulation showing the output of the new disturbance model	64
4.4	Trade-off curve by location with (a) heavy weighting in the CD, (b) equal weighting, and (c) heavy weighting in the MD, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.	70
4.5	CD variance by location with (a) heavy weighting in the CD, (b) equal weighting, and (c) heavy weighting in the MD, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.	71
4.6	(a) MD variance versus input variance trade-off at $t = 207$ sampling instants, and (b) CD variance versus input variance trade-off at $t = 207$ sampling instants. Simulation results represent variance of output over 100 control simulations.	72
4.7	MD and CD trade-off for different input variances, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.	72
4.8	CD, MD and input trade-off surface, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.	73
4.9	Comparison between PID controller performance, LQG controller performance, and optimum tradeoff curve	77
4.10	Effect of p value on the calculated variances for three independent simulation of a PID controller: (a) CD variance (b) MD variance	78
5.1	(a) ACF of a white noise sequence $\eta_{acf} = 1$ (b) ACF of closed-loop output for a suboptimal controller $\eta_{acf} = 0.46$	82

5.2	(a) ACF of closed-loop scanner output when control is weighted heavily in the machine-direction ($\eta_{acf} = 0.029639$, output variance = 0.5929) (b) ACF of closed-loop scanner output when control is weighted heavily in the cross-direction ($\eta_{acf} = 0.10176$, output variance = 0.31225) (c) ACF of closed-loop scanner output when control is weighted equally in both directions ($\eta_{acf} = 0.054525$, output variance = 0.15809) (d) ACF of open-loop scanner output obtained when the same disturbance used to generate (a), (b) and (c) is applied ($\eta_{acf} = 0.033094$, output variance = 1.8813)	84
5.3	(a) ACF of open-loop scanner output obtained when the spatial correlation coefficient $\rho = 0.4$ ($\eta_{acf} = 0.065577$, output variance = 1.4737), and (b) ACF of open-loop scanner output obtained when the spatial correlation coefficient $\rho = 0.1$ ($\eta_{acf} = 0.091782$, output variance = 1.6032)	85
5.4	Comparison of MD estimation algorithms with actual MD variation data. Estimation algorithms compared: (a) SSM, (b) EXPO and Kalman filter	87
5.5	Comparison of SSM CD estimation algorithm with actual CD variation data.	88
5.6	Comparison of EXPO and Kalman filter CD estimation algorithms with actual CD variation data. In EXPO, the historical CD profile is weighted by 0.7 and the current CD output is weighted by 0.3. Location 1 and 2 are shown in (a) while location 3 and 4 are shown in (b).	89
5.7	ACF of (a) closed-loop estimated MD output ($\eta_{acf} = 0.044439$) and (b) closed-loop estimated CD output ($\eta_{acf, \text{location 1}} = 0.042018$, $\eta_{acf, \text{location 2}} = 0.031967$, $\eta_{acf, \text{location 3}} = 0.035729$, $\eta_{acf, \text{location 4}} = 0.060291$)	91
5.8	ACF of (a) closed-loop estimated MD output ($\eta_{acf} = 0.05325$) and (b) closed-loop estimated CD output ($\eta_{acf, \text{location 1}} = 0.069308$, $\eta_{acf, \text{location 2}} = 0.040495$, $\eta_{acf, \text{location 3}} = 0.053446$, $\eta_{acf, \text{location 4}} = 0.084458$)	91
5.9	ACF of (a) closed-loop estimated MD output ($\eta_{acf} = 0.048107$) and (b) closed-loop estimated CD output ($\eta_{acf, \text{location 1}} = 0.041463$, $\eta_{acf, \text{location 2}} = 0.031762$, $\eta_{acf, \text{location 3}} = 0.03582$, $\eta_{acf, \text{location 4}} = 0.058847$)	92

List of Symbols, Nomenclature, or Abbreviations

\bar{y}	Nominal value of $y(x, t)$
η_{acf}	Performance index based on ACF of process output
\hat{U}_k	Vector of inputs at time k
\hat{Y}_k	Vector of estimated or predicted deviation of sheet properties at time k
κ	State space matrices of open-loop system
Λ	Diagonal matrix containing the eigenvalue of Q , obtained from the orthogonal decomposition of Q
∇	difference operator $(1 - q^{-1})$
\bar{A}	State space matrices of closed-loop time varying system
\bar{C}	State space matrices of closed-loop time varying system
\bar{G}	State space matrices of closed-loop time varying system
\bar{H}	State space matrices of closed-loop time varying system
\bar{K}	State space matrices of closed-loop time varying system
\bar{p}_k	Average of elements of P_k
\bar{y}_k^w	Weighted average of $y_{i,k}$
\bar{y}_k	Average of elements of Y_k
ϕ	AR polynomial of disturbance model

ρ	spatial correlation of disturbances ($0 < \rho < 1$)
Σ_a	Covariance matrix of a_t
θ	MA polynomial of disturbance model
A/A_t	State space matrices of open-loop system
a_1, a_2	mapping coefficients between c_j and x_j
a_t	vector of white noise as input to disturbance model
ACF	AutoCorrelation Function
$ARIMA$	Auto-Regressive Integrated Moving Average
B	interactor matrix
B/B_t	State space matrices of open-loop system
B_{max}	Largest acceptable ‘zigzag’ of actuators
C/C_t	State space matrices of open-loop system
c_j	center of downstream response of j^{th} actuator
C_{max}	Maximum rate of change of actuator set point between two consecutive sampling times
C_{obj}	Matrix that transforms original state-space model to state-space model with separate MD and CD terms for implementation in proposed objective function
CD	Cross Direction
CPA	Controller Performance Assessment
D/D_t	State space matrices of open-loop system
D_{max}	Largest acceptable difference between adjacent actuators
E_t	Kalman filter error between estimated and actual state
e_t	Measurement noise in the scanner output
g	parameters of the interactor matrix

G/G_t	State space matrices of open-loop system
H/H_t	State space matrices of open-loop system
I_m	Identity matrix of dimension $m \times m$
IMC	Internal Model Control
J	Linear Quadratic Gaussian (LQG) objective function
J	Objective function in coupled MD and CD control
K_D	Derivative gain of PID controller
K_I	Integral gain of PID controller
K_p	Proportional gain of PID controller
LHS	Left Hand Side
LQG	Linear Quadratic Gaussian
LTV	Linear Time Varying
m	number of measurement locations
M/M_t	State space matrices of open-loop system
MD	Machine Direction
$MIMO$	Multiple Input Multiple Output
n	number of CD actuator
$N(t)$	vector of disturbance at time t
N, M, Ψ	Matrices obtained from the SVD of the interactor matrix
n_{lags}	User defined parameter describing the number of lags of the ACF that are included in the calculation of η_{acf}
$P(t)$	Covariance matrix of states estimated by the Kalman filter
P_k	Transformed and scaled version of Y_k
$p_{i,k}$	Elements of P_k

PID	Proportional Integral Derivative
Q	Penalty matrix on \hat{Y}_k
q^{-1}	backshift operator
$q_{gmv}(z^{-1})$	input penalty transfer function
R	Penalty matrix on \hat{U}_k
r_u	Penalty on control action of $u_1:u_n$
r_v	Penalty on control move of v_k
RHS	Right Hand Side
SSM	Simple Scan Method
$U(t)$	vector of CD actuator moves at time t
$u_1 \dots u_n$	Set point of CD actuator
$u_{i,t,max}$	Maximum setpoint allowed for i^{th} actuator at time t
$u_{i,t,min}$	Minimum setpoint allowed for i^{th} actuator at time t
$u_{i,t}$	Setpoint for i th actuator at time t
$v(t)$	Mapped error profile
W_t	State noise in closed loop time varying system
W_{MD}	Weighting for CD cost in objective function
W_{MD}	Weighting for MD cost in objective function
x_j	upsteam location of j^{th} actuator
X_k	States of state-space model obtained by multiplying original state-space model with $C_o b_j$
$Y(t)$	vector of CD properties at time t
$y(x, t)$	Set point of MD actuator
$y(x, t)$	Set point of MD actuator

$y(x, t)$	Variation in y at location x and time t
y_i	Elements of Y_k
$y_R(x, t)$	Residual variations in $y(x, t)$ at time t
$y_{CD}(t)$	CD variations in $y(x, t)$ at time t
$y_{MD}(t)$	MD variations in $y(x, t)$ at time t
Z	Eigenvector matrix of Q , obtain from the orthogonal decomposition of Q

Chapter 1

Introduction

The term ‘sheet and film forming processes’ refers to a set of processes that have a 2-dimensional sheet or film as their output, and are typically found in the pulp and paper, plastics, metal rolling and coating industries (Bergh and MacGregor (1987)). Examples of products produced using sheet and film forming processes are metal sheets, tissue paper, adhesive stickers, and plastic bags. The economics governing these industries necessitates good control of these processes. The total capitalization of industries that rely on coating technology has been estimated to be over \$1 trillion (VanAntwerp et al. (2007)). Paper manufacture is the mainstay of the multi-billion dollar pulp and paper industry (Stewart (2000)). Goods manufactured through polymer-film extrusion range from windshield safety glass to large plastic bags. A significant fraction of the world’s metal is produced in a sheet form (VanAntwerp et al. (2000)).

Although the physical process in each of these industries is very different, the control objective for all sheet and film forming processes is similar. The objective is to maintain certain physical properties of the sheet or film uniform over both dimensions of the product during manufacture. In the next section, an overview of the common features of sheet and film forming processes is provided.

1.1 Overview of sheet and film forming processes

Fig. 1.1 depicts a highly simplified version of an industrial paper making process. This figure is used to illustrate the principal features of sheet and film forming processes. Feed flows into the headbox through an inlet pipe, and gets distributed onto a continuous wire mesh to form the sheet. The direction of flow of the

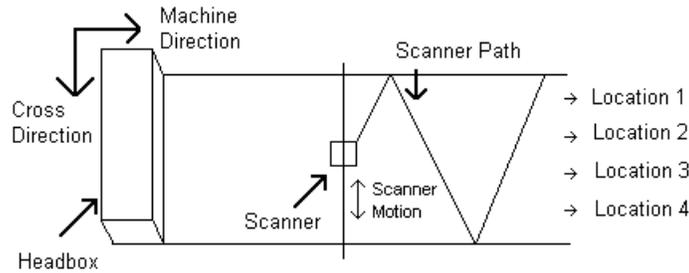


Figure 1.1: Typical sheet or film forming process

sheet in the machine with time is known as the machine direction (MD). The cross direction (CD) is perpendicular to the machine direction and corresponds to the spatial dimension of the process. The flowrate of feed to the headbox is manipulated using a valve and this valve is commonly used for machine-directional control. Henceforth, this valve is referred to as the MD actuator. The distribution of feed onto the continuous wire mesh in the headbox occurs through a set of actuators placed along the cross direction. These actuators are used for controlling the spatial profile, or CD profile, of the sheet, and are henceforth referred to as cross-directional (CD) actuators. The spatial region influenced by an actuator typically overlaps the regions influenced by neighbouring actuators, creating a coupling between the actuators and spatial outputs. In this paper making process, the manipulation of the CD actuators results in a redistribution of the net flow across the entire sheet, and therefore does not affect the MD variations of the sheet. However, in many other sheet or film forming processes, this is not the case, and CD actuators have an effect on the MD variations of the sheet.

The true complexity of a sheet or film forming process occurs due to the ‘scanning’ nature of the sensor typically found in the industry. This sensor, used to measure the physical property of interest, is mounted on a rack that moves between the edges of the sheet as shown in Fig. 1.1. The direction of the travel of the sheet and the motion of the scanner result in the scanner traversing a zig-zag path on the sheet. This makes the process periodic time-varying in nature, and causes difficulties in the separation of MD and CD variations of the process, for separate MD and CD control. Another complexity created by the measurement process is that the number of points measured by the scanning scanner over a full scan is typically more than the number of CD actuators. Therefore, the measured profile may be “transformed” to match the number of actuators being controlled, depending on the control scheme. The profile of the sheet at the location of the

sensor is henceforth referred to as the CD profile.

Sheet and film forming processes present a challenging control problem for a variety of reasons, including their two dimensional nature, the large numbers of inputs and outputs, the nature of the measurement process, the need for accurate mapping between actuator response and spatial location, coupling between actuator response and spatial locations, and constraints specific to the nature of the actuators. While traditional control approaches have been significantly successful in reducing process variations, advances in computer technology provide a platform for the implementation of the next generation of control technology to realize greater benefits (Heaven et al. (1994)). Some of the benefits of improving control in these processes include meeting higher quality specifications (Heaven et al. (1994)), reducing material consumption, increasing productivity from existing equipment and elimination of product rejects (VanAntwerp et al. (2007)).

1.2 Motivation and contributions

It is common in both academia and the industry to regard the two-dimensional control problem as two separate control problems, one along the direction in which the product is produced (Machine Direction, or MD), and another across the direction in which the product is produced (Cross Direction or CD). While this is acceptable, coupling the two control problems into a single problem of coupled MD and CD control can result in improved control (Bergh and MacGregor (1987)). Coupled MD and CD control is likely to perform better than separate MD and CD control because CD actuators can be used partially to control MD variations and MD actuators can be used partially to control CD variations. However, this has been recommended only for processes with fast changes in the cross direction, and the objective, which is to minimize output variations at all spatial locations during every time-step, offers little flexibility for tuning the control in favor of either direction. While separated MD and CD control has been explored extensively, coupled MD and CD control is relatively unexplored. Many questions regarding coupled MD and CD control are unanswered, such as:

- what is the link between MD and CD control and minimization of process variances?
- when MD and CD control are coupled, how can one tradeoff between MD

and CD control?

- Is it possible to employ coupled MD and CD control for processes with slow CD changes and still obtain good CD profiles, as this is the requirement for most processes?

A brief description of the physical sheet-forming process is outlined in Chapter 2. This is followed by the third chapter, where answers to the above questions are provided through the introduction of a new framework for analysing coupled MD and CD control. A simple relationship is shown to exist between the MD and CD control objectives and the minimization of process variance. This relationship suggests a tradeoff between the two control objectives in coupled control. Therefore, an LQG objective function with the ability to tradeoff between the two control objectives has been proposed. The implementation of this objective function in standard LQG framework is performed by modifying the standard coupled control model using a simple matrix. As a result of the modification, interaction parameters in the model are brought out clearly. The possible benefits of implementing coupled control can be judged using the gains of these interaction parameters, and separated MD and CD control can be interpreted as a special case where these interaction parameters are assumed to be zero. As most types of CD actuators are likely to have at least some effect on the MD control of the process, implementating the proposed coupled control scheme is likely to improve the performance of most sheet and film forming processes without compromising on the performance of the relatively more important direction.

Natural phenomena such as wear and tear of process equipment typically degrade the performance of the controller. Therefore, it is necessary to benchmark the performance of a controller for a given process in order to ensure that the benefits of improved control continue to accrue over a period of time. The LQG tradeoff curve is one method for benchmarking performance, and is the only method that takes control effort into consideration (Huang and Shah (1999)). In the fourth chapter, the LQG tradeoff curve is used to benchmark the performance of sheet and film forming processes, which are inherently Linear Time Varying (LTV) Multiple Input Multiple Output (MIMO) processes due to the periodic nature of the measurement process. Simulation results are used to verify the proposed method. The assessment of controller performance using closed loop data is discussed assuming full profile information is available at every time step, as scanning sensor

measurements provide insufficient information for fully establishing MD and CD variations. The methods developed can also provide information on the tradeoff between MD and CD control, MD control and input variance, and CD control and input variance in addition to the tradeoff between output variance and input variance. As an illustration of the use of the proposed method, the performance of PID controllers performing separated MD and CD control and the performance of LQG controllers performing coupled MD and CD control are analyzed and compared, and the benefit of coupled MD and CD control is demonstrated.

A recent development in CD control is the availability of full-scan sensors which measure the entire output profile at every time-step (Tyler and Morari (1995)). An important question that arises is the benefit of implementing a full-scan sensor against using a scanning sensor. In Tyler and Morari (1995), a method has been provided for analysing the effect of various scanning patterns on the closed loop scanning sensor variance. The results of Chapter 4 provide an alternative method for analysing the effect of various scanning patterns on the closed loop scanning sensor variance. The principal difference between these methods is that the method in Chapter 4 does not require the solution of any Riccati equations after the periodic Kalman filter gains are established and is significantly simpler to use, but the results are subject to truncation error created by truncating certain infinite series, while the method in Tyler and Morari (1995) requires the solution of additional Riccati equations, but avoids truncation error. An additional difference between the two methods is that the method in Tyler and Morari (1995) assumes that the periodicity of the process is limited to the measurement equation, while the method in Chapter 4 can be employed for processes with an arbitrarily time varying state and measurement equation.

While model based approaches for performance assessment are very useful, a model may not be available in many cases. Therefore, data driven techniques for performance assessment present a significant advantage in terms of their ease of implementation, as they do not need an *a priori* model. Such techniques for the performance assessment of sheet and film forming processes are discussed in the fifth chapter. Here, the autocorrelation function (ACF) of the scanning sensor is shown to be a useful tool for assessing the relative performance of CD and MD control without having to separate MD and CD variations. Such a tool has at least two uses: identifying the priority with which controllers need to be further examined, and determining the extent of spatial correlation in the

disturbances. Various methods for separating MD and CD variations are then compared in order to determine their suitability towards performance assessment of MD and CD controllers. The results of the comparison indicate that a Kalman filter is required to sufficiently capture information on MD and CD variations as is required for performance assessment. An index based on the ACF of the estimated MD and CD outputs is then proposed to judge the performance of the controllers. This index compares the closed-loop output with white noise, as an output of white noise indicates no further scope for control. The use of this index is illustrated with simulation examples.

1.3 Thesis outline

In the next chapter, a short description of the physical processes in industry is provided. This is followed by a discussion on the current state-of-the-art in modelling, control and performance assessment of these processes.

In the third chapter, a link is shown to exist between process variability and control over the temporal dimension (or Machine Direction, MD) and spatial dimension (or Cross-machine Directional, CD) of the process. Hence, a modification is proposed to the coupled control algorithm for sheet and film forming processes that provides the user with the ability to tradeoff between the temporal and spatial control objectives.

In the fourth chapter, a method for analysing the performance of sheet and film forming processes is proposed by extending the LQG tradeoff curve (proposed by Huang and Shah (1999)) to LTV MIMO processes. The coupled control problem discussed in the previous chapter is used to test the method, and simulation results are shown to agree with theoretical results from the method. In addition, a comparison between PID controllers performing decoupled control and LQG controller performing coupled control is provided to illustrate the use of the benchmark.

In the fifth chapter, data driven techniques for performance assessment of sheet and film forming processes are discussed. Here, the ACF of the scanning sensor is shown to be a useful tool for comparing the relative performance of CD and MD control without having to separate MD and CD variations. Following this, the various methods for separating MD and CD variations are compared in order

to determine their suitability towards assessing the performance of MD and CD controllers. An index, based on the ACF of the estimated MD and CD outputs, is then proposed to judge the performance of the individual controllers, and the use of this index is illustrated with simulation examples.

In the sixth chapter, conclusions of this work are provided and possible directions in which further research can be carried out are suggested.

Chapter 2

Literary review of sheet and film forming processes

2.1 Paper making process

Fig. 2.1 shows an industrial paper making machine. Most of today's paper making machines have a few hundred actuators and several hundred sensor lanes (VanAntwerp et al. (2000)). The width of the paper produced is typically around 7.5m and the speed at which the paper goes through the machine is in excess of 50 mph (VanAntwerp et al. (2000)). A schematic of the different sections of a paper making process is shown in Fig 2.2. A dilute suspension of fibers (in the 0.3 – 0.6% fibers range, the rest being water) is pumped into the headbox and onto a continuous wire frame (Fan (2003)). Three mechanisms are used to remove the water: gravity and suction devices at the wet end, a set of counter-rotating rollers in the press section and steam driers in the drying section (Fan (2003)). Caliper reduction and coating operations are performed in the post-drying section to improve paper quality (Fan (2003)), after which the paper is wound into a reel. Paper properties that are controlled in paper-making machines include basis weight, moisture, caliper and coat weight. A good review of the paper making process and advanced control applications of the same can be found in Smook (1992) and Dumont (1986), respectively.



Figure 2.1: Industrial paper making machine (adapted from Flickr (2009))

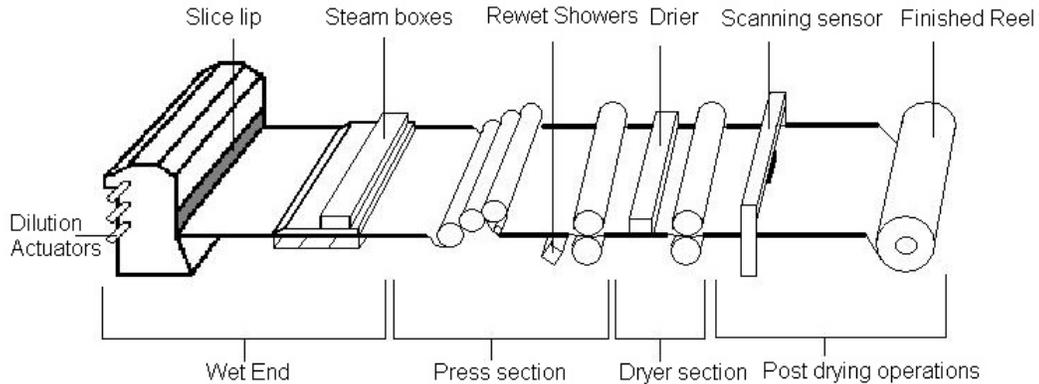


Figure 2.2: Schematic of a paper making process

Basis weight control

Measured in grams per square meter - gsm, basis weight values vary between 35 and 450 gsm (Maenpaa (2006)), depending on the grade of the paper. MD basis weight control is performed by manipulating the overall flowrate of pulp into the paper machine and the speed of the machine itself (Fan (2003)). Two types of actuators, slice-lip and dilution actuators, are commonly used for CD basis weight control. Slice-lip actuators work by locally deflecting the upper lip of a slice opening which controls the flow of fiber suspension out of the headbox (Maenpaa (2006)). Owing to the nature of the actuator, large differences between adjacent actuator levels are undesirable as they result in excessive strain and wear on the slice lip (Thake (1997)). Dilution actuators are a recent development in the design of paper machines. In this method, an array of actuators distributed across the headbox locally alter the concentration of the fibers by diluting it with

a flow of low consistency water (Vyse et al. (1996)). Dilution actuators provide a narrower spatial response and a much better bandwidth for control of basis weight profiles in comparison to slice lip actuators, but require more accurate mapping information (Maenpaa (2006)).

Moisture control

Moisture content of paper has an important influence on the paper strength (Maenpaa (2006)). Measured as a percentage of sheet weight, it is typically in the 5-9% by weight range (Fan (2003)). Some of the CD actuators employed for moisture control are: zone actuators, steam boxes, rewet shower actuators, and infrared heaters (Maenpaa (2006)). Zone actuators vary the local forces applied on the paper to improve the dewatering properties of the paper. Steam boxes alter the local temperature to vary the viscosity of water, which alters the dewatering properties. Rewet shower actuators use spray nozzles to rewet the paper locally to adjust the moisture profile. Infrared heaters alter the moisture content through infrared heating (Holik (2006)). In the MD direction, moisture control is performed by manipulating the overall steam flow into the dryer section (Fan (2003)).

Caliper control

The caliper of paper is typically in the order of μm . Many large cast-iron rolls are used to reduce the caliper of paper in an operation known as calendaring (Fan (2003)). A smoother sheet results in an improved paper quality, and a more uniform caliper profile improves the winding process and reduces sheet breaks in the printing press. Three technologies are used to control paper caliper: induction heating systems, confined air showers and zone controlled CD rolls (Maenpaa (2006)). Induction heating systems and confined air showers increase the diameter of the roll by heating the rolls externally, and thereby increase the pressure applied on the paper, while zone controlled CD rolls vary the pressure by mechanically deflecting the roll shell.

Coat weight control

Coating is primarily done to create a dense and homogenous surface that allows a high image definition (Maenpaa (2006)), and is measured in terms of optical reflection of the sheet (Fan (2003)). In this process, paper passes through two rolls known as application roll and backing roll. The application roll is continuously dipped in a pan of coating, and therefore applies excess coating on the paper that passes through it. The backing roll is used to support the paper as it passes through the machine. A flexible bent blade, attached to the pan containing the coating, is engaged at a low angle (10° to 15°) to the paper to remove the excess coating, which is deposited back into the pan. Coat weight control is done by controlling an array of motor-driven spindles, which locally affect compression force of the blade on the paper (Maenpaa (2006)). Changes in blade loading change the coat weight in a non-linear way (Ismail and Dumont (2003)). The coating process is non-stationary, and gain drifts and sign reversals due to wear of the blade are typical.

2.2 Plastics extrusion

Polymer film production operates by forming a tube of film which is inflated with compressed air to form a bubble. The resulting film is 60 times less thick than the original film (Wellstead et al. (2000)). Fig. 2.3 shows a typical blown film production line. Feed enters the main extruder through a feed hopper and a feed port, which streamlines the flow of raw material. The action of the screw moves the feed forward where it melts due to electrical heating and friction. Frictional heating may be significant enough to avoid the use of electric heating after start-up (VanAntwerp et al. (2000)). The film is then extruded and cooled using air. Additional air or liquid coolers are sometimes used to control the structure of the polymer. Imperfections in the die surface, disturbances in the air flow surrounding the film, and changes in polymer physical properties between runs cause variations in the thickness of the film. Thickness variations are controlled through CD actuators that vary the local width of the die gap, such as choke bars and flexible lips. Alternatively, the temperature of the polymer melt at the die surface can be controlled by local heating/cooling channels. Thickness may be measured using a sensor that revolves around the film tube, creating a spiral pattern, or using scanning gauges that measure the thickness after the product is flattened. For

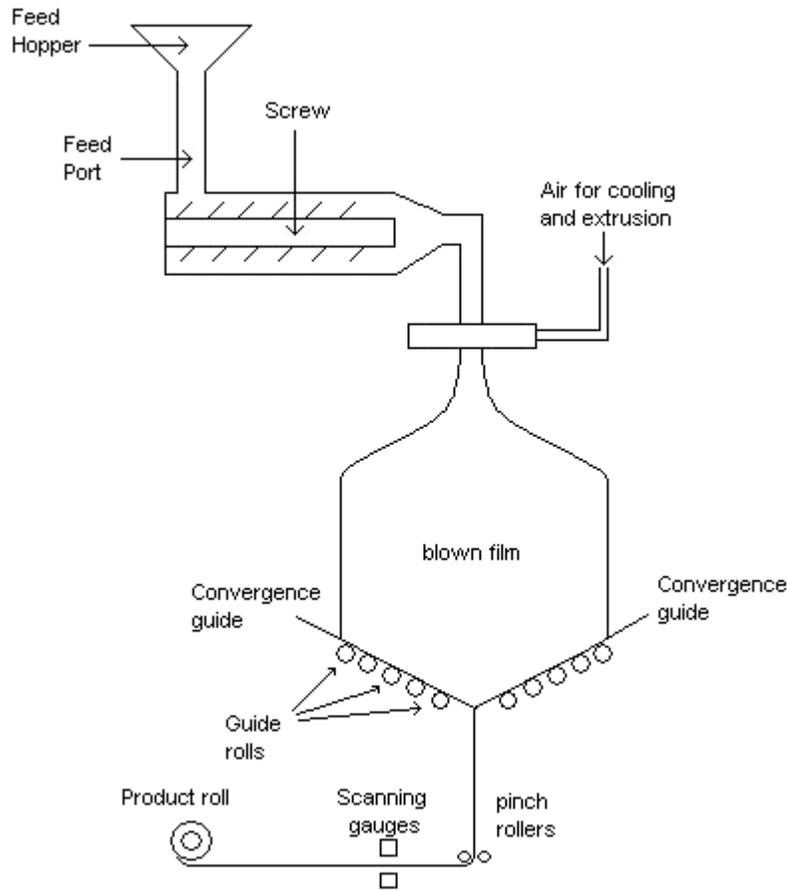


Figure 2.3: Schematic of a plastic blown film extrusion process

blown film lines, MD variations should be within $\pm 2\%$ while CD variations may be between $\pm 3\%$ to $\pm 15\%$ depending on the product. In certain cases, opacity and organic content are also measured.

2.3 Metal rolling

Metal rolling is performed on a Sendzimir mill in order to reduce the thickness of the sheet. A schematic of a Sendzimir mill is shown in Fig. 2.4. Gauge variations in the MD and CD are the primary defects found in flat-rolled steel products. Defective strips typically have long edges such that they do not lie flat or the caliper near the edges are lower than at the center. Edge trimming takes up to 2% of the total metal produced. It is estimated that a reduction of 1% of edge trimming may save \$3 million per year (Norbury Jr (1996)). Shapemeters and X-

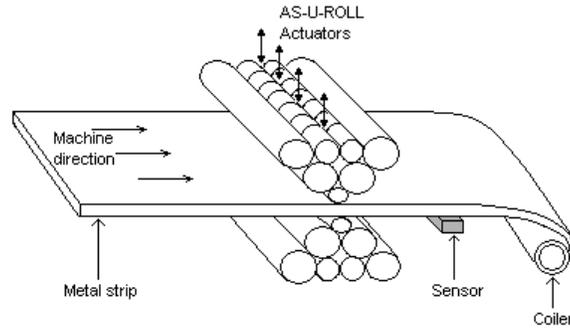


Figure 2.4: Schematic of a Sendzimir mill

ray profile sensors are used to provide simultaneous profile measurements across the machine.

2.4 Coating

Coating processes are commonly used to produce an even coat of adhesives for the productions of laminates of two sheets or films. This technique is used to produce photographic films, medical, pharmaceutical, and food packaging, and products from bumper stickers to optical and medical media storage devices (VanAntwerp et al. (2000)). Feed rollers supply a substrate which passes between a roller and a stainless steel die. The die has a cavity to apply the adhesive to the substrate, whose flowrate is determined by a controlled pump. The die gap can be varied by a set of actuators that locally deform the die. The operation of the die gap actuators is similar to the operation of slice-lip actuators used in paper basis weight control. After coating, the substrate goes through a drier and the coat weight is measured through a traversing sensor.

2.5 Scanning versus stationary sensors

With the advancement of sensing technology, a new kind of sensors that scan the full width of the sheet during every time step has become possible. These sensors, known as ‘full scan sensors’, greatly reduce errors related to the estimation of the states of the process. An analysis on the influence of sensor location and placement on the covariance of the closed-loop sensor output has been presented in Tyler

and Morari (1995). In Chang et al. (2001), image-based analysis is proposed to discriminate between different scanning patterns when more than one scanning sensor is present on the same gauge. The measures used for discrimination are the mean and standard deviations of the time to the nearest neighbour.

2.6 Modelling

Due to their nature as distributed parameter systems, the underlying dynamics of sheet and film forming processes are best described using partial differential equations (PDE) in space and time. These equations can be developed based on the physics of the process. However, such an approach is difficult and prone to uncertainties. For example, the parameters of the PDEs are likely to change with the physical properties of the sheet and in some processes, such as paper manufacture, changes in the physical properties can be naturally expected along the machine direction due to the drying process. Still, approximating the PDEs as linear and time-invariant allows the use of two-dimensional transfer function models. The identification and control of two dimensional transfer-function models for sheet and film forming processes has been studied in Wellstead et al. (2000), Gacon et al. (1996) and the references within. However, in general, two dimensional models are likely to be computationally too expensive for practical use (Wellstead et al. (1996), Thake (1997)), and, therefore, discussions of methods that use such models are omitted from this thesis.

As discussed in section 2.7, it is common to separate the MD and CD models into two separate control problems. In this approach, the MD model is typically an ARMA process, and needs little further description. The CD models are MIMO transfer functions, and are built based on the assumption that the response of the CD actuators can be separated into their dynamics and CD gain. In other words, the CD actuator response is modelled as the product of a gain vector and a scalar transfer function. This scalar transfer function (dynamics of the CD actuator) is usually a low-order AR process, such as first-order. If it is further assumed that the dynamics of all CD actuators are the same, then the scalar gain vectors of all CD actuators are combined together to form the *interactor* matrix. The input-output model is of the form shown below:

$$Y(t) = \frac{B}{A(z^{-1})}U(t) + N(t) \tag{2.1}$$

where $Y(t)$ is the vector of CD properties at time t ,

$$Y(t) = \begin{bmatrix} y_1 & \dots & y_j & \dots & y_m \end{bmatrix}'$$

$U(t)$ is the vector of CD actuator moves at time t ,

$$U(t) = \begin{bmatrix} u_1 & \dots & u_i & \dots & u_n \end{bmatrix}'$$

B is the interactor matrix, which is an $m \times n$ matrix of steady-state CD gains, and is a non-square matrix in general, $A(z^{-1})$ is a scalar polynomial representing the dynamics of the CD actuators, and $N(t)$ is the disturbance model.

The noise dynamics in eqn. (2.1) are modelled using an ARIMA process with correlated white noise as its input (Bergh and MacGregor (1987)):

$$\phi(q^{-1}) \nabla I_m N_t = \theta(q^{-1}) a_t$$

where ∇ is the difference operator, I_m is the identity matrix of dimension m , ϕ and θ are AR and MA polynomials of q^{-1} , respectively, and a_t is a vector white noise process with correlation matrix Σ_a , of the form

$$\Sigma_a = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^m \\ \rho & 1 & \rho & \ddots & \vdots \\ \rho^2 & \rho & 1 & \ddots & \rho^2 \\ \vdots & \ddots & \ddots & \ddots & \rho \\ \rho^m & \dots & \rho^2 & \rho & 1 \end{pmatrix}$$

where $0 < \rho < 1$. The value of ρ determines the extent of spatial correlation in the noise process, and depends on the nature of the process being modeled. Due to the time-varying nature of the scanning sensor, estimation of the MA process $\theta(q^{-1})$ is difficult, and therefore models of the form $\phi(q^{-1}) \nabla I_m N_t = a_t$ are typically used. The integrator in the noise model is used to model the drifting nature of the disturbances in sheet and film forming processes. A more detailed explanation for the use of an integrator is presented in Rawlings and Chien (1996).

Early work on CD control assumed a square interactor matrix as the number of measured locations were usually equal to the number of actuator locations. However, due to advances in sensing technology, the number of sensing locations have dramatically increased in recent years. If all the sensing locations are used

in the model, the size of the interactor matrix would become computationally too large. Therefore, most of today's CD algorithms reduce the dimension m , by performing a linear transformation, to match the dimension n (VanAntwerp et al. (2007)). In Heaven et al. (1994), it is claimed that several simulations and analytical studies indicate better control when the original profile is used instead of the transformed profile. In Gorinevsky et al. (2000), however, it is claimed that the performance of a mapped controller can be made to match the performance of a minimum-variance controller by using a mapping window of the same shape as the CD actuator response. The number of sensing locations before transformation is usually in the order of hundreds to tens of thousands while the number of actuator locations is usually in the order of hundreds. According to Gorinevsky et al. (2000), in most practical cases, $3 \leq m/n \leq 12$.

In order to reduce the number of parameters estimated, and increase the confidence on the estimates of the parameters, further assumptions are typically made on the structure of the interactor matrix. The assumptions made depend on the underlying process and can be crucial in determining the success of a CD system (Featherstone and Braatz (1997)). The structure assumed for the interactor matrix may be one of the following three structures (VanAntwerp et al. (2007)):

1. Centrosymmetric

In this structure, it is assumed that the machine is exactly symmetric about the center of the CD profile. Hence, the response of actuators on either side of the machine placed at an equal distance from the center of the machine are the same. One of the key features of a centrosymmetric model is that actuators at the edges may have a different model compared to actuators away from the edges. This allows edge effects to be modelled more accurately. An

example of a centrosymmetric interactor matrix is shown below

$$\begin{pmatrix} g_{1,0} & g_{2,-1} & g_{3,-2} & 0 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ g_{1,1} & g_{2,0} & g_{3,-1} & \ddots & 0 \\ g_{1,2} & g_{2,1} & g_{3,0} & \ddots & g_{p,-2} & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & g_{2,2} & g_{3,1} & \ddots & g_{p,-1} & g_{p+1,2} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & g_{3,2} & \ddots & g_{p,0} & g_{p+1,1} & g_{p,2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & \ddots & g_{p,1} & g_{p+1,0} & g_{p,1} & \ddots & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & g_{p,2} & g_{p+1,1} & g_{p,0} & \ddots & g_{3,2} & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & g_{p+1,2} & g_{p,-1} & \ddots & g_{3,1} & g_{2,2} & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & g_{p,-2} & \ddots & g_{3,0} & g_{2,1} & g_{1,2} \\ 0 & \ddots & g_{3,-1} & g_{2,0} & g_{1,1} \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & g_{3,-2} & g_{2,-1} & g_{1,0} \end{pmatrix}$$

Further assuming that the matrix is symmetric leads to a centrosymmetric symmetric interactor matrix.

2. Toeplitz

In this model structure, it is assumed that all actuators have the same response, and, therefore, the machine is identical along the cross-direction. The resulting interactor matrix is band diagonal. An example of a Toeplitz interactor matrix is shown below

$$\begin{pmatrix} g_0 & g_{-1} & g_{-2} & 0 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ g_1 & g_0 & g_{-1} & \ddots & 0 \\ g_2 & g_1 & g_0 & \ddots & g_{-2} & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & g_2 & g_1 & \ddots & g_{-1} & g_{-2} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & g_2 & \ddots & g_0 & g_{-1} & g_{-2} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & \ddots & g_1 & g_0 & g_{-1} & \ddots & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & g_2 & g_1 & g_0 & \ddots & g_{-2} & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & g_2 & g_1 & \ddots & g_{-1} & g_{-2} & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & g_2 & \ddots & g_0 & g_{-1} & g_{-2} \\ 0 & \ddots & g_1 & g_0 & g_{-1} \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & g_2 & g_1 & g_0 \end{pmatrix}$$

Due to the assumption that all actuators have a similar response, edge effects are likely to be poorly modeled, and can lead to poor control near the edges.

3. Circulant Symmetric

For processes with a circulant symmetry, such as in blown film extrusion, the resulting interactor matrix shows a circulant symmetric structure. An example of a circulant symmetric interactor matrix is shown below

$$\begin{pmatrix} g_0 & g_1 & g_2 & 0 & \dots & \dots & 0 & \dots & 0 & g_2 & g_1 \\ g_1 & g_0 & g_1 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & g_2 \\ g_2 & g_1 & g_0 & \ddots & g_2 & 0 & \ddots & \ddots & 0 & \ddots & 0 \\ 0 & g_2 & g_1 & \ddots & g_1 & g_2 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & g_2 & \ddots & g_0 & g_1 & g_2 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 0 & \ddots & g_1 & g_0 & g_1 & \ddots & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & g_2 & g_1 & g_0 & \ddots & g_2 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & g_2 & g_1 & \ddots & g_1 & g_2 & 0 \\ 0 & \ddots & 0 & \ddots & \ddots & 0 & g_2 & \ddots & g_0 & g_1 & g_2 \\ g_2 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & g_1 & g_0 & g_1 \\ g_1 & g_2 & 0 & \dots & \dots & 0 & \dots & 0 & g_2 & g_1 & g_0 \end{pmatrix}$$

The centrosymmetric and Toeplitz interactor matrices can be considered as transformed versions of a large circulant symmetric matrix (Laughlin et al. (1993)). Furthermore, in Laughlin et al. (1993), a description of the effect of the interaction parameters (g_i) on the singular values of the interactor matrix is presented. These singular values are important as good CD control can only be performed if the interactor matrix is well conditioned, or has few near-zero singular values (Laughlin et al. (1993)). Presence of zero singular values is an indication of a poorly designed process (Stewart (2000)).

2.6.1 Mapping

Mapping is performed in the paper industry due to shrinking of paper in the cross-direction as it dries (Stewart (2000)). This phenomenon is illustrated in Fig. 2.5. As the paper dries, the width of the sheet in the cross-direction reduces. As a result, the response of the CD actuators cannot be related directly to the CD

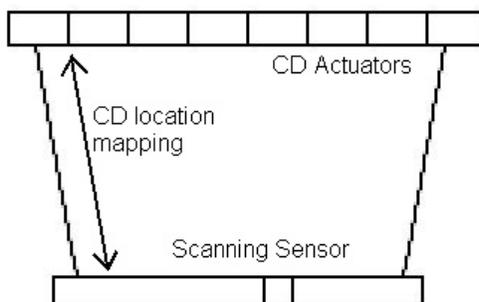


Figure 2.5: Shrinking of paper along the CD

location measured by the sensor. The problem of relating the CD locations before drying to sensor locations after drying is known as the mapping, lane identification or alignment problem (Duncan (1989)). The most common model for mapping is a linear model which assumes a linear relationship between the center of the downstream response c_j and the position of the j^{th} actuator x_j (VanAntwerp et al. (2007)) :

$$c_j = \alpha_1 + \alpha_2 x_j$$

The identification of this linear model is a separate problem in itself and has been considered by many researchers. Publications in this area include Adamy (1997), Anson et al. (2008) and the references within. Poor identification of the mapping model or changes in the mapping with time are some causes for poor CD control.

2.6.2 Constraints

Constraints in the CD process are of the following types (VanAntwerp et al. (2007)):

- Limit constraints

Limit constraints define the physical upper and lower limits of the actuators by specifying their maximum and minimum set points. A limit constraint is typically of the form

$$u_{i,t,min} \leq u_{i,t} \leq u_{i,t,max}$$

where $u_{i,t}$, $u_{i,t,min}$, and $u_{i,t,max}$ refers to the setpoint, minimum setpoint and maximum setpoint that actuator i can take at time t .

- Actuator stress constraints

These constraints, also known as first order bending moment constraints, are enforced to prevent excessive wear or breakage of the actuators. They are especially important in slice-lip type actuators (section 2.1) as large differences between setpoints of adjacent actuators may lead to breaking of the slice-lip. They have the form:

$$|u_{i+1,t} - u_{i,t}| < D_{max}$$

where D_{max} is the largest acceptable difference between adjacent actuator set points.

- Bending moment constraints

Bending moment constraints, also called second order bending moment constraints, are especially important in slice-lip type actuators as they specify the maximum amount of ‘bending’ the slice-lip is allowed to undergo. They are of the form

$$|u_{i+1,t} - 2u_{i,t} + u_{i-1,t}| < B_{max}$$

where B_{max} represents the maximum amount of ‘bending’ the slice-lip is allowed to undergo.

- Rate of change constraint

These constraints are enforced to avoid large actuator moves to prevent actuator wear. They are of the form

$$|u_{i,t} - u_{i,t+1}| < C_{max}$$

where C_{max} represents the maximum allowable change in the actuator set point between subsequent time steps.

2.6.3 Identification

According to Chen and Subbarayan (1999), the performance of CD control applications depends highly on the accurate modelling of the CD actuators. Identification of CD processes is usually done using an industry standard bump-test, where individual CD actuators are sent a ‘bump’ signal and the resulting profile is used to estimate parameters of the interactor matrix. The MD dynamics of the CD actuators are estimated using the spatial center of the response profile and the input signal, while the vector of CD gains for each CD actuator is obtained using the steady-state response profile.

In Heaven et al. (1993), PRBS signals were sent to actuators whose responses were spatially isolated in order to estimate model parameters of a paper-machine. It is found that using a common term for the MD dynamics of all CD actuators results in better parameter estimates. Actuator response profiles were modelled using a second-order critically damped response, along with a linear model for mapping, and found to fit well. A quadratic penalty function (QPF) based controller (discussed in section 2.7.1) is implemented on the machine and better performance is observed compared to PI controllers. In Gorinevsky and Gheorghe (2003), iterative algorithms for identifying CD response profiles and time dynamics that use bump test data and form the core of an industrial identification tool for CD processes are presented. The tool is reported to have been successfully applied to many paper mills and identification results from industrial data has been presented. The identification tool also identifies models for actuator alignment, response shape and process time-response based on bump test results.

Although the bump test is commonly used, it gives poor estimates of the singular values of the interactor matrix (Featherstone and Braatz (1997)), and, for some singular values, the sign of the gains are estimated incorrectly. This sign mismatch causes saturation of control actions along the directions in which the sign of the singular values are incorrectly identified and leads to a zig-zag pattern in the control actions and output, known as ‘picketing’. In Featherstone and Braatz (1997), an iterative algorithm for the design of experiments and estimation of model gains has been provided, and its use is illustrated using a simulated 5-input 5-output blown-film extrusion process. Combined with an SVD-based robust control algorithm, the model is shown to provide good control. In comparison, a QPF controller with a model estimated from bump-test data is shown to perform

poorly.

Orthogonal polynomial representations have been considered quite frequently for CD models, beginning with Kjaer et al. (1995). The advantage of using orthogonal polynomials is the parsimonious representation of the profile and the filtering of high-frequency uncontrollable modes (Heath (1996)). In Kjaer et al. (1995) it is reported that the method has been successfully applied online using an LQG formulation in Heath (1993) to an industrial plastics film extruder.

In Chen and Subbarayan (1999), a method for estimating the CD actuator response model from two-dimensional data obtained using a full-scan sensor has been proposed. The new method simultaneously identifies individual two-dimensional responses of multiple arrays of CD actuators, and reduces the required testing duration and deviations from normal production. The method is reported to have been successfully tested on a production paper-machine and yielded accurate response models for CD control.

2.7 Control

Variations in web characteristics are thought to be the sum of three types of variations. If the variations are represented by $y(x, t)$, with x representing the cross-direction and t representing the machine-direction, then $y(x, t)$ can be written as (Corcadden and Duncan (2000)):

$$y(x, t) = \bar{y} + y_{MD}(t) + y_{CD}(x) + y_R(x, t)$$

where \bar{y} represents the average of the variations, $y_{MD}(t)$ represents variations in the machine direction, $y_{CD}(x)$ represents variations in the cross-direction and $y_R(x, t)$ represents residual variations. In Wang et al. (1993b), it is reported that most MD basis weight variations in paper can be attributed to disturbances occurring prior to the headbox, while most CD basis-weight variations can be attributed to flow patterns inside the headbox. Residual variations typically arise from disturbances such as sensor noise or initial offset. A detailed treatment of the types of variations, frequencies and some typical causes for paper manufacture are presented in Cutshall (1990) and Stewart (2000). High frequency disturbances and disturbances with bandwidth greater than twice the actuator spacing are known to be uncontrollable (Heath (1996), and Duncan (1989)). The problem of estimating

individual variance components based on data measured using dwelling, single, and three gauge scanner configurations has been discussed in Bishop (1963).

As mentioned earlier, two approaches exist for handling the two dimensional control problem. The first and popular approach is to separate the control problem into two separate control loops, with each loop controlling one dimension of the problem. This configuration, known as separate MD and CD control, makes use of the following assumptions:

1. CD profile variations are sufficiently slow that they may be measured and controlled without affecting MD control
2. CD actuators do not have any effect on MD control due to the nature of the actuator

As an example of the second assumption, consider the slice-lip actuators for CD basis weight control on a paper machine. These actuators work by locally deforming the upper lip of the headbox outlet using mechanical screws. This deformation causes a redistribution of flow away from that location but does not affect the total flow out of the headbox. Although slice-lip CD actuators do not show any MD effects, other CD actuators may have an MD effect (Bergh and MacGregor (1987)).

Since many of the singular values of the interactor matrix are usually close to zero, the use of integral action in the CD direction could make the closed-loop system unstable (Fan (2003)). This is due to sensitivity of the controller to uncertainty in the directions corresponding to low singular values. The result is a continuous and erroneous increase in the magnitude of the actuator action in these directions. Hence, integral action is typically applied only to the MD direction.

2.7.1 Separate MD and CD control

This configuration has been widely adopted by both industry and academia. Measurements from the scanning sensor are separated into machine-directional and cross-directional components using one of many such algorithms. The control of each direction is then performed separately, using separate actuators and controllers.

Algorithms for separating MD and CD variations

The most popular algorithm for separation MD and CD variations, proposed by Dahlin (1970), is termed as Exponential Multi-Scan Trending (EXPO). The average of the measurements over a single scan is taken to be the machine-directional component of the variations. MD control can therefore be performed only a few times per scan. To get the cross-directional profile, the MD value is subtracted from the measured profile. The resulting CD profile is weighted against its long-term history value to get an exponentially filtered CD profile that is used for control. As process disturbances are expected to be correlated, these methods may give poor profile estimates (Bergh and MacGregor (1987)). Another issue with the CD algorithm is that it is very sluggish in identifying CD disturbances and results in sluggish control (Thake (1997)).

A modification of the EXPO algorithm has been presented in Balderud and Wilson (2001), where MD and CD values are re-estimated for each sampling time instead of at the end of a scan. The principal advantage of this method is that profile updates are made sooner than EXPO schemes for the sections of the sheet measured. The method is described as follows. Scanner measurements from the beginning of the scan, or from some point in the current scan, upto the current sampling instant, are grouped into a measurement vector and fed into separate MD and CD estimation algorithms. The MD estimation algorithm uses a portion of the zero mean CD profile estimated using the CD estimation algorithm. The portion used consists of the locations included in the measurement vector fed to the algorithm. The deviation of the measurement vector from the previously estimated CD profile at these locations is averaged and used to update the MD estimate using a Kalman filter. To estimate the CD profile, the CD estimation algorithm independently estimates the MD variations at every time step. The MD estimate of the CD estimation algorithm is subtracted from the measured profile and the resulting vector is used to update the CD profile of the section of the sheet measured. The CD profile is updated using an exponential filter with a variable forgetting factor. The forgetting factor used depends on the time elapsed since the location was last sampled. The method is reported to have been applied on a 5-layer board machine in Skoghall, Sweden, as part of a project described in Wilson and Balderud (2000).

In Wang et al. (1993b), moisture content variations and basis-weight varia-

tions in paper are estimated using a non-linear and a linear model, respectively. CD variations are estimated using an exponential forgetting and resetting least-squares algorithm (EFRA), and an extended Kalman filter is used for MD variations. Some model parameters of the moisture content model can be identified online while the entire ARMA model can be identified online for basis-weight variations. Offline results using industrial data are presented in both Wang et al. (1993b) and Dumont et al. (1993). The online identification of parameters makes the algorithms suitable for adaptive control (Wang et al. (1993b)).

In Bergh and MacGregor (1987), a periodic Kalman filter has been proposed for estimating process variations, with the periodic gains calculated offline. The method is described in section 2.7.2. In a similar approach, a dual Kalman filter that simultaneously estimates both MD and CD variations has been presented in Chen (1988).

MD control schemes

The primary limitation for MD control is the large time-delay between the MD actuator and the scanning gauge (Bergh and MacGregor (1987)). As a result, most MD control schemes use dead-time compensation techniques such as Smith Predictor or Dahlin controllers. Some of the first applications of MD control on paper machines are minimum variance controllers, such as in Astrom (1967), and self turning regulators, such as in Borisson and Wittenmark (1974). Wang et al. (1993a) reports that most MD basis-weight control applications in the industry use non-adaptive PI control, or Dahlin control, or Smith Predictor. For the paper industry, the headbox is the key unit where most MD control algorithms can be implemented, and the technical challenges associated with the headbox are covered in Brewster and Bjerring (1970).

In Bialkowski (1983), MD variations are modelled as a first order plus time-delay process with uncertainty in the time-delay. A Kalman filter with an augmented state-space matrix is used for estimating MD variations and LQG theory is used for control. A dead-band and a tuning parameter are used to detune the Kalman filter in accordance with the uncertainty in the time-delay. The tuning parameter increases robustness while compromising performance. The algorithm is reported to have been applied on five paper machine basis weight control systems, one paper machine moisture control system, and one bleach plant pulp brightness control system.

In Wang et al. (1993a), a GPC controller for the control of MD basis-weight variations has been developed to work in conjunction with adaptive estimation methods in Wang et al. (1993b). Simulation results are presented that favourably compare the performance of the controller with a non-adaptive PI controller tuned to provide a critically damped closed loop response. In Baki et al. (2001), three MD basis-weight control schemes (PI, Smith Predictor, and PID with derivative filtering) are compared using a pilot-scale process in UMIST. The PID controller is found to give the best transient and steady-state performance. Derivative action is reported to provide little benefit in the control of processes with large time delay due to the limited predictive capabilities of PID controllers.

CD control schemes

Early work on CD control (Breecher and Bareiss (1970)) used a steady-state model to describe the effect of slice-lip opening on CD basis weight profile. The models were simplified by assuming that the response of a CD actuator is symmetric and identical to all other CD actuators, in order to improve parameter estimation. These control schemes were concerned only with steady-state optimal solutions, and the control moves are typically calculated by inverting the model, resulting in excessive control moves (Boyle (1978)). As steady-state is assumed, long intervals are required between control moves, thereby reducing the effectiveness of control. Owing to the inversion of the model, they are termed Model Inverse (MI) schemes. Model Inverse control schemes are still found in the industry today (VanAntwerp et al. (2007)).

In order to avoid the excessive control actions required by early linear control schemes (such as in Breecher and Bareiss (1970)), Quadratic Programming (QP) formulation were introduced in Boyle (1978) and Boyle (1977). While bending moment constraints were incorporated as soft constraints, limit constraints were incorporated as hard constraints. However, the implementation of the QP formulation was found to be infeasible on existing hardware. Therefore, in Boyle (1977), all constraints were implemented as soft constraints by penalizing them in the control objective, and feasible solutions were obtained by appropriately modifying the input penalty matrix. This approach, termed Quadratic Penalty Formulation (QPF), has been compared favorably in Chen and Wilhelm Jr. (1986) to the QP algorithm. In particular, the ability to increase robustness by increasing the

penalty on control actions, and the smoother control actions in the QPF algorithm has been cited as reasons to suggest QPF as the more practical algorithm. Over 40 installations of a commercial package that uses the QPF algorithm has also been reported in Chen and Wilhelm Jr. (1986), and sample field results are presented. However, these schemes require an accurate model and are generally not suitable for ill/poorly conditioned processes (VanAntwerp et al. (2007)). Hence, the weaknesses of a linear control schemes are that constraints are satisfied only by sufficiently penalizing the control action in the objective function, which can lead to sluggish control, and that robustness of the controller to inaccuracies in the model is not guaranteed (VanAntwerp et al. (2007)).

An alternative and quicker approach to solving the QP problem is presented in (VanAntwerp and Braatz (2000)). In this work, the control actions at each sampling instant are approximated as an ellipsoid with fixed center, directions, and relative lengths but of varying size. The size of the ellipsoid is determined online, in order to reduce conservatism, through interval bisection, and is chosen as the largest ellipsoid that fits into the feasible region bounded by the polytope of constraints. As the directions of the ellipsoid may not match the direction of the control action in the optimal solution of the constrained problem, the success of the algorithm depends on the match between the two. In VanAntwerp and Braatz (2000), the authors have chosen the directions of the ellipsoid based on the SVD of the interaction matrix, and, therefore, these directions are the same as the unconstrained solution. The performance of the algorithm then depend on the similarity between the solution of the unconstrained problem and the solution of the constrained problem. Since certain singular values of the identified model may have poor confidence intervals, the inputs in these directions are set to zero in order to avoid picketing. From the simulation results in the paper, the authors concludes that constraint handling is actually unnecessary for some (but not all) web processes, provided that the control algorithm does not attempt to manipulate the process in uncontrollable directions. The algorithm is termed Robust Ellipsoid (RE) algorithm.

Another approach to reduce computation time for MPC algorithms is to formulate the MPC with constraints only on the current set of inputs (Zheng (1999)). Although this approximation may cause a reduction in closed-loop performance, simulation studies for a 20×20 paper machine example in Zheng (1999) suggest negligible loss.

The use of orthogonal basis functions to extract and control lower order profile information has been proposed by Kristinsson and Dumont (1993) and Ringwood and Grimble (1990), for basis weight control in paper and caliper control in metal rolling, respectively. In both works, Gram polynomials are used to extract the lower order profile information, and parameters of the polynomials are controlled directly instead of controlling the profile. As mentioned earlier, the advantage of this method is the parsimonious representation of the profile, and the filtering of high-frequency uncontrollable models. Kristinsson and Dumont (1993) also states that the mapping problem is eliminated as the entire CD profile is controlled simultaneously. Heath (1996) proposed the use of the Fourier series for blown bubble film extrusion (due to the lack of physical boundaries in the process) and Chebyshev polynomials for bounded web processes. The paper shows that high spectral components of the CD profile are uncontrollable and low spectral components are controllable only if the Gram matrix is sufficiently high, which depends on the number of actuators. The parsimonious representation of the profile using Chebyshev polynomials allows QP algorithms to be implemented with lower computation time. This is demonstrated in Wellstead et al. (1998), where Chebyshev polynomials are used to control an industrial polymer film extrusion process. The QP solver makes use of an iterative scheme known as Mixed Weight Least Squares algorithm (MWLS) found in Rossiter and Kouvaritakis (1993).

In Corcadden and Duncan (2000), Generalized Minimum Variance (GMV) controllers are developed assuming correlated and uncorrelated disturbances, and it is shown that modelling the correlations leads to better control. The disturbances are assumed to be of two types. The first is an infrequent, spatially correlated, low-spatial-frequency disturbances that typically occurs during grade changes, application and removal of the process head, and process drift. The second is random disturbances that occurs during every sampling time interval and may be spatially correlated. Both types of disturbances are assumed to occur simulatenously. In Stewart et al. (1998), a robust GMV controller is proposed that minimizes an objective function of the form

$$J = E \{ |v(t+d)|^2 + |q_{gmv}(z^{-1})u(t)|^2 \} \quad (2.2)$$

where $u(t)$ is the controller action, $q_{gmv}(z^{-1}) = rq(z^{-1})$ is the input penalty transfer function, and $v(t)$ is the mapped error profile and is related to the measured error profile, $p(t)$, through the interactor matrix, G , by $v(t) = G^T p(t)$. This map-

ping is performed to remove uncontrollable components of the measured profile, and results in a ‘square’ input-output equation. Diagonalization is performed on the square input-output equation using a matrix M to decouple the MIMO system into n SISO subsystems, and a GMV controller that minimizes an equation of the form 2.2 is established for each SISO subsystem. Robustnesses to uncertainties in the eigenvalues of $G^T G$ is obtained by choosing a sufficiently large value for r .

Robust control has been applied to sheet and film forming processes beginning with the work of Laughlin et al. (1993). The motivation for this work is that robust control design algorithms such as DK-iterations (Doyle (1987)) and the robust decentralized controller design method of Skogestad et al. (1988) require large computation times and produce high order controllers which are unsuitable for CD processes due to their large-scale nature. In Laughlin et al. (1993), robust stability and performance are first defined for SISO systems, based on the definitions of nominal stability and performance, for a set of possible models containing the uncertainty description. The relationship between SISO and MIMO stability is then established by relating parameter uncertainty in the interactor matrix to eigenvalue uncertainty for a circulant symmetric interactor matrix, and further relating the eigenvalue uncertainty to gain uncertainty for a SISO process, with the last step utilizing transformations. Therefore, the design of a robustly stable MIMO controller is reduced to the design of a robustly stable SISO controller. The design of a robustly stable IMC SISO controller is discussed in Laughlin et al. (1986). An assumption that the interactor matrix is positive definite is made.

In Duncan (1994), methods similar to Laughlin et al. (1993) are developed for robust control of sheet and film forming processes using LQ type-controllers assuming arbitrary interaction matrices, and conditions for robust performance with multiplicative input and output uncertainties are derived in terms of robust performance of SISO subsystems. The assumption made is that the dynamics of the controller can be represented with a scalar term. The controller is therefore the product of a precompensator and scalar dynamics. In VanAntwerp et al. (2001), robust control of sheet and film forming processes under various types of input, output and plant uncertainties are considered, and related literature on robust control of CD processes are easily found in the references within.

Adaptive control has been applied for CD control in a few instances. In Ismail and Dumont (2003), dual adaptive control is applied on an industrial paper coat-

ing machine. The controller used is a linear controller with integral action and soft constraints on control moves instead of control action. A constraint checking algorithm ensures that no constraints are violated. If there is a violation, a QP algorithm with penalty on control effort subject to the standard constraints mentioned in section 2.6 recalculates the control moves. The authors remark that their focus was on developing a dual control algorithm and therefore did not attempt to develop an efficient QP algorithm. Another discussion on the use of adaptive control is in Chen et al. (2008) where a neural network based adaptive control scheme has been used to identify a non-linear model for caliper control on a polymer film extrusion process and has been successfully applied to an industrial process.

A good review of CD control technologies is available in VanAntwerp et al. (2007). As mentioned earlier, the alternative approach to separate MD and CD control is to couple the two control problems into one control problem and simultaneously control both objectives. This configuration is discussed in the next section.

2.7.2 Coupled MD and CD control

An alternative approach for handling the two dimension control problem is to couple the MD and CD control problem into a single control problem and control both objectives simultaneously. The only work known to this author on coupled MD and CD control is Bergh and MacGregor (1987). Since process disturbances are expected to be well correlated, profile estimation is done using a time-varying Kalman filter, and an LQG objective function is used for state-feedback control. The periodic nature of the process causes the steady-state Kalman filter gains to be periodic, and can therefore be computed offline. The LQG objective function minimized using state feedback control is given below

$$J = \lim_{N \rightarrow \infty} \min_{U_t, U_{t+1}, U_{t+2} \dots U_{t+N}} E \left\{ \sum_{k=t}^{t+N} \hat{Y}_k^T Q \hat{Y}_k + U_k^T R U_k \right\} \quad (2.3)$$

where \hat{Y}_k is the vector of estimated or predicted deviation of sheet properties at time k at various locations across the sheet, U_k is the vector of inputs at time k , Q and R are the output and input weighting matrices, respectively, and N is the control horizon. When Q and R are chosen to be diagonal matrices, the control

objective is to minimizing the weighted variance of future states while penalizing the weighted variance of the input. Simulation results using a sample model in Bergh and MacGregor (1987) are presented in section 3.4.

2.8 Performance Assessment

The objective of Controller Performance Assessment (CPA) is to evaluate the performance of a controller in order to identify the need to retune a controller or reidentify the process model. The motivation behind this area arises from viewing the controller as an asset that needs to be managed in order to ensure high returns. Research on CPA began with the ground-breaking study by Harris (1989), where it was shown that routine operating data can be used to estimate the minimum achievable output variance for a process. Since then, the area of controller performance assessment has grown sizably, with some applications being reported Jelali (2006). The most recent review of CPA and its application is provided in Jelali (2006).

One of the very few publications on performance assessment of sheet and film forming processes is on the assessment of CD processes in Duncan et al. (1999). In this paper, it is assumed that both the plant model and interactor matrices are completely known but the disturbance model is not completely know. An ARMAX structure is assumed for the disturbance model. A minimum-variance controller that is detuned inorder to account for constraints is assumed. Algebraic manipulations provide an expression for the difference between the optimum and actual output in terms of actual output and input for the assumed model and controller. This difference is used for performance assessment. The delay-free part of the MA polynomial in the disturbance model is estimated using MD data, and the complete MA polynomial is estimated using filtered output and input. While calculating the difference between optimal and actual outputs, a unitary matrix obtained from the SVD of the interactor matrix is used to extract the controllable modes in the output. The method is tested on data from a 600-output, 235-input paper machine and shows streaks of relatively poor performance. The streaks are attributed to the detuning of the existing controller in order to avoid control of higher order spatial modes.

In Duncan et al. (2000), a minimum variance controller that controls only the controllable and predictable disturbances is used to propose a performance

index for CD processes. The complete model is assumed to be known. In a related work in Taylor and Duncan (2005), however, this assumption is relaxed and the disturbance model is assumed to be unknown. The identification of the noise model is solved by assuming a vector autoregressive (VAR) model and using a Bayesian approach for parameter estimation. The method is tested on a plastic film extruder and indicates that 11% of observed variations could have been eliminated by a minimum variance controller.

Chapter 3

Trade-off in coupled control of sheet and film forming processes

3.1 Introduction

The control of sheet and film forming processes is typically divided into two separate problems of Machine-Direction (MD) and Cross-Direction (CD). Separate control of MD and CD variations has been studied extensively, and is commonly found in the industry. A good review of MD control schemes is presented in Dumont (1986) and a review of CD control schemes is presented in VanAntwerp et al. (2007). The following assumptions are usually made in order to separate the control problems: (i) CD variations are relatively slow compared to the MD variations (ii) CD actuators, such as slice-lip actuators which are used in basis weight control of paper machines, do not have an MD effect as they merely redistribute the flow locally. In many processes, such as basis-weight control of paper, CD control is relatively more important as poor CD control results in windability and rollability problems.

While separate MD and CD control has been explored extensively, coupled MD and CD control is a relatively unexplored topic. In coupled MD and CD control, CD actuators can be used to control MD variations and MD actuators can be used to control CD variations, resulting in better control than separate MD and CD control. One reason for considering coupled control is that, in many cases, CD actuators may have an effect on MD control Bergh and MacGregor (1987), and these interactions need to be considered. In the opinion of the authors,

another justification for coupled MD and CD control is that CD actuators can help overcome the effect of large time delay typically found in the MD actuators on closed loop performance. Coupled MD and CD control, however, is relatively unpopular since minimizing the variance may not result in adequate control of CD variations as they are slower than MD variations. Therefore, they have so far been proposed only for processes with fast CD variations. In this chapter, we propose a coupled MD and CD control scheme that will allow one the flexibility to trade-off between MD and CD control, and therefore allow good CD control for processes with slow CD variations. In addition, we show the existence of a simple relationship between the variance of the outputs and MD and CD objectives.

In section 3.2, it is shown that minimizing the variance of the states corresponds to weighting the MD and CD control objectives equally in a linear quadratic objective function. When the variance of the states is equally weighted, that is, the output weighting matrix Q is a multiple of the identity matrix, the objective function can be interpreted as minimizing, with equal weighting, the MD and CD control objectives. When the variance of the states is not equally weighted, or, for any symmetric, positive definite Q matrix in general, the objective function can be interpreted as minimizing, with equal weighting, the MD and CD control objectives defined on a new set of variables that are scaled and/or transformed versions of the original variables. The results of section 3.2 suggest that a trade-off can be established between the MD and CD control objectives by decoupling them in the objective function. This flexibility is important because the control loop may be well tuned in one direction but poorly tuned in the other direction, leading to overall unsatisfactory results. Therefore, a two-dimensional objective function that allows this trade-off is proposed in section 3.3. Implementation of the objective function in a standard LQG framework, and guidelines for appropriate choice of machine and cross-direction weights are discussed.

In section 3.5, model parameters that relate to the interaction of MD and CD control are derived and used as a tool to judge the potential for improvement by coupling of MD and CD control. This analysis is presented using the sample model in section 3.4, and the trade-off between MD and CD control is illustrated by simulations in section 3.6. Fully weighting the control in the machine direction is seen to result in excellent control of the average of the profiles with time, but poor profiles at individual time instants. Weighting the control heavily in the cross direction is seen to result in excellent profiles at every time instant but

poor control of the average profiles with time. Thus a trade-off can be made by adjusting the relative weightings appropriately.

3.2 Decoupling of MD and CD control objectives

The Linear Quadratic Gaussian (LQG) objective function for coupled MD and CD control proposed in Bergh and MacGregor (1987) is given below

$$J = \min_{U_t, U_{t+1}, U_{t+2} \dots U_{t+N}} \lim_{N \rightarrow \infty} E \left\{ \sum_{k=t}^{t+N} \hat{Y}_k^T Q \hat{Y}_k + U_k^T R U_k \right\} \quad (3.1)$$

where \hat{Y}_k is the vector of estimated or predicted deviation of sheet properties at time k at various locations across the sheet, and U_k is the vector of inputs a time k with the form:

$$U_k = \begin{pmatrix} \nabla v_k \\ u_{1,k} \\ \vdots \\ u_{n,k} \end{pmatrix}$$

where v_k is the set-point of a ‘lumped’ actuator that affects all outputs uniformly, and $u_1 \dots u_n$ are the set-point of actuators that are designed to affect only outputs in their neighbourhood. The actuators are described in greater detail in section 3.4. The penalty matrix R takes the form

$$R = \begin{pmatrix} r_v & 0 \\ 0 & r_u \end{pmatrix}$$

where r_v is the penalty on the control move of v_k and r_u is the penalty on the control actions of $u_1 \dots u_n$. When Q and R are chosen to be diagonal matrices, the control objective is to minimize the weighted variance of future states while penalizing the variance of the input.

3.2.1 When Q is a multiple of the identity matrix

Consider the first term of the objective function, $\hat{Y}_k^T Q \hat{Y}_k$, and assume the special case where Q is of the form $Q = r \times I$, where I is the identity matrix. The first term can be expanded as:

$$\hat{Y}_k^T Q \hat{Y}_k = \sum_{i=1}^n r y_{i,k}^2 = \sum_{i=1}^n r ((y_{i,k} - \bar{y}_k) + \bar{y}_k)^2 \quad (3.2)$$

$$= \sum_{i=1}^n r (y_{i,k} - \bar{y}_k)^2 + \sum_{i=1}^n r \bar{y}_k^2 + 2 \sum_{i=1}^n r \bar{y}_k (y_{i,k} - \bar{y}_k) \quad (3.3)$$

where y_i represents the i th element of \hat{Y}_k and \bar{y}_k is the simple average of $y_{i,k}$ at the time instant k : $\bar{y}_k = \frac{1}{n} \sum_{i=1}^n y_{i,k}$. Using this definition, the third term in the LHS of eqn. (3.3) disappears to give

$$\hat{Y}_k^T Q \hat{Y}_k = r \sum_{i=1}^n (y_{i,k} - \bar{y}_k)^2 + r n \bar{y}_k^2 \quad (3.4)$$

The interpretation of the expansion is simple. The first term represents the cross-directional cost function and the second term represents the machine directional cost function. Eqn. (3.4) shows that minimizing the variance of the states in the standard LQG formulation constitutes minimizing the cross-directional and machine directional costs with equal weight. Although this insight is developed for the case where Q is a multiple of the identity matrix, it can be generalized to other forms of the Q matrix also, as shown below.

3.2.2 When Q is a diagonal matrix

Suppose the diagonal elements of Q are not equal (the elements of Q are written as q_i), the expansion is then rewritten as

$$\begin{aligned} \hat{Y}_k^T Q \hat{Y}_k &= \sum_{i=1}^n (\sqrt{q_i} y_{i,k})^2 \\ &= \sum_{i=1}^n (\sqrt{q_i} y_{i,k} - \bar{y}_k^w)^2 + \sum_{i=1}^n (\bar{y}_k^w)^2 + 2 \sum_{i=1}^n \bar{y}_k^w (\sqrt{q_i} y_{i,k} - \bar{y}_k^w) \end{aligned} \quad (3.5)$$

where \bar{y}_k^w is the weighted average of $y_{i,k}$ at the time instant k with $\sqrt{q_i}$ as the weights: $\bar{y}_k^w = \frac{1}{n} \sum_{i=1}^n \sqrt{q_i} y_{i,k}$. Using this definition, it can be shown that the third term of the eqn. (3.5) disappears to give

$$\hat{Y}_k^T Q \hat{Y}_k = \sum_{i=1}^n (\sqrt{q_i} y_{i,k} - \bar{y}_k^w)^2 + \sum_{i=1}^n (\bar{y}_k^w)^2$$

Therefore, when the diagonal elements of the Q matrix are not equal, using $\hat{Y}_k^T Q \hat{Y}_k$ as the cost function is similar to weighting or scaling the outputs $y_{i,k}$ with $\sqrt{q_i}$. The machine direction cost is defined on the weighted average of the outputs \bar{y}_k^w instead of the simple average \bar{y}_k , and, similarly, the cross-directional cost is defined on the deviation of the weighted output from the weighted average of the outputs, instead of the deviation of the output from its simple average.

3.2.3 When Q is any symmetric, positive definite matrix

If Q is non-diagonal, but symmetric and positive definite, orthogonal decomposition is performed on Q to get $Q = Z^T \Lambda Z_k$. Since the matrix Λ is diagonal, we perform the transformation $P_k = \Lambda^{1/2} Z Y_k$ to gives us the following equation

$$\hat{Y}_k^T Q \hat{Y}_k = P_k^T P_k = \sum_{i=1}^n (p_{i,k})^2 = \sum_{i=1}^n (p_{i,k} - \bar{p}_k)^2 + \sum_{i=1}^n (\bar{p}_k)^2$$

where λ_i s are the diagonal elements of Λ and the average \bar{p}_k is given by $\bar{p}_k = \frac{1}{n} \sum_{i=1}^n p_{i,k}$. Therefore, when Q is symmetric and positive definite, the outputs are scaled and transformed into p_k . The machine direction cost is defined on the average, \bar{p}_k , of the transformed and scaled outputs, and, similarly, the cross-directional cost is defined on the deviation of the scaled and transformed output from its average.

3.3 Decoupled objective function and its implementation

Therefore, the first term of the objective function in equation (4.2), $\hat{Y}_k^T Q \hat{Y}_k$, is rewritten as

$$W_{MD} n \bar{p}_k^2 + \left[(P_k - \bar{P}_k)^T W_{CD} (P_k - \bar{P}_k) \right] \quad (3.6)$$

were $P_k = \Lambda^{1/2} Z \hat{Y}_k$ represents the scaled/transformed outputs, \bar{P}_k is the average written in a vector form:

$$\bar{P}_k = \bar{p}_k \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}_n^T,$$

W_{MD} is the relative machine directional cost, and W_{CD} is the relative cross directional cost. Note that W_{MD} and W_{CD} are scalars. When both W_{MD} and W_{CD} are chosen as 1, the cost function in (3.6) is exactly the same as $\hat{Y}_k^T Q \hat{Y}_k$. The purpose of W_{MD} and W_{CD} is to trade-off between control in the machine and cross-direction. Hence, the proposed objective function is

$$J = \min_{u_t, u_{t+1}, \dots, u_{t+N-1}} \lim_{N \rightarrow \infty} E \left\{ \sum_{k=t}^{t+N} \left([(P_k - \bar{P}_k)^T W_{CD} (P_k - \bar{P}_k)] + W_{MD} n \bar{p}_k^2 + U_k^T R U_k \right) \right\} \quad (3.7)$$

To calculate this objective function, two new states of \bar{p}_k and $(P_k - \bar{P}_k)$ are introduced into the state space model of the system by multiplying the existing state-space model by a matrix C_{obj} defined below. Hence, if the existing states are \hat{Y}_k , the resulting state vector for the new state-space model of the system is:

$$X_{k+1} = C_{obj} Y_k = \begin{bmatrix} P_{k+1} \\ \bar{p}_{k+1} \\ P_{k+1} - \bar{P}_{k+1} \end{bmatrix} \quad (3.8)$$

The form of C_{obj} is given below

$$C_{obj} = \begin{bmatrix} I_n \\ \frac{1}{n} \underbrace{[1 \dots 1]}_n \\ I_n - \frac{1}{n} \underbrace{\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \dots & \vdots \\ 1 & \dots & 1 \end{pmatrix}}_{n \times n} \end{bmatrix} \Lambda^{1/2} Z \quad (3.9)$$

When $n = 3$, for example, C_{obj} is given as

$$C_{obj} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 - \frac{1}{3} = \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & 1 - \frac{1}{3} = \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & 1 - \frac{1}{3} = \frac{2}{3} \end{bmatrix} \Lambda^{1/2} Z \quad (3.10)$$

The objective function in eqn. (3.7) in terms of the new states is then

$$J = \min_{u_t, u_{t+1}, \dots, u_{t+N-1}} \lim_{N \rightarrow \infty} E \left\{ \sum_{k=t}^{t+N} X_k^T \Phi X_k + U_k^T R U_k \right\} \quad (3.11)$$

where Φ is the cost function for the new state-vector and is given by

$$\Phi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & W_{MD}n & 0 \\ 0 & 0 & W_{CD} \end{bmatrix}$$

The advantage of decoupling the objective function in eqn. (4.2) to the objective function in eqn. (3.11) is that a tradeoff can be obtained between MD and CD controls in eqn. 3.11. However, one may pose the question ‘how can one ensure that the input magnitude remains approximately the same while varying W_{MD} and W_{CD} ?’ Although not an exact method, an approximate method for maintaining same magnitude of input is suggested here. If W_{MD} and W_{CD} are chosen as 1, the cost function (eqn. (3.7)) is exactly the same as $\hat{Y}_k^T Q \hat{Y}_k$. A logical idea is to change the weightings W_{MD} and W_{CD} such that the sum remains the same ($W_{MD} + W_{CD} = 2$). This has the effect of maintaining the total output cost ($X_k^T \Phi X_k$) equal to $\hat{Y}_k^T Q \hat{Y}_k$ if $(P_k - \bar{P}_k)$ and \bar{p}_k are of approximately the equal magnitude, and may hence result in the same input magnitude. The success of this method relies on the relative magnitudes of $(P_k - \bar{P}_k)$ and \bar{p}_k . In the numerical example (section 3.6), it is observed that, for the simulated model, keeping the sum of W_{MD} and W_{CD} constant at 2 results in approximately the same magnitude of inputs when the machine-direction or the cross-direction is weighted heavily. This is only an approximate method, however, and does not guarantee that the input magnitude will remain the same.

3.4 Sample model

The following model used by Bergh and MacGregor (1987), will be adopted here as an example. The model has one output, which is the scanner measurement, and five inputs. There are four locations that are periodically measured using the scanning sensor, and the actual value of the physical property at each location is represented by $Y_{i,t}$, where the index i denotes the location and the index t denotes time. v is a lumped actuator that affects all inputs uniformly and is primarily used for MD control. u_i s are a set of cross-directional actuators that manipulate the profile in the neighbourhood of Y_i and are hence primarily used for CD control. The input-output transfer function model, as in Bergh and MacGregor (1987), is given as

$$\begin{bmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \\ Y_{4,t} \end{bmatrix} = \begin{bmatrix} \frac{-0.024}{1-0.88z^{-1}} & 2 & 0.8 & 0 & 0 \\ \frac{-0.024}{1-0.88z^{-1}} & 0.8 & 2 & 0.8 & 0 \\ \frac{-0.024}{1-0.88z^{-1}} & 0 & 0.8 & 2 & 0.8 \\ \frac{-0.024}{1-0.88z^{-1}} & 0 & 0 & 0.8 & 2 \end{bmatrix} \begin{bmatrix} v_{t-7} \\ u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \\ u_{4,t-1} \end{bmatrix} + N_t \quad (3.12)$$

The disturbance $N_t = \begin{bmatrix} N_{1,t} & N_{2,t} & N_{3,t} & N_{4,t} \end{bmatrix}^T$ is modelled as an ARIMA process of the form

$$(1 - 0.9z^{-1}) I_4 \nabla N_t = A_t \quad (3.13)$$

where ∇ is the backshift operator, I_4 is the identity matrix of dimension 4×4 , and $A_t = \begin{bmatrix} a_{1,t} & a_{2,t} & a_{3,t} & a_{4,t} \end{bmatrix}^T$ is a vector white noise sequence with the following covariance matrix:

$$\Sigma = \begin{pmatrix} 1 & 0.8 & 0.64 & 0.512 \\ 0.8 & 1 & 0.8 & 0.64 \\ 0.64 & 0.8 & 1 & 0.8 \\ 0.512 & 0.64 & 0.8 & 1 \end{pmatrix} \times 3.5 \times 10^{-4}$$

The reader is referred to section 2.6 to understand the structure of the model.

The presence of an integrator in the disturbance dynamics implies that the variance of the disturbance grows in an unbounded manner. This is useful for the sake of control design when integral action is considered; however, in practice, the disturbance cannot grow unboundedly. Furthermore, integral action in the CD

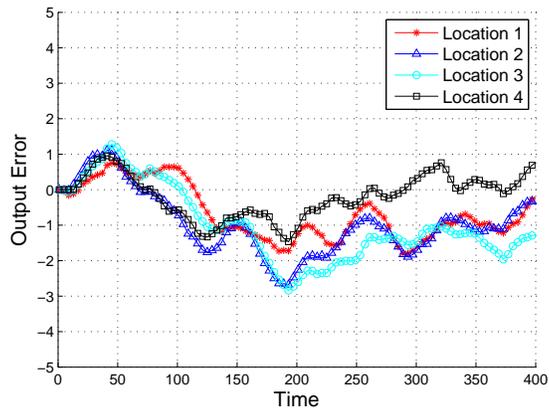


Figure 3.1: Open-loop simulation showing the output of the new disturbance model

direction could result in an unstable controller due to sensitivity to uncertainties in the model (Fan (2003)). Hence, the integrator in eqn. (4.25) is replaced with $(1 - 0.99z^{-1})$ to avoid the use of an intergrator in the CD direction. The new model for the disturbances process is therefore:

$$(1 - 0.9z^{-1})(1 - 0.99z^{-1})I_{4 \times 4}N_t = A_t$$

Open-loop simulations of the above model are shown in Fig. 3.1, and it is observed that this model sufficiently captures the drifting nature of the disturbance.

The transfer function model in eqn. (4.24) can be converted to a state-space model and the scanning sensor can be incorporated in the state space model using a time-varying measurement (C) matrix which varies according to the path taken by the scanner across the surface (Bergh and MacGregor (1987)). The time varying state space system is given by

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + \kappa a_t \\ y_t &= C_t x_t + e_t \end{aligned}$$

where the measurement noise, e_t , is modelled as white noise with variance R_t . A periodically varying Kalman filter can be used to estimate the unmeasured states (Bergh and MacGregor (1987)). The Kalman filter equation is:

$$\hat{x}(t+1|t) = A\hat{x}(t|t-1) + Bu(t) + K_{est}(t)(y(t) - C(t)\hat{x}(t|t-1))$$

where $K_{est}(t)$ is the Kalman gain. The Kalman gain is calculated as

$$K_{est}(t) = AP(t)C^T(t)[C(t)P(t)C^T(t) + R(t)]^{-1}$$

where $P(t)$ is the covariance matrix of the estimates and is found by the solving the following time varying Riccati equation:

$$P(t) = -AP(t)C^T(t)[C(t)P(t)C^T(t) + R(t)]^{-1}C(t)P(t)A^T + AP(t)A^T + GQG^T$$

In practice, the time-varying Riccati equation can be solved before-hand to give periodically varying Kalman filter gains and periodically varying covariance matrices (Bergh and MacGregor (1987)). An approach of calculating the Kalman filter gains by lifting the state-space model is presented in Tyler and Morari (1995). The LQG controller gain is obtained in a standard manner from the constant A and B matrices by solving the discrete algebraic Riccati equation.

Fig. 3.2 shows the open and closed loop response for this system when LQG control is performed with the output cost matrix (Q) and input cost matrices (R) chosen as

$$Q = \mathbf{I}, R = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \end{pmatrix}$$

Fig. 3.3 shows the inputs for the closed-loop response and Fig. 3.4 shows the closed-loop CD profiles of the system at various time points. The noise model is not included in the controller design, but is used in the estimation process.

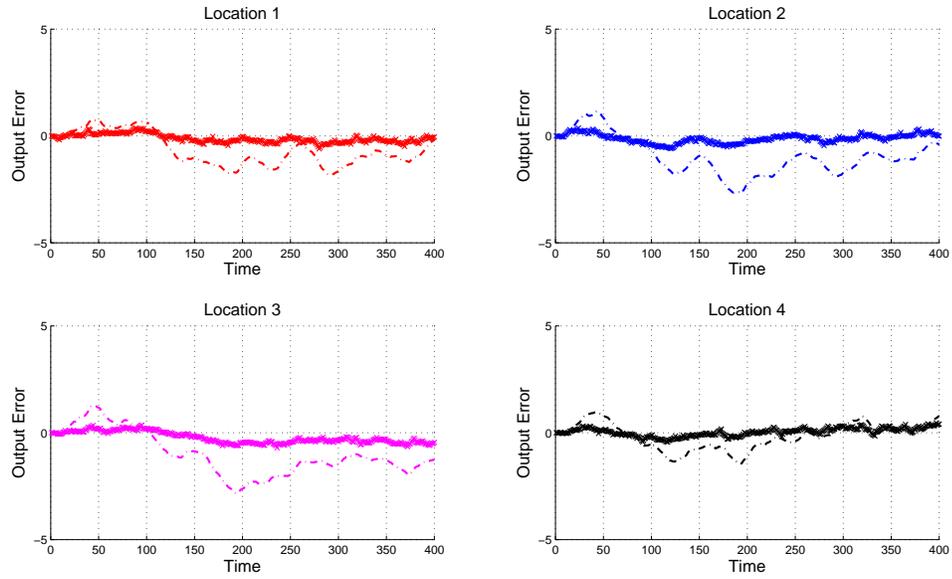


Figure 3.2: Open (dash-dot line) and closed loop (starred line) response under LQG control for a 5-input, 4-output system.

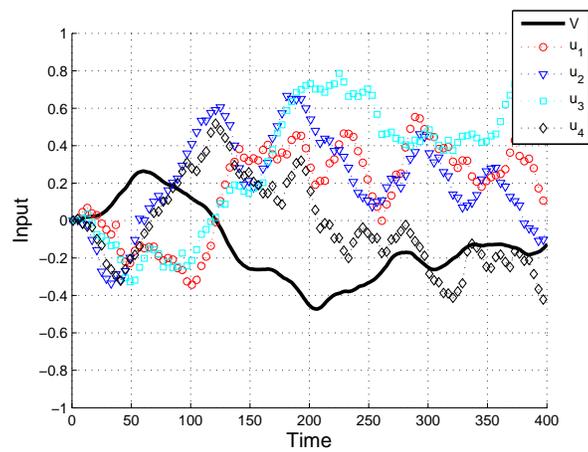


Figure 3.3: Control actions for closed loop response shown in Fig. 3.2

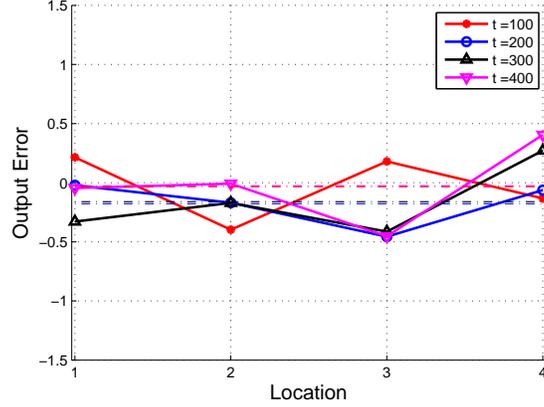


Figure 3.4: Output profiles at various time points for closed loop response in Fig. 3.2 (–. shows average profile).

3.5 Analysis of interaction parameters

The interactions parameters between MD and CD control are easily seen in the new states space model obtained by multiplication of the original state-space model with the C_{obj} matrix (refer to eqn. (3.9)). Consider the model described in section 3.4. Multiplying the input-output model (eqn. (3.12)) with C_{obj} (obtained using $Q = I$) gives the following additional states:

$$\begin{bmatrix} \bar{y}_t \\ Y_{1,t} - \bar{y} \\ Y_{2,t} - \bar{y} \\ Y_{3,t} - \bar{y} \\ Y_{4,t} - \bar{y} \end{bmatrix} = \begin{bmatrix} \frac{-0.024}{1-0.88z^{-1}} & 0.7 & 0.8 & 0.8 & 0.7 \\ 0 & 1.3 & -0.1 & -0.9 & -0.7 \\ 0 & 0.1 & 1.1 & -0.1 & -0.7 \\ 0 & -0.7 & -0.1 & 1.1 & 0.1 \\ 0 & -0.7 & -0.9 & -0.1 & 1.3 \end{bmatrix} \begin{bmatrix} v_{t-7} \\ u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \\ u_{4,t-1} \end{bmatrix} \quad (3.14)$$

$$+ \begin{bmatrix} \underbrace{\frac{1}{n}[1..1]}_n \\ I_n - \frac{1}{n} \underbrace{\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \dots & \vdots \\ 1 & \dots & 1 \end{pmatrix}}_{n \times n} \end{bmatrix} N_t \quad (3.15)$$

Eqn. (3.14) shows the interaction terms clearly. The steady-state gains for MD from the CD actuators are seen as 0.7 for edge actuators and 0.8 for other

actuators. Similarly, the gain for CD from MD actuators are seen to be zero. These terms indicate the potential for improvement available from coupling MD and CD control. As the CD gain from the MD actuator is zero, the MD actuator cannot improve CD control in any manner. This is because, in this model, the MD actuator can only affect the profile in a uniform manner. If, however, the MD actuator is modelled as affecting the profile in a non-uniform manner, a potential exists for CD profile control using the MD actuator. For CD actuators which fundamentally do not affect the MD control, such as slice-lip actuators, the effect of $u_{i,s}$ on \bar{y} is likely to be zero. In such cases, unless the MD actuators affect the profile in a non-uniform manner, no benefit is derived from coupling MD and CD control.

Therefore, for this model, the trade-off is entirely a result of using CD actuators for MD control versus using CD actuators for CD control. Separate MD and CD control may be seen as a special case of coupled MD and CD control where the interaction parameters are not taken into account (and assumed to be zero). Eqn. (3.14) also shows the importance of removing the scan average when identifying CD models. The CD gains are substantially different after removing their effect on the average. However, this effect is likely to be low when the process is of large dimension.

3.6 Results

Consider the model in section (3.4). Assume that the original output cost matrix is $Q = I$. The objective function is decoupled as

$$\begin{aligned} Y_k^T Q Y_k &= \sum_{i=1}^4 (y_{i,k} - \bar{y}_k)^2 + 4\bar{y}_k^2 \\ &= W_{CD} \sum_{i=1}^4 (y_{i,k} - \bar{y}_k)^2 + 4W_{MD}\bar{y}_k^2 \end{aligned} \quad (3.16)$$

where $W_{MD} = 1$ and $W_{CD} = 1$. Simulating the system with the right hand side of eqn. (3.16) as the objective function and $W_{MD} = W_{CD} = 1$ gives exactly the same results as in Figs. 3.2 and 3.3.

The control is weighted heavily in the machine direction by choosing $W_{MD} = 1.875$ and $W_{CD} = 0.125$. For this simulation, the closed loop response is shown in

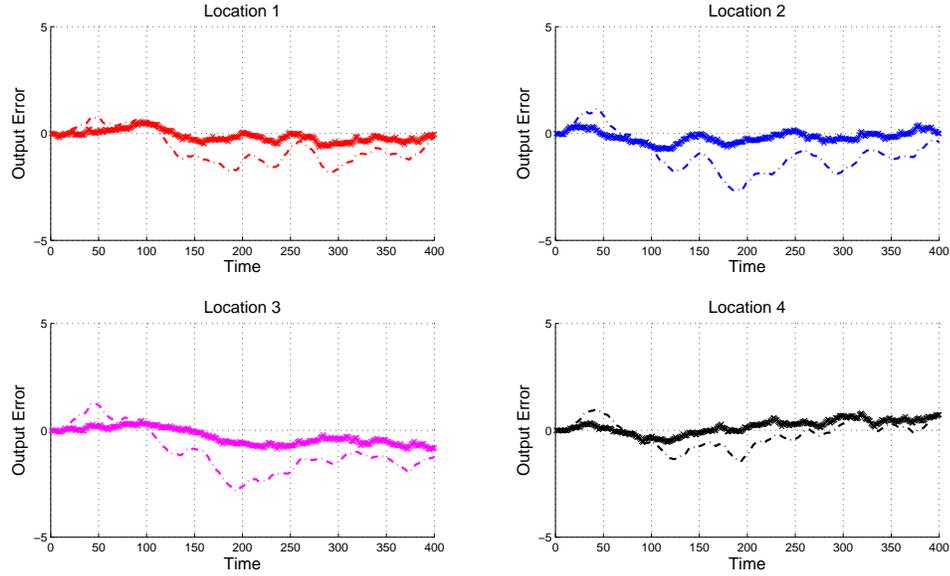


Figure 3.5: Open (dash-dot line) and closed loop (starred line) response under LQG control with heavy weighting in the machine direction

Fig. 3.5, along with the control action in Fig. 3.6, and output profiles in Fig. 3.7.

The closed loop output profiles in Fig. 3.5 are seen to be diverging quite easily. For example, at time $t = 300$, outputs at location 3 and location 4 are substantially away from each other, and in opposite sides of the zero-line. This is illustrative of the poor weighting in the cross-direction. Fig. 3.6 suggests the reason for the poor profile control: all inputs are seen moving in the same direction, except inputs u_4 and V . Input V moves in the opposite direction to the CD actuators as its gain is negative (equation 3.14). The reason for the combined movement of the CD actuators is because the focus of the inputs is on bringing the entire profile as close to the origin as possible, without any consideration of the cross-direction profile. As a result, the average profile is much closer, but the profiles show a more wavy nature. This is shown in Fig. 3.7, where the profiles show significant overlap, but some points, such as location 4 at $t = 400$, are quite far from their neighbours. Since integral action is employed in the MD actuator, MD control is already quite good, and so the improvement towards MD control is relatively small.

The control is weighted heavily in the cross direction by choosing $W_{MD} = 0.125$ and $W_{CD} = 1.875$. This simulation is performed with the same disturbances as in

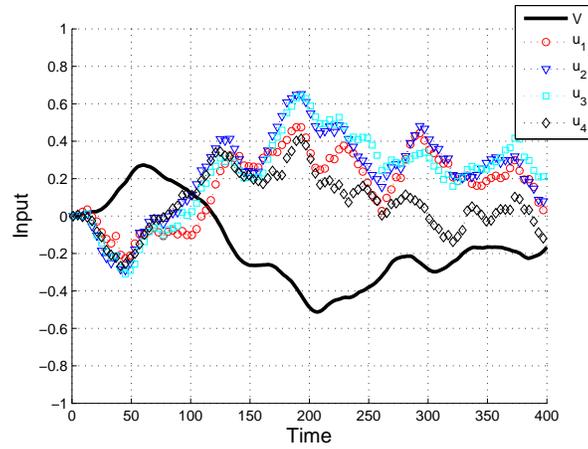


Figure 3.6: Input for closed loop response under LQG control with heavy weighting in the machine direction

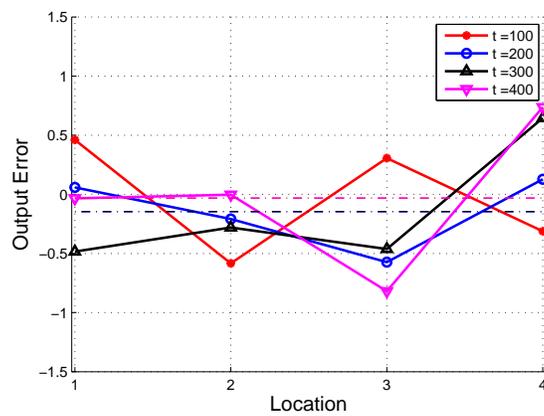


Figure 3.7: Output profile at various time points for closed loop response under LQG control with heavy weighting in the machine direction (— line shows average profile).

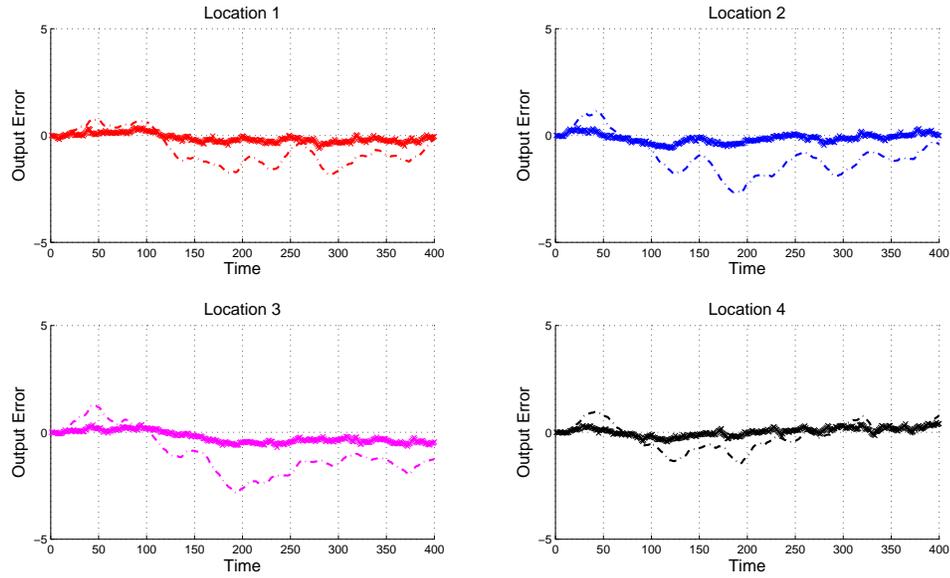


Figure 3.8: Open (dash-dot line) and closed loop (starred line) response under LQG control with heavy weighting in the cross direction

Figs. 3.2 and 3.5. The resulting output is displayed in Fig. 3.8, along with the control action in Fig. 3.9, and the output profile in Fig. 3.10.

The output profiles of Fig. 3.10 are much flatter in comparison to the profiles in Fig. 3.7 and Fig. 3.4, but are also substantially spaced out from each other. Fig. 3.8 shows this clearly as the output for the different locations are seen to move close together, in contrast to the output in Fig. 3.5. Comparing Figs. 3.6, and 3.8 shows that the inputs are of approximately the same magnitudes (within 1.5 and -2), but the CD inputs are further apart from each other when the cross direction is weighted heavily (Fig. 3.8) and close together when the machine direction is weighted heavily (Fig. 3.6). This suggests that the trade-off is a direct consequence of having the choice of varying the inputs more freely either over time or over space. For an easier comparison of the zero-mean cross-directional (CD) profiles, the mean-subtracted output profiles at various time-points are presented in Fig. 3.11, after offsetting according to the time point to enhance visualization.

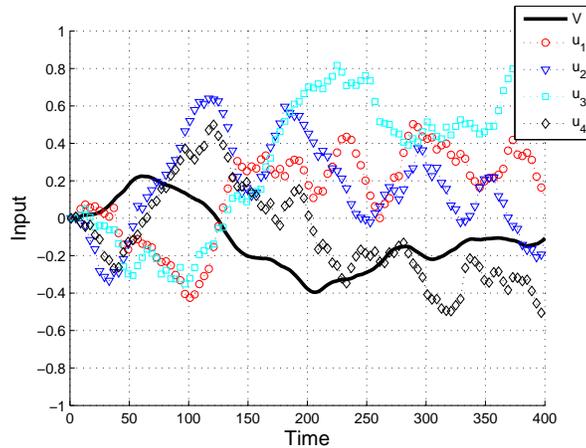


Figure 3.9: Input for closed loop response under LQG control with heavy weighting in the cross direction

3.7 Conclusion

In this chapter, several questions regarding coupled MD and CD control are addressed. In section 3.2, a link has been provided between process variability along MD and CD directions respectively. It is shown that minimizing the variance of the states corresponds to weighting the MD and CD control objectives equally. When the states are equally weighted, the MD and CD control objectives are simply defined as the variances of n times the average output, and the variance of the deviations of the outputs from the average, respectively. When the variance of the states is not equally weighted, the MD and CD control objectives are defined on a new set of variables that are scaled and/or transformed versions of the original variables. This analysis has been used to propose a two dimensional objective function with the ability to trade-off between MD and CD control in section 3.3. The implementation of this objective function under LQG control and the appropriate choice of control weightings are also discussed. In section 3.5, interaction parameters in the model are viewed as a way to identify potential benefits of coupling MD and CD control, and in section 3.6, a numerical example is used to illustrate the concepts and provide an explanation for the existence of the trade-off. The following remarks are in order:

- Conservativeness in the control action can be reduced by iteratively choosing R such that the control actions are large enough but do not violate constraints.

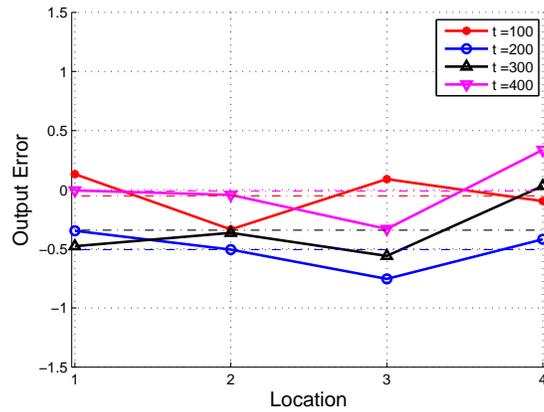


Figure 3.10: Output profile for closed loop response under LQG control with heavy weighting in the cross direction (— shows average profile).

- Other constraints, such as first and second order bending constraints, can be accommodated by introducing soft constraints in the objective function. The reader is referred to Chen and Wilhelm Jr. (1986) to see how this is done.

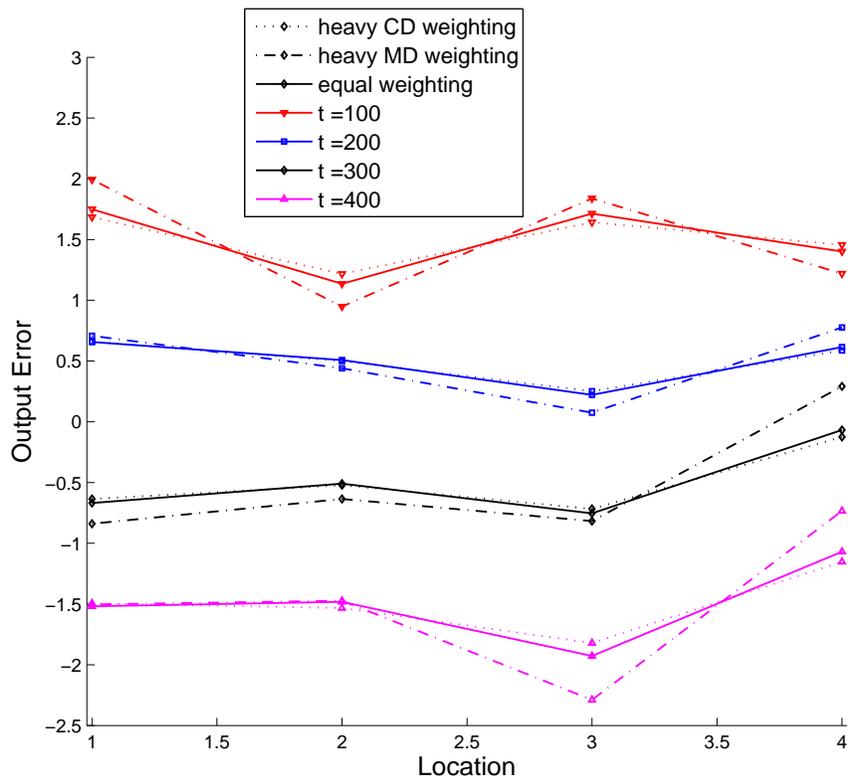


Figure 3.11: Comparison of zero-mean cross-directional profiles for different MD and CD weightings. The following offsets are used: +1.5 for $t = 100$, +0.5 for $t = 200$, -0.5 for $t = 300$, -1 for $t = 400$.

Chapter 4

LQG benchmark for sheet and film forming processes

4.1 Introduction

The Linear Quadratic Gaussian (LQG) tradeoff curve was introduced by Huang and Shah (1999) as an alternative to the minimum-variance benchmark commonly used in Controller Performance Assessment (CPA). This benchmark addresses an outstanding issue of the minimum-variance benchmark which is its inability to take control action into consideration. Minimum-variance controllers are typically not desirable in practice as they result in excessive control actions. The LQG tradeoff curve, however, provides the optimum achievable variance by any linear controller for a given variance of the input (Boyd and Barratt (1991)).

The form of the LQG tradeoff curve is shown in Fig. 4.1. The performance achievable by any linear controller is always at or above the optimum trade-off curve. The x -asymptote of the curve shows the minimum variance, and the y -intercept shows the open loop variance. The point on the trade-off curve with the same x -value as the actual performance (x_1) shows the output variance achievable (y_2) with the given input variance (x_1). Similarly, the point on the trade-off curve that has the same y -value as the actual performance (y_1) shows the minimum input variance required (x_2) for the given output variance (y_1).

In this chapter, a model based method for obtaining the tradeoff curve for sheet and film forming processes, which are a special case of LTV MIMO processes as they have a periodic measurement equation, is developed using the LQG

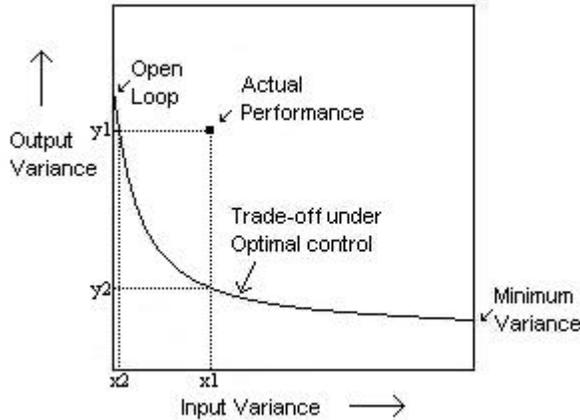


Figure 4.1: LQG tradeoff curve

benchmark. Although methods for performance assessment of sheet and film forming processes already exist (such as Taylor and Duncan (2005) and Duncan et al. (1999)), knowledge of the full profile is typically assumed to be known during control and performance monitoring, and, therefore, the effect of the scanning nature of the sensor usually found in these processes is not considered. The proposed method takes into account the effect of the scanning sensor and provides the variability of the output for a given variability in the input. In addition, the proposed method also provides information on the variability of the closed loop states of the system, which can be used to obtain the tradeoff curve for the individual locations on the sheet and the trade-off curve between MD and CD control for a given input variance. The latter tradeoff can be utilized in controller design to identify the appropriate choice for the relative weights between MD and CD control in coupled control of sheet and film forming processes. Using the proposed method, one can also judge the impact of additional sensors on closed-loop controller performance. Hence, the method is of interest to both vendors and practising control engineers alike.

In section 4.2, the concept of the LQG trade-off curve is extended to time-varying MIMO processes. Section 4.3 discusses the method used for evaluating the closed loop performance of sheet and film forming processes assuming full profile information is available. Such information can be obtained by passing the produced sheet or film through a scanning sensor at a very slow pace in the

machine direction. The sample process model used for simulation (adopted from Bergh and MacGregor (1987)), and the coupled control schemes used to control the process are then briefly described in section 4.4. Simulation results are presented in section 4.5 and section 4.6 concludes the chapter.

4.2 LQG trade-off curve for time-varying MIMO processes

In this section, a method to obtain the LQG tradeoff curve for a time-varying MIMO process is presented. There are three steps to obtaining the trade-off curve:

1. Create the time-varying state space model of the closed loop system
2. Convert the state-space model into time-varying impulse response coefficients
3. Obtain the H_2 norm of the closed-loop system for various input penalties

A state-space model of the open loop system of the form in eqn. (4.1) is assumed to be known.

4.2.1 Time-varying state space model of the closed loop system

Consider the following time-varying state space model

$$\begin{aligned}x_{t+1} &= A_t x_t + B_t u_t + G_t e_t \\y_t &= C_t x_t + D_t u_t + H_t e_t + M v_t\end{aligned}\tag{4.1}$$

with the LQG objective function

$$J = \min_{u_t, u_{t+1}, u_{t+2} \dots} \lim_{\substack{t_0 \rightarrow -\infty \\ t_1 \rightarrow -\infty}} \frac{1}{t_1 - t_0} E \left\{ \sum_{k=t_0}^{t_1} x_k^T Q x_k + \lambda u_k^T R u_k \right\}\tag{4.2}$$

The state-feedback controller $u_t = -L_t \hat{x}_{t|t-1}$ is obtained by solving the time-varying discrete algebraic Riccati equation. The reader is referred to Chapter 6,

Kwakernaak and Sivan (1972), to see how the controller is obtained. An analogous result for the continuous time problem is presented in Chapter 5, Kwakernaak and Sivan (1972). Since the estimated state, $\hat{x}_{t|t-1}$, is related to the actual state, x_t , and the estimation error, $E_{t|t-1}$, by $\hat{x}_{t|t-1} = x_t - E_{t|t-1}$, we get $u_t = -L_t(x_t - E_{t|t-1})$. Substituting for u_t in eqn. (4.1) gives

$$x_{t+1} = A_t x_t - B_t L_t (x_t - E_{t|t-1}) + G_t e_t \quad (4.3)$$

$$\begin{aligned} &= (A_t - B_t L_t) x_t + B_t L_t E_{t|t-1} + G_t e_t \\ y_t &= C_t x_t - D_t L_t (x_t - E_{t|t-1}) + H_t e_t + M_t v_t \\ &= (C_t - D_t L_t) x_t + D_t L_t E_{t|t-1} + H_t e_t + M_t v_t \end{aligned} \quad (4.4)$$

The equation for the Kalman Filter used to estimate the states x_{t+1} is:

$$\begin{aligned} \hat{x}_{t+1|t} &= A_t \hat{x}_{t|t-1} + B_t u_t + K_{est,t} (y_t - C_t \hat{x}_{t|t-1} - D_t u_t) \\ &= (A_t - K_{est,t} C_t) \hat{x}_{t|t-1} + B_t u_t + K_{est,t} (y_t - D_t u_t) \end{aligned} \quad (4.5)$$

Subtracting eqn. (4.5) from eqn. (4.3) gives

$$\begin{aligned} x_{t+1} - \hat{x}_{t+1|t} &= A_t x_t + B_t u_t + G_t e_t - (A_t - K_{est,t} C_t) \hat{x}_{t|t-1} \\ &\quad - B_t u_t - K_{est,t} (y_t - D_t u_t) \end{aligned} \quad (4.6)$$

$$\begin{aligned} &= A_t x_t + G_t e_t - (A_t - K_{est,t} C_t) \hat{x}_{t|t-1} \\ &\quad - K_{est,t} (C_t x_t + D_t u_t + H_t e_t + M_t v_t - D_t u_t) \end{aligned} \quad (4.7)$$

$$\begin{aligned} &= (A_t - K_{est,t} C_t) (x_t - \hat{x}_{t|t-1}) + (G_t - K_{est,t} H_t) e_t \\ &\quad - K_{est,t} M_t v_t G_{u,t} W(t) \end{aligned} \quad (4.8)$$

$$E_{t+1|t} = (A_t - K_{est,t} C_t) E_{t-1|t} + (G_t - K_{est,t} H_t) e_t - K_{est,t} M_t v_t \quad (4.9)$$

The closed-loop state-space model is obtained by adding the states $E_{t-1|t}$ to x_t and combining eqns (4.3), (4.4) and (4.9) as:

$$\begin{aligned}
\begin{bmatrix} x_{t+1} \\ E_{t+1|t} \end{bmatrix} &= \begin{bmatrix} (A_t - B_t L_t) & B_t L_t \\ 0 & (A_t - K_{est,t} C_t) \end{bmatrix} \begin{bmatrix} x_t \\ E_{t-1|t} \end{bmatrix} + \\
&\begin{bmatrix} G_t & 0 \\ (G_t - K_{est,t} H_t) & -K_{est,t} M_t \end{bmatrix} \begin{bmatrix} e_t \\ v_t \end{bmatrix} \quad (4.10) \\
y_t &= \begin{bmatrix} C_t - D_t L & D_t L \end{bmatrix} \begin{bmatrix} x_t \\ E_{t-1|t} \end{bmatrix} + \begin{bmatrix} H_t & M_t \end{bmatrix} \begin{bmatrix} e_t \\ v_t \end{bmatrix}
\end{aligned}$$

This state-space model is well known for its use in the development of the separation principle. The state-space matrices of eqn. (4.10) are redefined in terms of new variables $\bar{A}, \bar{G}, \bar{C}, \bar{H}$, and \bar{K} as follows

$$\begin{aligned}
\bar{A}_t &= \begin{bmatrix} (A_t - B_t L_t) & B_t L_t \\ 0 & (A_t - K_{est,t} C_t) \end{bmatrix}, \quad (4.11) \\
\bar{G}_t &= \begin{bmatrix} G_t & 0 \\ (G_t - K_{est,t} H_t) & -K_{est,t} M_t \end{bmatrix}, \\
\bar{C}_t &= \begin{bmatrix} C_t - D_t L \\ D_t L \end{bmatrix}, \bar{H}_t = \begin{bmatrix} H_t & M_t \end{bmatrix}_t, \bar{K}_t = \begin{bmatrix} -L_t & L_t \end{bmatrix}^T
\end{aligned}$$

The new states of eqn. (4.10) are $X_t = \begin{bmatrix} x_t \\ E_{t-1|t} \end{bmatrix}$ and the inputs are $W_t = \begin{bmatrix} e_t \\ v_t \end{bmatrix}$. The closed loop state-space model of the system in terms of the new variables is

$$\begin{aligned}
X_{t+1} &= \bar{A}_t X_t + \bar{G}_t W_t \quad (4.12) \\
y_t &= \bar{C}_t X_t + \bar{H}_t W_t
\end{aligned}$$

Another closed-loop state-space model of the system, giving the relationship between the states and the input to the plant is

$$\begin{aligned}
X_{t+1} &= \bar{A}_t X_t + \bar{G}_t W_t \quad (4.13) \\
u_t &= \bar{K}_t X_t
\end{aligned}$$

4.2.2 Time-varying impulse response coefficients

In order to obtain the time-varying impulse response coefficients from the closed-loop state space models, the non-commutative property of a time-varying system (Huang (2002)) has to be used. This property is described as follows: for any two LTV transfer functions $A(q^{-1}, t)$ and $B(q^{-1}, t)$, the non-commutative property implies that $A(q^{-1}, t) \times B(q^{-1}, t) \neq B(q^{-1}, t) \times A(q^{-1}, t)$. As a result of the non-commutivity property, performance assessment algorithms for LTV processes are not simple extensions of performance assessment algorithms for LTI processes. Care must be taken during multiplication of LTV polynomials, as the results obtained from assuming the commutivity property can be very different from the results obtained through the non-commutivity property. The time-varying impulse response coefficients of the closed-loop state space model in eqn. (4.12) are obtained as shown below.

The state-space model in eqn. (4.12) is easily converted to the transfer function form below:

$$\begin{aligned} y(t) &= [\bar{C}_t(qI - \bar{A}_t)^{-1}\bar{G}_t + \bar{H}_t] W(t) \\ &= \bar{C}_t (q (I - \bar{A}_{t-1}q^{-1}))^{-1} \bar{G}_t W(t) + \bar{H}_t W(t) \end{aligned} \quad (4.14)$$

$$= \bar{C}_t (I - \bar{A}_{t-1}q^{-1})^{-1} \bar{G}_{t-1} W(t-1) + \bar{H}_t W(t) \quad (4.15)$$

Since

$$(I - \bar{A}_{t-1}q^{-1}) (I + \bar{A}_{t-1}q^{-1} + \bar{A}_{t-1}\bar{A}_{t-2}q^{-2} + \dots) = I \quad (4.16)$$

the term $(I - \bar{A}_{t-1}q^{-1})^{-1}$ can be replaced by $(I + \bar{A}_{t-1}q^{-1} + \bar{A}_{t-1}\bar{A}_{t-2}q^{-2} + \dots)$ to give

$$\begin{aligned} y(t) &= \bar{C}_t (I + \bar{A}_{t-1}q^{-1} + \bar{A}_{t-1}\bar{A}_{t-2}q^{-2} + \dots) \bar{G}_{t-1} W(t-1) + \bar{H}_t e(t) \\ &= \bar{C}_t \bar{G}_{t-1} W(t-1) + \bar{C}_t \bar{A}_{t-1} \bar{G}_{t-2} W(t-2) + \bar{C}_t \bar{A}_{t-1} \bar{A}_{t-2} \bar{G}_{t-3} W(t-3) + \dots \\ &\quad + \bar{H}_t W(t) \\ &= \bar{C}_t \sum_{i=1}^{\infty} \left[\left(\prod_{k=1}^{i-1} \bar{A}_{t-k} \right) \bar{G}_{t-i} W(t-i) \right] + \bar{H}_t W(t) \\ &= G_{y,t} W(t) \end{aligned} \quad (4.17)$$

Similarly, we can obtain the impulse response form of the state-space model in eqn. (4.13) using the steps:

$$\begin{aligned}
u(t) &= [\bar{K}_t(qI - \bar{A}_t)^{-1}\bar{G}_t] W(t) \\
&= \bar{K}_t (q(I - \bar{A}_{t-1}q^{-1}))^{-1} \bar{G}_t W(t) \\
&= \bar{K}_t (I - \bar{A}_{t-1}q^{-1})^{-1} \bar{G}_{t-1} W(t-1)
\end{aligned}$$

Using eqn. 4.16 again,

$$\begin{aligned}
u(t) &= \bar{K}_t (I + \bar{A}_{t-1}q^{-1} + \bar{A}_{t-1}\bar{A}_{t-2}q^{-2} + \dots) \bar{G}_{t-1} W(t-1) \\
&= \bar{K}_t \bar{G}_{t-1} W(t-1) + \bar{K}_t \bar{A}_{t-1} \bar{G}_{t-2} W(t-2) + \bar{K}_t \bar{A}_{t-1} \bar{A}_{t-2} \bar{G}_{t-3} W(t-3) + \dots \\
&= \bar{K}_t \sum_{i=1}^{\infty} \left[\left(\prod_{k=1}^{i-1} \bar{A}_{t-k} \right) \bar{G}_{t-i} W(t-i) \right] \tag{4.18} \\
&= G_{u,t} W(t)
\end{aligned}$$

4.2.3 H_2 norm of the closed-loop system

The H_2 norm of the system between the closed-loop output and noise is obtained from the impulse response matrices in eqn. (4.17) as

$$\begin{aligned}
\|G_{y,t}\|_2 &= \int_0^{\infty} (G_{y,t}) (G_{y,t})^T dt \\
&= \left[\bar{H}_t \bar{H}_t^T + \bar{C}_t \bar{G}_{t-1} \bar{G}_{t-1}^T \bar{C}_t^T + \bar{C}_t \bar{A}_{t-1} \bar{G}_{t-2} \bar{G}_{t-2}^T \bar{A}_{t-1}^T \bar{C}_t^T \right. \\
&\quad \left. + \bar{C}_t \bar{A}_{t-1} \bar{A}_{t-2} \bar{G}_{t-3} \bar{G}_{t-3}^T \bar{A}_{t-2}^T \bar{A}_{t-1}^T \bar{C}_t^T + \dots \right] \tag{4.19} \\
&= \bar{H}_t \bar{H}_t^T + \bar{C}_t \sum_{i=1}^{\infty} \left[\left(\prod_{k=1}^{i-1} \bar{A}_{t-k} \right) \bar{G}_{t-i} \bar{G}_{t-i}^T \left(\prod_{k=1}^{i-1} \bar{A}_{t-i+k}^T \right) \right] \bar{C}_t^T \tag{4.20}
\end{aligned}$$

Similarly, the H_2 norm of the system between the closed-loop input and noise is obtained as

$$\begin{aligned}
\|G_{u,t}\|_2 &= \sum_{t=0}^{\infty} (G_{u,t}) (G_{u,t})^T dt \\
&= \left[\bar{K}_t \bar{G}_{t-1} \bar{G}_{t-1}^T \bar{K}_t^T + \bar{K}_t \bar{A}_{t-1} \bar{G}_{t-2} \bar{G}_{t-2}^T \bar{A}_{t-1}^T \bar{K}_t^T + \right. \\
&\quad \left. \bar{K}_t \bar{A}_{t-1} \bar{A}_{t-2} \bar{G}_{t-3} \bar{G}_{t-3}^T \bar{A}_{t-2}^T \bar{A}_{t-1}^T \bar{K}_t^T + \dots \right] \quad (4.21)
\end{aligned}$$

$$= \bar{K}_t \sum_{i=1}^{\infty} \left[\left(\prod_{k=1}^{i-1} \bar{A}_{t-k} \right) \bar{G}_{t-i} \bar{G}_{t-i}^T \left(\prod_{k=1}^{i-1} \bar{A}_{t-i+k}^T \right) \right] \bar{K}_t^T \quad (4.22)$$

If the eigenvalues of \bar{A}_t are within the unit circle for all t , then higher order terms in the RHS of eqns. (4.19) and (4.21) contribute by smaller amounts to the LHS. Therefore, the infinite series can be truncated to calculate the H_2 norm. From the H_2 norm of the output and the input, the LQG trade-off curve is obtained by varying values of λ in eqn. (4.2).

Note that the above algorithm requires one to select a time-point of interest for obtaining the tradeoff curve, as eqns. (4.20) and (4.22) require one to select a value for t . This should be expected since the variance of a time-varying process is intuitively time-varying as the output variance at each time-instant will be dependent on the process at that time-instant. However, as sheet and film forming processes are periodic in nature, the variances of sheet and film forming processes will be periodic during steady-state behaviour. Hence, a trade-off curve that is generated for any single point in the period of the process using a sufficiently large value for t , to allow for steady-state behaviour, is sufficient for performance analysis.

4.3 Performance evaluation using full profile information

In this section, the assessment of controller performance from closed loop data is discussed assuming that property measurements are available for the entire profile at each sampling instant. Such information could be obtained by passing the produced sheet or film through a scanning sensor at a very slow pace along the machine-direction. Due to the periodicity of the process, the profile information obtained for the section of the sheet measured needs to be segregated. The

segregation of the data for performance assessment is discussed below.

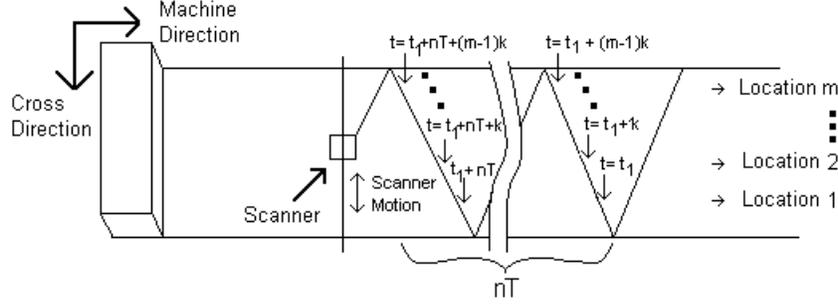


Figure 4.2: Measurements times for section of sheet used for performance assessment

In Fig. 4.2, a section of the output sheet between times $t = t_1$ and $t = t_1 + nT + (m - 1)k$ is shown, where T is the time taken for each scan, n is the number of scans, m is the number of locations of interest, and k is the sampling time. Note that $T = mk$. Assume that full profile information is available for this portion of the sheet. Let the forward scan be defined as the scanning sensor moving diagonally from location 1 to location m , and the backward scan be defined as the scanning sensor moving diagonally from location m to location 1. If $y_{i,t}$ represents the property measurement for location i at time t , then the data obtained for the section of the sheet shown in Fig. 4.2 can be represented using the matrix below:

$$\begin{bmatrix} y_{1,t_1+nT+(m-1)k} & y_{2,t_1+nT+(m-1)k} & \cdots & y_{m,t_1+nT+(m-1)k} \\ \vdots & \vdots & & \vdots \\ y_{1,t_1+k} & y_{2,t_1+k} & \cdots & y_{m,t_1+k} \\ y_{1,t_1} & y_{2,t_1} & \cdots & y_{m,t_1} \end{bmatrix} \quad (4.23)$$

When there are m locations of interest, the period of the process is $2m$ measurements, assuming the commonly used scanning pattern shown in Fig. 4.2. Let each of these periodic cases be denoted by p . Each value of p corresponds to a particular point in the scanning pattern of the sensor. For example, when $m = 4$ the resulting period is 8 sample times, and $1 \leq p \leq 8$. For this case, $p = 1$ corresponds to the measurement of location 1 during the forward scan, $p = 2$ corresponds to the measurement of location 2 during the forward scan, and so on until $p = 8$ corresponds to the measurement of location 1 during the backward scan. Therefore, the p value uniquely determines the position of the scanning sensor. Hence each

row in the data matrix above can be assigned a p value according to the location measured by the scanning sensor at that time instant. The p values assigned for the above matrix is shown below.

$$\left[\begin{array}{cccc}
 y_{1,t_1+nT+(m-1)k} & y_{2,t_1+nT+(m-1)k} & \cdots & y_{m,t_1+nT+(m-1)k} \\
 \vdots & \vdots & & \vdots \\
 y_{1,t_1+2t+2k} & y_{2,t_1+2T+2k} & \cdots & y_{m,t_1+2T+2k} \\
 y_{1,t_1+2T+k} & y_{2,t_1+2T+k} & \cdots & y_{m,t_1+2T+k} \\
 y_{1,t_1+2T} & y_{2,t_1+2T} & \cdots & y_{m,t_1+2T} \\
 y_{1,t_1+(2m-1)k} & y_{2,t_1+(2m-1)k} & \cdots & y_{m,t_1+(2m-1)k} \\
 \vdots & \vdots & & \vdots \\
 y_{1,t_1+T} & y_{2,t_1+T} & \cdots & y_{m,t_1+T} \\
 y_{1,t_1+(m-1)k} & y_{2,t_1+(m-1)k} & \cdots & y_{m,t_1+(m-1)k} \\
 \vdots & \vdots & & \vdots \\
 y_{1,t_1+2k} & y_{2,t_1+2k} & \cdots & y_{m,t_1+2k} \\
 y_{1,t_1+k} & y_{2,t_1+k} & \cdots & y_{m,t_1+k} \\
 y_{1,t_1} & y_{2,t_1} & \cdots & y_{m,t_1}
 \end{array} \right] \rightarrow p = \begin{cases} m & \text{if } n \text{ is even} \\ 2m & \text{if } n \text{ is odd} \end{cases}$$

The data matrix above is therefore separated into $2m$ data sets according to the p value, and only the data set with a p value that corresponds to the same p value as the tradeoff curves is used for variance calculations. The p value of the tradeoff curve is determined by the location of the scanning sensor at the time point chosen for computing the H_2 norms in eqn. (4.19) and (4.21).

4.3.1 Evaluating the performance by location

To evaluate the performance for a particular location, data for the selected location is extracted from the appropriate data set and its variance is calculated to obtain the output variance. The total input variance is obtained as the sum of the variance of all inputs from the appropriate data set. The resulting input and output variances are used to evaluate the performance for the given location along with the appropriate tradeoff curve. An example of the tradeoff curve between output at each location and total input variance will be shown in Fig 4.4.

4.3.2 Evaluating the performance by direction (machine/cross-direction)

The MD ‘output’ is defined as:

$$\bar{y}(t) = \frac{1}{m} \sum_{i=1}^m y(i, t).$$

Therefore, the output at all locations are averaged at each time point in the appropriate data set to obtain the MD output. The variance of the resulting MD output data is used along with the total input variance to evaluate the performance of the MD controller. An example of the MD tradeoff curve will be shown in Fig 4.6 (a).

The CD ‘output’ for the i^{th} location is defined as

$$y_{CD}(i, t) = y(i, t) - \bar{y}(t)$$

The variance of this value is calculated for each of the m location in the appropriate data set, and this is totalled to form the total CD variance. This is used along with the total input variance and the appropriate tradeoff curve to judge the performance of the CD controller. An example of the CD tradeoff curve will be shown in Fig 4.6 (b). It is also possible to obtain the tradeoff between the variance of individual CD outputs and input variance, and an example of this kind of tradeoff curve will be shown in Fig 4.5. Section 4.5 illustrates how the algorithm in section 4.2 is modified to obtain the trade-off between MD variance and input variance, total CD variance and input variance, and individual CD variance and input variance.

By segregating the data according to the p value, and using only the appropriate data set for variance calculations, the periodic nature of sheet and film forming process has been taken into consideration.

4.4 Coupled control of sheet and film forming processes

In this section, the model and control scheme used for the simulations in section 4.5 are described.

4.4.1 Model

The following model used by Bergh and MacGregor (1987), will be adopted here as an example. The model has one output, which is the scanner measurement, and five inputs. There are four locations that are periodically measured using the scanning sensor, and the actual value of the physical property at each location is represented by $Y_{i,t}$. v is a lumped actuator which affects all inputs uniformly, and is used primarily for MD control. u_i s are used primarily to manipulate the cross-directional profile as they influence Y_i and its neighbours in a non-uniform way. The input-output transfer function model, as in Bergh and MacGregor (1987), is given as

$$\begin{bmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \\ Y_{4,t} \end{bmatrix} = \begin{bmatrix} \frac{-0.024}{1-0.88z^{-1}} & 2 & 0.8 & 0 & 0 \\ \frac{-0.024}{1-0.88z^{-1}} & 0.8 & 2 & 0.8 & 0 \\ \frac{-0.024}{1-0.88z^{-1}} & 0 & 0.8 & 2 & 0.8 \\ \frac{-0.024}{1-0.88z^{-1}} & 0 & 0 & 0.8 & 2 \end{bmatrix} \begin{bmatrix} v_{t-7} \\ u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \\ u_{4,t-1} \end{bmatrix} + N_t \quad (4.24)$$

The noise dynamics, N_t , are modelled as an ARIMA process of the form

$$(1 - 0.9z^{-1}) I_4 \nabla N_t = A_t \quad (4.25)$$

where ∇ is the backshift operator, I_4 is the identity matrix of dimension 4×4 , and $A_t = [a_{1,t} \ a_{2,t} \ a_{3,t} \ a_{4,t}]^T$ is a vector white noise sequence with the following covariance matrix:

$$\Sigma = \begin{pmatrix} 1 & 0.8 & 0.64 & 0.512 \\ 0.8 & 1 & 0.8 & 0.64 \\ 0.64 & 0.8 & 1 & 0.8 \\ 0.512 & 0.64 & 0.8 & 1 \end{pmatrix} \times 3.5 \times 10^{-4}$$

The reader is referred to section 2.6 to understand the structure of the model.

The presence of an integrator in the disturbance dynamics implies that the number of terms in the LHS of eqns. (4.20) and (4.22) that need to be considered for calculating the variance is very large. In addition, the presence of the integrator implies that the disturbance grows in an unbounded manner. This is useful for the sake of control design when integral action is considered; however, in practice, the

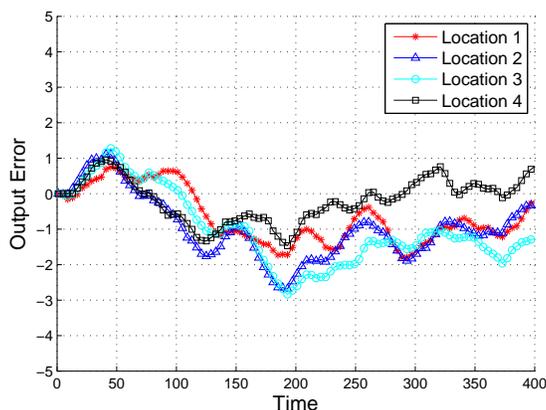


Figure 4.3: Open-loop simulation showing the output of the new disturbance model

disturbance cannot grow unboundedly. Furthermore, integral action in the CD direction could result in an unstable controller due to sensitivity to uncertainties in the model (Fan (2003)). Hence, the integrator in eqn. (4.25) is replaced with $(1 - 0.99z^{-1})$. The new model for the disturbances process is then $(1 - 0.9z^{-1})(1 - 0.99z^{-1})N_t = a_t$. Open-loop simulations of the new model are shown in Fig. 4.3, and it is observed that this model sufficiently captures the drifting nature of the disturbance.

4.4.2 State estimation and control

The transfer function model in eqn. (4.24) can be converted to a state-space model and the scanning sensor can be incorporated in the state space model using a time-varying C matrix which varies according to the path taken by the scanner across the surface (Bergh and MacGregor (1987)). The time varying state space system is given by

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + \kappa a_t \\ y_t &= C_t x_t + e_t \end{aligned} \quad (4.26)$$

where e_t is measurement noise modelled as white noise with variance R_t . Let n denote the number of locations of interest. The state-space model in eqn. (4.26) is multiplied by the matrix C_{obj} defined below to get a new state-space model. Here Λ and Z are derived from the orthogonal decomposition of the control output

weighting matrix, Q , into $Q = Z^T \Lambda Z$.

$$C_{obj} = \left[\begin{array}{c} I_n \\ \frac{1}{n} \underbrace{[1 \dots 1]}_n \\ I_n - \frac{1}{n} \underbrace{\begin{pmatrix} 1 & \dots & 1 \\ \vdots & \dots & \vdots \\ 1 & \dots & 1 \end{pmatrix}}_{n \times n} \end{array} \right] \Lambda^{1/2} Z \quad (4.27)$$

The state vector in the new state space model of the system is written as

$$X_{k+1} = \begin{bmatrix} P_{k+1} \\ \bar{p}_{k+1} \\ P_{k+1} - \bar{P}_{k+1} \end{bmatrix} = C_{obj} Y_k \quad (4.28)$$

with \bar{p}_{k+1} representing the 'machine/time' dimension 'output', and $P_{k+1} - \bar{P}_{k+1}$ representing the 'cross-machine/space' dimension 'outputs'.

A periodically varying Kalman filter is used to estimate the unmeasured states Bergh and MacGregor (1987). The Kalman filter equation is:

$$\hat{x}(t+1|t) = A\hat{x}(t|t-1) + Bu(t) + K_{est}(t)(y(t) - C(t)\hat{x}(t|t-1))$$

where $K_{est}(t)$ is the Kalman gain and is given by

$$K(t) = AP(t)C^T(t)[C(t)P(t)C^T(t) + R(t)]^{-1}$$

where $P(t)$ is the covariance matrix of the estimates and is given by the solution of the following time varying Riccati equation

$$\begin{aligned} P(t) = & -AP(t)C^T(t)[C(t)P(t)C^T(t) + R(t)]^{-1}C(t)P(t)A^T \\ & + AP(t)A^T + GQG^T \end{aligned}$$

In practice, the time-varying Riccati equation can be solved before-hand to give periodically varying Kalman filter gains and periodically varying covariance matrices Bergh and MacGregor (1987).

The Linear Quadratic Gaussian (LQG) objective function in terms of the new states is given as

$$J = \min_{u_t, u_{t+1}, \dots, u_{t+N-1}} \lim_{N \rightarrow \infty} \mathbb{E} \left\{ \sum_{k=t}^{t+N} X_k^T \Phi X_k + U_k^T R U_k \right\}$$

where X_k are the estimated or predicted new states of the system, U_k is the vector of inputs a time k with the form:

$$U_k = \begin{pmatrix} \nabla v_k \\ u_{1,k} \\ \vdots \\ u_{n,k} \end{pmatrix},$$

R is the cost function on the input and Φ is the cost function for the new state-vector and is given by

$$\Phi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & W_{MD}n & 0 \\ 0 & 0 & W_{CD} \end{bmatrix}$$

where W_{MD} is the relative machine directional cost, and W_{CD} is the relative cross-machine directional cost. When W_{MD} and W_{CD} are chosen as 1, this cost function is exactly the same as $Y_k^T Q Y_k$. Note that both W_{MD} and W_{CD} are scalars. The purpose of W_{MD} and W_{CD} is to trade-off between control in the machine and cross-direction. The LQG gain is obtained from the constant A, B, Q and R matrices through the discrete algebraic Riccati equation. That is, the system is time-invariant as far as the design of the controller is concerned since A and B are time-invariant and only C is time-varying.

4.5 Simulation results

The model and control scheme in section 4.4 are simulated to verify the analytical results obtained using the method in section 4.2. The objectives are to find:

- the optimum trade-off curves between input and output variance for each location, for different weightings of MD and CD control

- the trade-off between CD variance and input variance for each location, for different weightings of MD and CD control
- the trade-off between MD variance and input variance, for different weightings of MD and CD control
- the extent of trade-off between MD and CD control, for different input variances

Since the process is time-varying, all of the above can only be found for a particular time instant. However, the periodic nature of the process limits the number of time-points that need to be considered after steady-state behaviour is reached, as the tradeoff curves will also become periodic. Therefore, the tradeoff curve is generated for steady state behaviour at a particular point in the period of the process and this tradeoff curve is used for performance analysis. In the following discussion, by input variance we mean $tr(var(U_k)) = var(\nabla v_k) + var(u_{1,k}) + var(u_{2,k}) + var(u_{3,k}) + var(u_{4,k})$.

In order to calculate the matrices in eqn. (4.10), the Kalman-filter gains and system matrices are required at each time-step. This is easily obtainable in practice if a Kalman filter is already employed for estimating the MD and CD components. Alternatively, the Kalman-filter gains can be calculated offline and substituted in the algorithm. In this work, the gains are obtained by solving the Riccati equation offline. As the controller and Kalman filter are designed independently, the calculation of the Kalman-filter gains is done only once. The controller, however, is recalculated for each setting of the relative weights W_{MD} and W_{CD} , and for each value of λ . For calculating the H_2 norm, the infinite series in eqn. (4.20) and eqn. (4.22) are truncated based on a tolerance for the contributions from higher terms. That is, since higher order terms contribute to the sum by smaller amounts as the order of the term increases, terms that contribute less than a value, say $\epsilon = 10^{-2}$, are not included in the calculation of the norm. For this simulation, the first 175 terms are used to calculate the H_2 norms.

In order to ensure that the algorithm in section 4.2 is truly ‘robust’, the system in section 4.4 is simulated and the variance at the time point of interest (t in eqn. (4.20) and eqn. (4.22)) is calculated based on independent simulations that are performed according to each setting of the controller. Therefore, the output of

a 100 independent simulations at the particular time-point of interest are used for variance calculation. For example, the variance of location 1 in Fig. 4.4 represents its variance over 100 independent simulations of the system for that controller setting for the time-point of interest. In order to allow for the time-varying Kalman filter to reach its steady-state behaviour, the time-point chosen was 207 sampling instants after the simulation begins. At this time-instant, the sensor is at location 2, and has just measured location 3 in the previous time instant (corresponding to $p = 7$). However, after steady-state behaviour was reached, the results were dependent on the p value only, and not the time-point itself.

As mentioned previously, the algorithm can be constructed to provide the variance of the closed-loop states of the system directly. This is done by modifying the \bar{C}_t matrix in eqn. (4.11) to measure the appropriate states from the state-space model in eqn. (4.12). Since the state space model (4.12) represents the augmented state-space system, the first few states of the state-space model correspond to the states X_{k+1} in eqn. (4.28), which are obtained by multiplying the model in eqn. (4.26) with the C_{obj} matrix. Therefore, if the trade-off between the output variance and total inputs variance is desired, then \bar{C}_t will take the form:

$$\bar{C}_t = \left[\begin{bmatrix} I_n & 0 & 0 \end{bmatrix} \quad \dots \right]^T$$

If instead, the trade-off between MD variance and input variance is desired, then \bar{C}_t will take the form:

$$\bar{C}_t = \left[\begin{bmatrix} 0 & n & 0 \end{bmatrix} \quad \dots \right]^T$$

For obtaining the trade-off between CD variance and input variance, \bar{C}_t will take the form:

$$\bar{C}_t = \left[\begin{bmatrix} 0 & 0 & I_n \end{bmatrix} \quad \dots \right]^T$$

The trade-off curve for each location is shown in Fig. 4.4. Three cases are shown: heavy weighting in the cross-direction (a), equal weighting (b) and heavy weighting in the machine-direction (c). In all the three graphs, the y -axis represents the location, and the x - and z -axes represent the input and output variance, respectively. As input variances increases, the output variance is seen to decrease

as expected. When CD control is heavily weighted, the variance of the outputs are large and approximately the same for all locations. However, the variances are smaller when the weightings are equal, and when MD control is weighted heavily. Heavily weighting MD control is seen to lead to lower variance in the central locations, and a slightly larger variance at the edges, suggesting poor profiles around the edges. The first reason is the scanning pattern, since the maximum time that elapses between subsequent measurements is highest at the edges, resulting in larger estimation errors for some values of p . The second reason is the availability of fewer actuators to control edge locations in comparison to central locations. In this model, 2 actuators influence the profile at the edges, against 3 for central locations (refer to eqn. (4.26) or eqn. (4.29)).

Fig. 4.5 shows the trade-off between CD variance and input variance for each location. Note the different y -axis scales in Fig. 4.5 (a), (b) and (c). Again, as expected, increasing the MD control weight results in much larger CD variances. An interesting observation is that even if CD control is heavily weighted, the variance at location 1 is substantially higher than that of other locations. This is because, for the chosen value of p ($p = 7$), the scanner has previously measured location 1 six sampling instances before, and, therefore, the output at location 1 is prone to large estimation errors. Through this analysis, the result of estimation error on CD control is seen for a given p value.

Fig. 4.6 (a) shows the trade-off between MD variance and input variance. As expected, increasing the MD control weight results in much lower MD variance for the same input variance. Aggregating the CD variance over all locations results in a tradeoff curve similar to Fig. 4.6 (a) and this is shown in Fig. 4.6 (b).

The trade-off between MD and CD variance is shown in Fig. 4.7 for various input variances. As input variance increases, the process reaches minimum-variance control and so the curves converge to a point. The x -asymptote of the curves shows the minimum variance in the CD control for a given input variance, and the y -asymptote shows the minimum variance in the MD directions for a given input variance. The data is plotted as a trade-off surface in Fig. 4.8 along with six other input variances for better visualization.

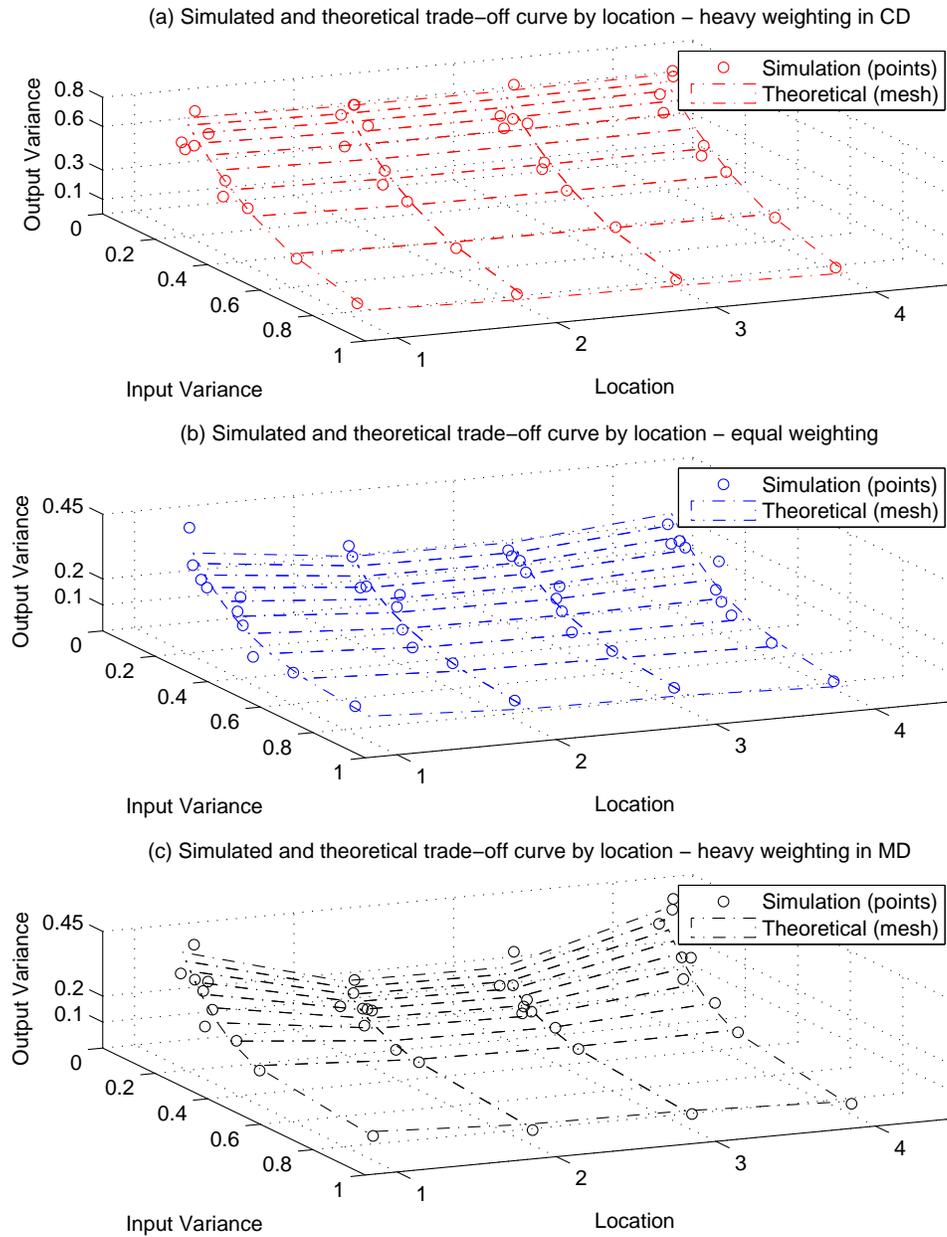


Figure 4.4: Trade-off curve by location with (a) heavy weighting in the CD, (b) equal weighting, and (c) heavy weighting in the MD, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.

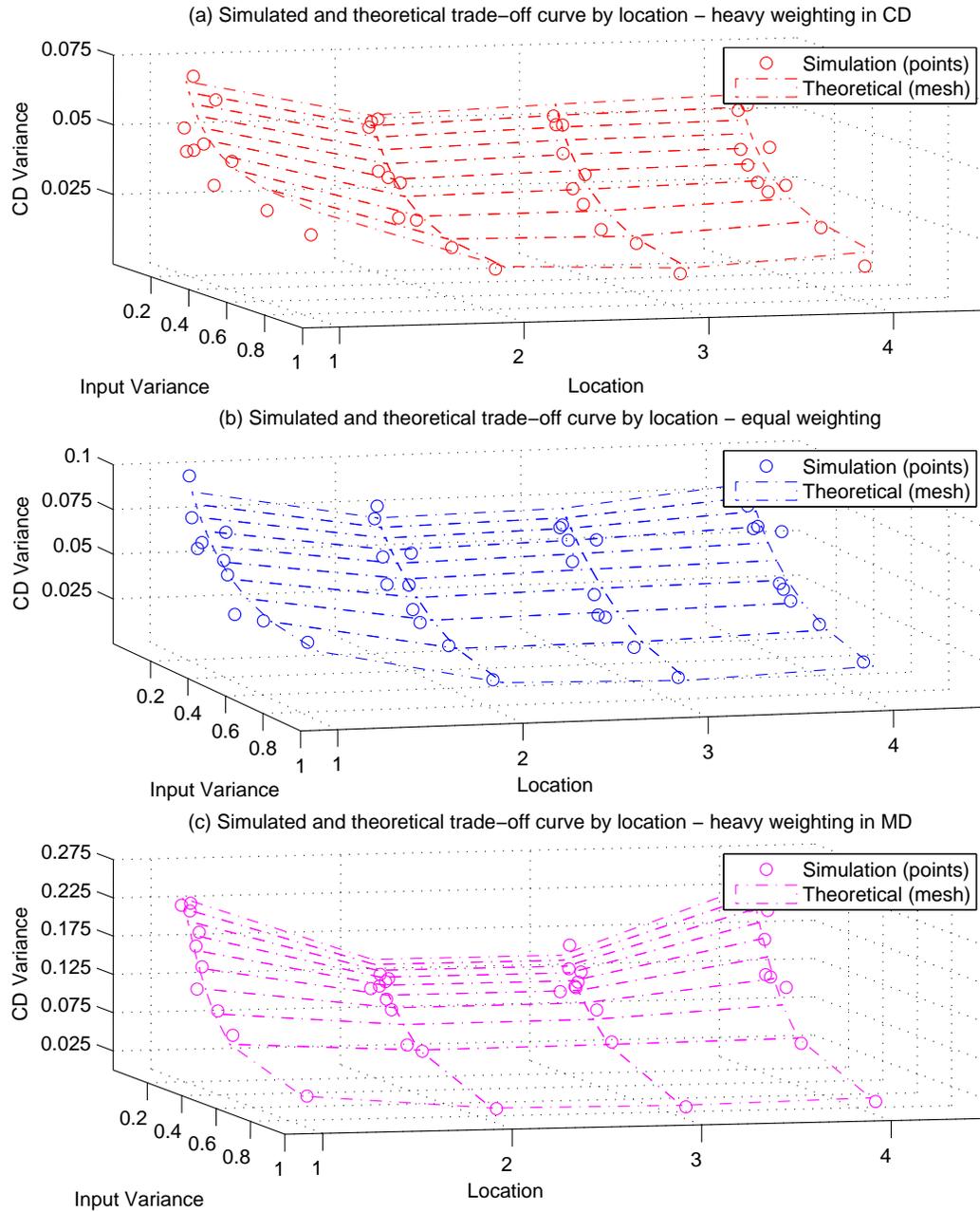


Figure 4.5: CD variance by location with (a) heavy weighting in the CD, (b) equal weighting, and (c) heavy weighting in the MD, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.

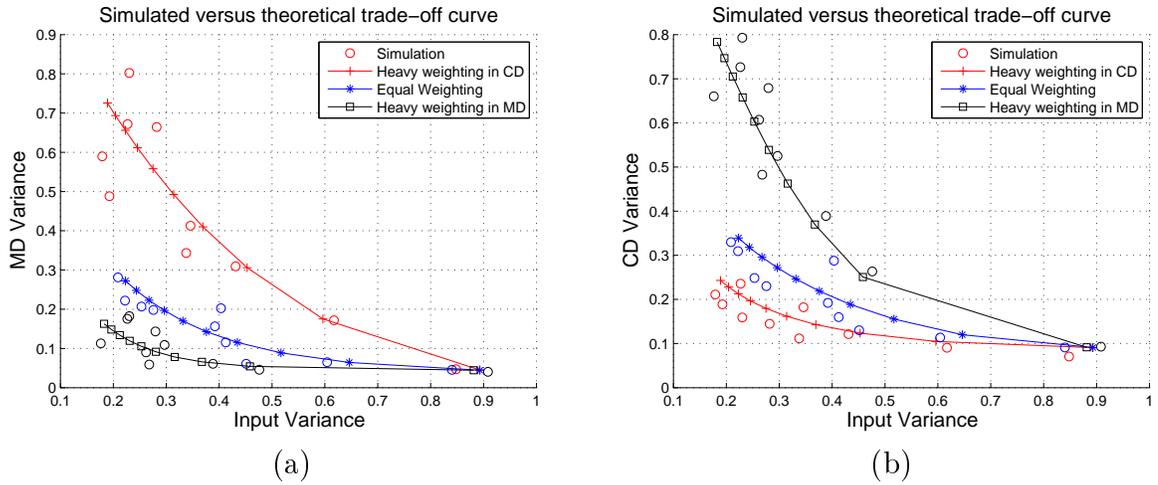


Figure 4.6: (a) MD variance versus input variance trade-off at $t = 207$ sampling instants, and (b) CD variance versus input variance trade-off at $t = 207$ sampling instants. Simulation results represent variance of output over 100 control simulations.

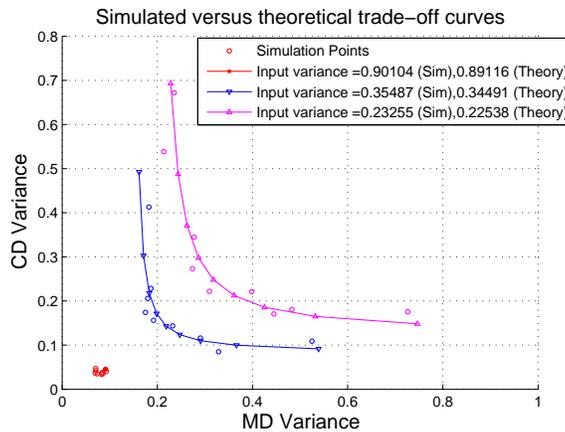


Figure 4.7: MD and CD trade-off for different input variances, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.

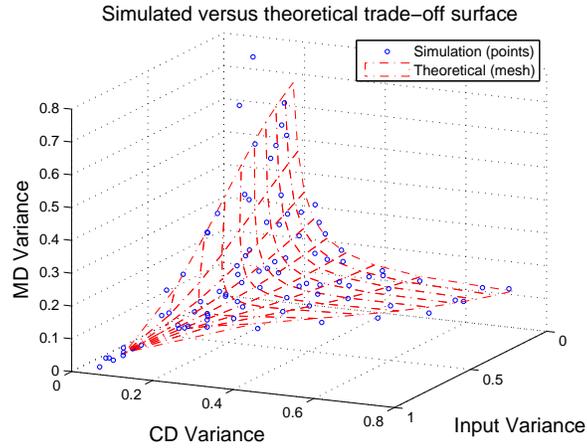


Figure 4.8: CD, MD and input trade-off surface, at $t = 207$ sampling instants. Simulation results represent variance of output of 100 control simulations.

4.5.1 Comparison between PID and LQG controllers as an example

In order to illustrate the use of the above tradeoff curves, the performance of PID controllers performing decoupled control is compared with the performance of coupled LQG controllers using the above curves. For this analysis, the model employed is obtained by multiplying eqn. (4.24) with the C_{obj} matrix in eqn. (4.27). The resulting model has separate models for MD and CD variations and is therefore similar to industrial practice. Assuming that the output weighting matrix is Q is chosen as I , we get the following model:

$$\begin{bmatrix} \bar{y}_t \\ Y_{1,t} - \bar{y} \\ Y_{2,t} - \bar{y} \\ Y_{3,t} - \bar{y} \\ Y_{4,t} - \bar{y} \end{bmatrix} = \begin{bmatrix} \frac{-0.024}{1-0.88z^{-1}} & 0.7 & 0.8 & 0.8 & 0.7 \\ 0 & 1.3 & -0.1 & -0.9 & -0.7 \\ 0 & 0.1 & 1.1 & -0.1 & -0.7 \\ 0 & -0.7 & -0.1 & 1.1 & 0.1 \\ 0 & -0.7 & -0.9 & -0.1 & 1.3 \end{bmatrix} \begin{bmatrix} v_{t-7} \\ u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \\ u_{4,t-1} \end{bmatrix} + C_{obj}N_t \quad (4.29)$$

The MD PID controller is designed based on the MD transfer function $\bar{y}_t = \frac{-0.024}{1-0.88z^{-1}}v_{t-7}$ and, since the gain of the transfer function is negative, the MD controller has negative gains. The CD PID controllers are assumed to have an SVD structure described in VanAntwerp et al. (2000). The SVD of the CD model is given as:

$$svd \left(\begin{bmatrix} 1.3 & -0.1 & -0.9 & -0.7 \\ 0.1 & 1.1 & -0.1 & -0.7 \\ -0.7 & -0.1 & 1.1 & 0.1 \\ -0.7 & -0.9 & -0.1 & 1.3 \end{bmatrix} \right) = \underbrace{\begin{bmatrix} N_1 & N_2 & N_3 & N_4 \end{bmatrix}}_N \underbrace{\begin{bmatrix} 2.49 & 0 & 0 & 0 \\ 0 & 1.64 & 0 & 0 \\ 0 & 0 & 0.7065 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_\Psi \underbrace{\begin{bmatrix} M_1 & M_2 & M_3 & M_4 \end{bmatrix}}_{M^T}$$

where N and M^T are unitary matrix whose inverses are given by N^T and M respectively, and Ψ is a diagonal matrix of singular values. The CD controller takes the form

$$K(s) = M \begin{bmatrix} \frac{K_{CD,1}(s)}{2.49} & 0 & 0 & 0 \\ 0 & \frac{K_{CD,2}(s)}{1.64} & 0 & 0 \\ 0 & 0 & \frac{K_{CD,3}(s)}{0.7065} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} N^T$$

where $K_{CD,1}, K_{CD,2}$ and $K_{CD,3}$ are the three CD PID controllers. Hence, the transfer function to be controlled for all three CD controllers are decoupled SISO models of the form $y_{t,CD} = u_{t-1,CD}$. As the model for the three controllers are alike, the same settings are adopted for all three controllers.

For both PID and LQG controllers, MD and CD variation data obtained from the Kalman filter are used for control. The simulation run length was 10000 sample times in order to minimize the effect of statistical variations. The 10000 samples obtained from each simulation are divided into 8 parts according to the p value, and data from $p = 7$ is used for performance analysis as the tradeoff curves above have been generated for the same p value. Such a large sample size is necessary as the disturbance model has a near-integrating effect, and therefore causes much of the output to be similar over short run lengths. For processes with a sampling time of only a few seconds, however, such a run length translates to only a few hours at most. Many independent simulations are performed for each controller setting to show the extent of statistical variation.

As a starting point, a PID controller that used an IMC based tuning rule was used for MD control while no control was implemented in the CD directions. The tuning was performed using the *sisotool* utility in MATLAB, and the dominant closed loop time-constant required for tuning the controller was set at a significant detuned value of 2.5 sampling times. The resulting controller was:

$$v = (-16.22 + 25.83z^{-1} - 10.2z^{-2}) \overline{y}_t$$

Although the control in the MD direction was good with small MD variances, the input variance was very large (6.0857, 6.1769 and 6.0919). The need for further modification of the tuning parameters was obvious. Reasonably well tuned settings for the PID controllers were then determined after some trial and error. Table 4.1 shows the settings for the PID controllers that were attempted, and Fig. 4.9 (a) and (b) shows the performance of these controllers along with the optimum tradeoff curves from Fig. 4.9 (a) and (b). Little improvement was observed by introducing derivative action to MD control.

The IMC tuning rule for CD control was obtained from Skogestad (2003) as $C(s) = \frac{K_I}{s} = \frac{1}{s(\tau_c + \theta)}$, where $\tau_c = 2$ is the chosen closed loop time constant, $\theta = 1$ is the time delay and $C(s) = \frac{1}{3s}$ is the controller. This control scheme, and some other PID controllers were attempted for CD control. The performance of the PI MD controller in conjunction with the proportional CD controller is seen to be quite close to the optimum tradeoff curve for heavy weighting in the cross-direction, as shown in Fig. 4.9 (a) and (b). When the CD controllers are changed to PI, however, the performance is much further away from the optimum trade-off curve and there is a slight deterioration of MD performance. This is due to the increased effect of interactions, as the effect of CD actuators on MD control is not taken into consideration by the PID controllers, and CD actuators do not aid MD control in any way.

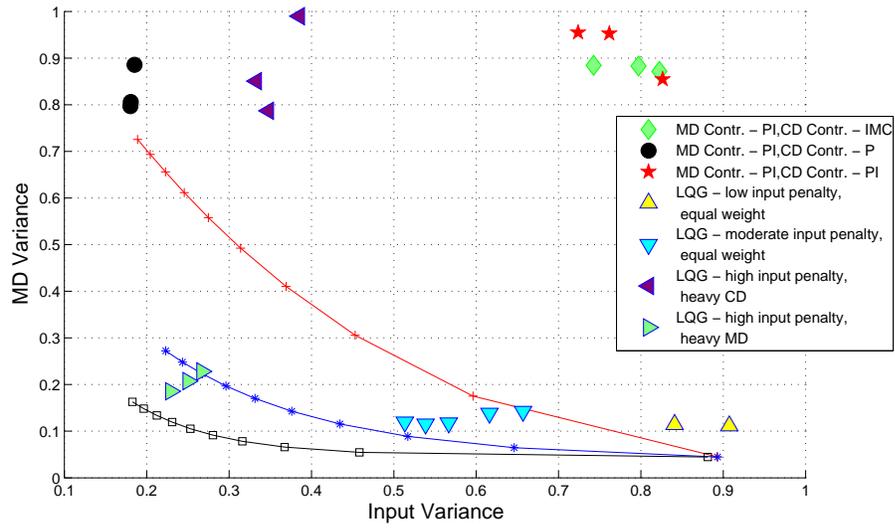
Having attained the best PID settings, the performance of LQG controllers is now evaluated. Table 4.2 shows the various settings of the LQG controllers that were simulated and the representation used to display their performance in Fig 4.9 (a) and (b). Given low input penalty, the LQG controller is seen to easily out perform the best PID controller, especially in terms of MD variance. While the performance of the LQG controllers is close to the optimum tradeoff curve for moderate input penalties, the performance is substantially above the optimum

MD/CD controller type:	K_P	K_I	Representation in Fig 4.9 (a) and (b)	Remark
PID/None (Tuned based on IMC tuning rules)	(see text)	(see text)	-	Large input variance (6.0857, 6.1769 and 6.0919)- omitted from Fig 4.9 (a) and (b)
PI/IMC	-0.70/0	-0.5/0.33	Green diamond (\diamond)	Far from optimum surface
PI/P	-0.70/0.8	-0.5/-	Black circle (\circ)	Close to optimum surface
PI/PI	-0.70/0.8	-0.5/0.3	Red star (\star)	Far from optimum surface

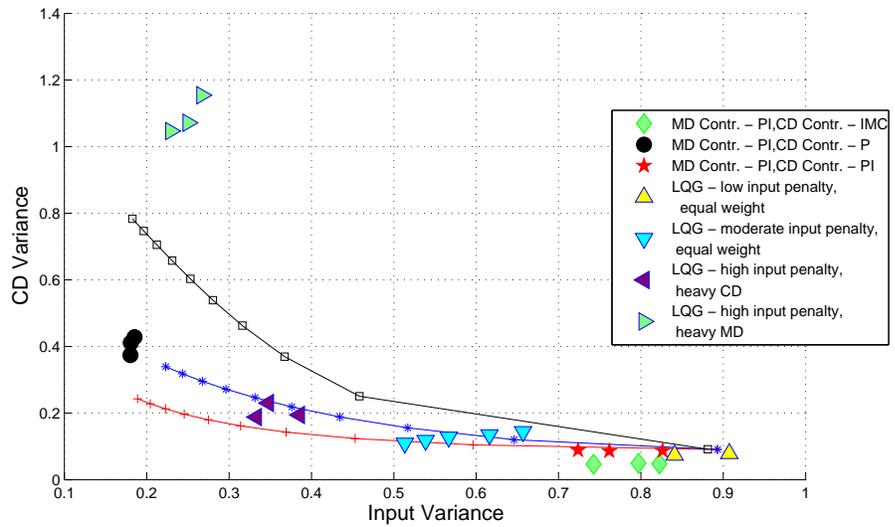
Table 4.1: Controller settings for the PID controllers

Weighting	Q	r_v	r_u	Representation in Fig 4.9 (a) and (b)
Heavy in CD	I	0.5	0.01	Orange triangle (\triangle)
Equal	I	0.1	0.3	Blue triangle (∇)
Heavy in CD	I	4	1.4	Purple triangle (\triangleleft)
Heavy in MD	I	5	2	Green triangle (\triangleright)

Table 4.2: Settings for the LQG controllers



(a)



(b)

Figure 4.9: Comparison between PID controller performance, LQG controller performance, and optimum tradeoff curve

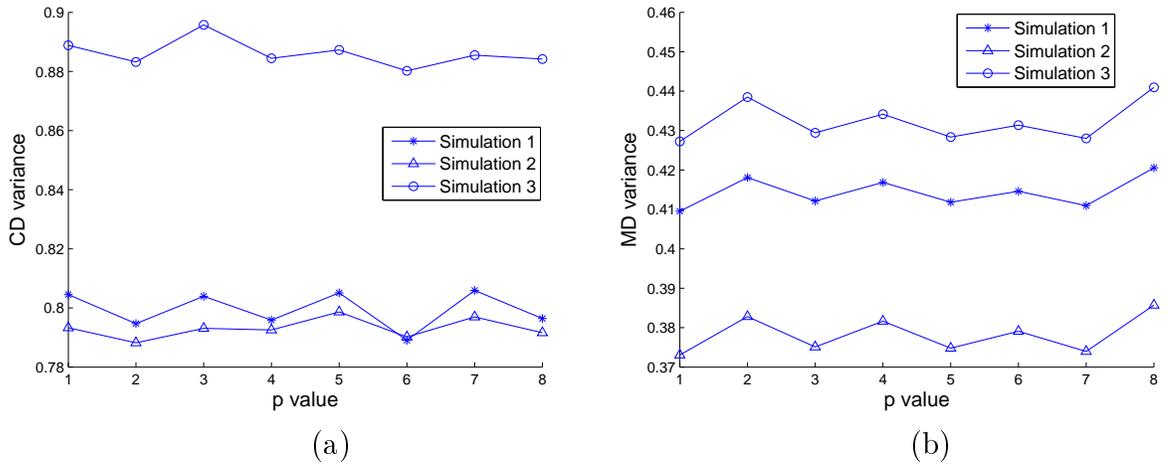


Figure 4.10: Effect of p value on the calculated variances for three independent simulation of a PID controller: (a) CD variance (b) MD variance

tradeoff curve for larger input penalty. One possible reason could be that the performance of the LQG controllers is estimated from data sampled from the simulation. Another possible reason could be the numerical errors involved in obtaining the optimum tradeoff curve, such as the online estimation of Kalman filter gains and the truncation of the infinite series used to calculate the H_2 norms. However, going by the proximity of the LQG controller data to the curve in general, we see that the calculated tradeoff curve is a good representative of the actual tradeoff curve

In order to see the effect of p value on the results, the MD and CD variances obtained for different p values from four independent simulations of the same PID controller are presented in Fig. 4.10 (a) and (b). It is clear from these figures that the calculated variance follows a pattern according to the p values. This underscores the importance of ensuring that the p value used for variance calculations is the same as the p value of the trade-off curve.

4.6 Conclusion

In this chapter, the LQG trade-off curve has been extended to analyse the performance of MIMO systems with time-varying dynamics. This is accomplished through several steps. Closed-loop state-space models are obtained between the output and noise and between the input and noise, and these are converted to impulse response forms. The impulse response coefficients are used to calculate

the H_2 norm of the systems, from which closed-loop input/output variances are obtained. This is done for each setting of the input penalty and is plotted as a tradeoff curve. As the variance of a time-varying process is time-varying, the tradeoff curve varies with time. For a periodic process, however, the number of time-points that need to be considered to identify the tradeoff during steady-state behaviour is limited by the periodicity of the process.

The use of the proposed method in the evaluation of controller performance in sheet and film forming processes is then discussed in section 4.4. The advantage of the proposed method over pre-existing method is its ability to take the estimation error created by the time-varying sensor that is typically used in these processes into account. The proposed method has been supported by simulation results in section 4.5 that employs the model and control scheme described in section 4.4. It is shown that the variance of the states can also be obtained using minor modifications of the proposed method, and the usefulness of this feature has been demonstrated by obtaining the tradeoff curve as a function of location, and as a function of direction (machine/cross-machine direction).

A comparison between PID controllers performing decoupled control, which is the configuration typically found in the industry, and LQG controllers performing coupled control is then performed in section 4.5.1 in order to illustrate the use of the proposed method. As expected, the performance of the LQG controllers is substantially superior to that of PID controllers, as they are much closer to the optimum trade-off curves. When the focus is on good CD control, the results of section 4.5.1 suggest that implementing an LQG controller could decrease the MD variance to a tenth of the MD variance obtained using a PID controller, even if both CD and MD PID controllers are very well tuned.

Chapter 5

Data driven techniques for performance assessment of sheet and film forming processes

5.1 Introduction

The main objective of Controller Performance Assessment (CPA) is to provide an online automated procedure that delivers information to plant personnel for determining whether specified performance targets and response characteristics are being met by the controlled process variables and that evaluates the performance of the control system Jelali (2006). The motivation behind this area arises from viewing the controller as an asset that needs to be managed in order to ensure high returns. Research on CPA began with the ground-breaking study by Harris (1989) where it was shown that routine operating data can be used to estimate the minimum output variance that can be achieved for a process. Since then, the area of controller performance assessment has grown sizably, with many applications being reported Jelali (2006).

Data driven approaches for Controller Performance Assessment (CPA) are more beneficial than model-based approaches because the need to identify a model *a priori* is eliminated in data driven approaches. This is useful since one may not always possess a model of the process, or, if a model is available, the process may have change over time due to natural phenomena such as wear and tear, thereby rendering the model poor. In addition, the model could have been incorrectly identified and therefore may not represent the true process. Hence, a data-driven

approach to CPA is always more desirable than a model based approach. As mentioned earlier, research on CPA began with the ground-breaking study by Harris (1989) where it was shown that routine operating data can be used to estimate the minimum output variance that can be achieved for a process. Hence, the first method for CPA was data-driven.

In this chapter, data-driven techniques for CPA of sheet and film forming processes are discussed. Section 5.2 introduces a measure of performance based on the ACF of the closed-loop output that compares the signal with white noise. In section 5.3, a method is described that allows one to qualitatively assess the relative performance of MD and CD control, using the ACF of the scanning sensor output. This information is useful in determining the priority with which controllers may need to be accessed, and also in identifying the extent of spatial correlation in the disturbance (ρ). In combination with the variance of the scanning sensor, the ACF of the scanning sensor provides a complete, but approximate, picture of the performance of the controller. Various techniques for separation of MD and CD variations from the scanning sensor output are then compared in order to determine their suitability towards individual performance assessment of MD and CD controllers in section 5.4.1. This is followed by a demonstration of the use of the ACF based performance indices for assessment of MD and CD controller performance using the ACF of the estimated MD and CD outputs, in section 5.4.2. The resulting indices allow one to monitor changes in controller performance over time in order to identify the need to retune a controller or reidentify the process model. These techniques could also be used to distinguish between various controller tunings and identify the controller tuning with the best performance.

5.2 ACF based performance index

The ACF of the closed-loop output can be used to create an index for controller performance. When a controller is performing optimally or near optimally, and has not hit any limits posed by the constraints, the closed-loop output of the process resembles noise and has the characteristics of a white-noise sequence, if the time-delay of the process is one sampling instant. Since the ACF of a white-noise sequence is unity for zero lag and zero for higher lags, the ACF of the closed-loop output is ideally unity for zero lag and zero for higher order lags. This ACF plot is shown in Fig. 5.1 (a). When a controller is performing suboptimally,

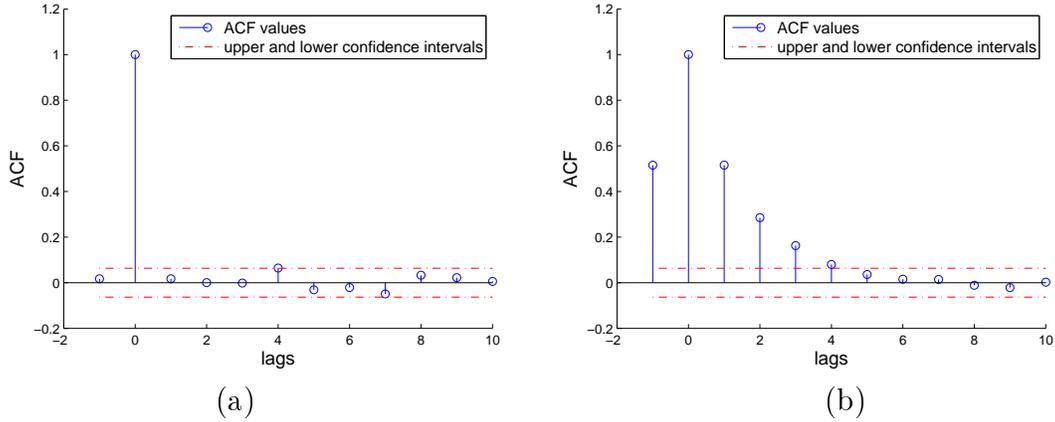


Figure 5.1: (a) ACF of a white noise sequence $\eta_{acf} = 1$ (b) ACF of closed-loop output for a suboptimal controller $\eta_{acf} = 0.46$

the ACF of the closed-loop output will have non-zero values for lags higher than zero, for processes with unit time delay. An example of this ACF is shown in Fig. 5.1 (b). Other reasons for the ACF to delay slowly include:

1. the influence of time decay limiting the performance achievable by the controller
2. the influence of constraints limiting the performance achievable by the controller

Therefore, an index can be created to access the performance of the controller using the ACF of the output. This index, η_{acf} , is defined as:

$$\begin{aligned} \eta_{acf} &= \frac{\sum_{l=0}^{n_{lags}} |\text{ACF OF WHITE NOISE}(lag = l)|}{\sum_{l=0}^{n_{lags}} |\text{ACF OF CLOSED-LOOP OUTPUT}(lag = l)|} \\ &= \frac{1}{\sum_{l=0}^{n_{lags}} |\text{ACF OF CLOSED- LOOP OUTPUT}(lag = l)|} \end{aligned}$$

An increase in the value of the index indicates better performance.

5.3 ACF of the scanning sensor output

The ACF of the scanning sensor output can be used as a preliminary tool to qualitatively identify if MD control or CD control is relatively poor. In order to illustrate this, simulations of the coupled control scheme in Chapter 3 are

performed under different weightings of MD and CD control and the ACFs of the resulting scanning sensor outputs are compared. Fig. 5.2 (a) shows the ACF of the scanning sensor output when control is heavily weighted in the machine direction, Fig. 5.2 (b) shows the ACF of the scanning sensor output when control is heavily weighted in the cross direction, and Fig. 5.2 (c) shows the ACF of the scanning sensor output when control is equally weighted in both directions.

It is easy to see the pattern in the figures. Improved CD control leads to an increase in the correlation for all lags except lags that are multiples of the periodicity of the process. The converse is true for improved MD control. The correlation decreases for all lags except lags that are multiples of the periodicity of the process. The periodicity of the process depends on the scanning pattern employed, but for the commonly used pattern considered in this thesis (displayed in Fig. 1.1), the periodicity is given by $2m$, where m is the number of measurements taken during a scan. For the simulation results in Fig. 5.2, the process model had four outputs, measured one at a time, resulting in a periodicity of 8 sampling times.

The reasoning for the pattern in the ACF of the scanning sensor output is as follows. Improved CD control causes the output profile at a given time to become closer together, and differences between subsequent locations are smaller by definition. Therefore, an improvement in CD control leads to greater correlation between subsequent locations and causes the entire ACF curve to exhibit larger correlation for all lags except lags that are the multiples of the periodicity of the process. For lags that are multiples of the periodicity, however, the correlation corresponds to the one between values measured for the same location at different intervals in time, since the same location is measured after a period according to the definition of a period. Hence, improved CD control does not affect the ACF at lags that are multiples of the periodicity of the process. Consequently, improved MD control leads to larger correlations for lags that are multiples of the periodicity of the process but does not affect the ACF for other lags.

As a result of the above analysis, it is clear that the nature of the ACF of the scanning sensor output provides information on the relative performance of MD and CD control. If the ACF demonstrates large peaks for lags that are multiples of the periodicity of the process, then CD control is poor relative to MD control. If, on the other hand, the ACF demonstrates comparable peaks for all lags, then MD

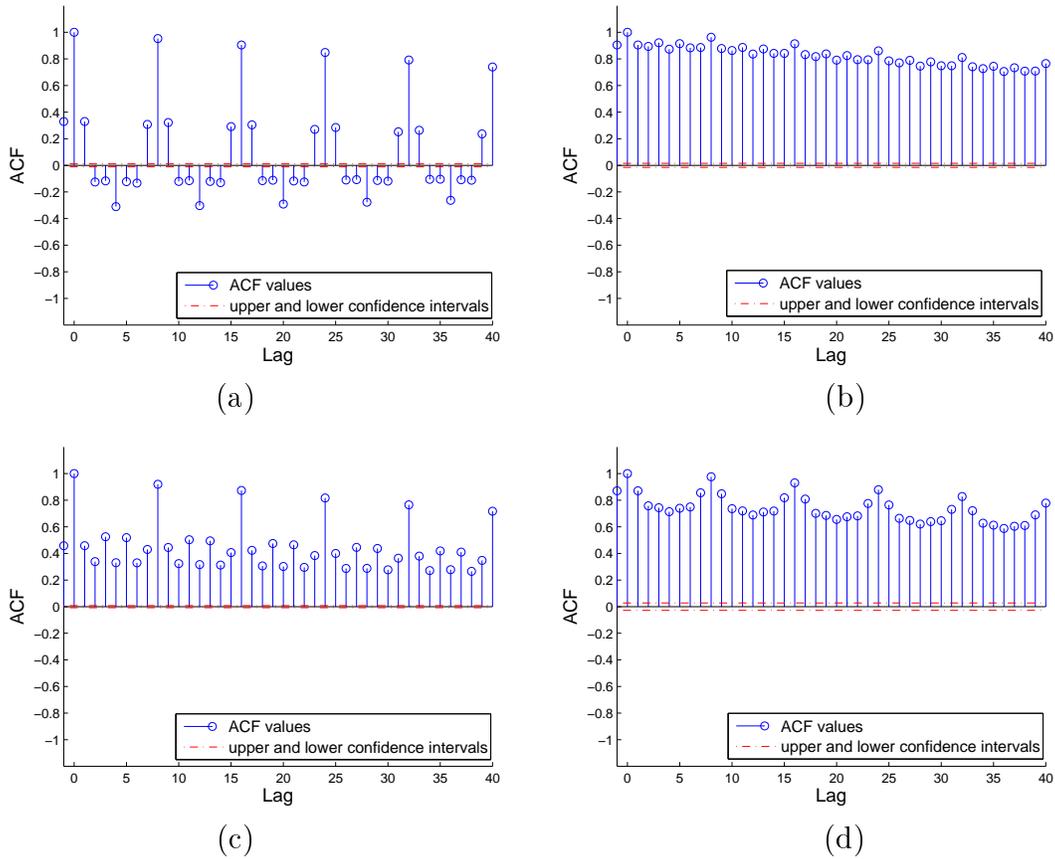


Figure 5.2: (a) ACF of closed-loop scanner output when control is weighted heavily in the machine-direction ($\eta_{acf} = 0.029639$, output variance = 0.5929) (b) ACF of closed-loop scanner output when control is weighted heavily in the cross-direction ($\eta_{acf} = 0.10176$, output variance = 0.31225) (c) ACF of closed-loop scanner output when control is weighted equally in both directions ($\eta_{acf} = 0.054525$, output variance = 0.15809) (d) ACF of open-loop scanner output obtained when the same disturbance used to generate (a), (b) and (c) is applied ($\eta_{acf} = 0.033094$, output variance = 1.8813)

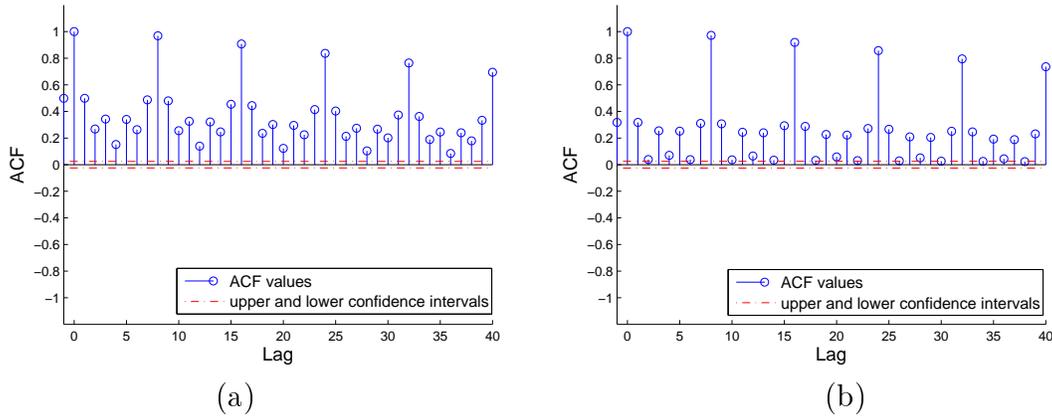


Figure 5.3: (a) ACF of open-loop scanner output obtained when the spatial correlation coefficient $\rho = 0.4$ ($\eta_{acf} = 0.065577$, output variance = 1.4737), and (b) ACF of open-loop scanner output obtained when the spatial correlation coefficient $\rho = 0.1$ ($\eta_{acf} = 0.091782$, output variance = 1.6032)

control is poor relative to CD control. When MD and CD controls are equally good, the ACF demonstrates small, and positive, correlation for all lags with slightly larger peaks at lags that are multiples of the periodicity of the process, such as in Fig. 5.2 (c).

As mentioned earlier, the ACF of the scanning sensor output can also be used to assess the extent of spatial correlation in the disturbance by running the plant under open loop conditions. Fig. 5.2 (d) shows the ACF of the scanning sensor output under open-loop conditions when the same disturbance used to generate the other three ACFs is employed. In this figure, it is observed that the cross-direction correlations are substantial compared to the machine direction correlation, and therefore indicates a substantial amount of spatial correlation. The spatial correlation coefficient used for the simulations in Fig. 5.2 is $\rho = 0.8$. As a comparison, the ACF of the scanning sensor output under open-loop conditions when the spatial correlation coefficient is set to $\rho = 0.4$ is shown in Fig. 5.3 (e), and the case when the spatial correlation coefficient is set to $\rho = 0.1$ is shown in Fig. 5.3 (f). The run length used to generate the ACFs was 10,000 samples in order to reduce statistical variations in the ACF.

Although the ACF of the scanning sensor is a useful tool for distinguishing between MD and CD control, it not a useful tool to access the overall performance of the controller. As shown in the caption of Fig. 5.2, even when the variance of the scanning reduces, the η_{acf} value for the controller does not reduce, and even

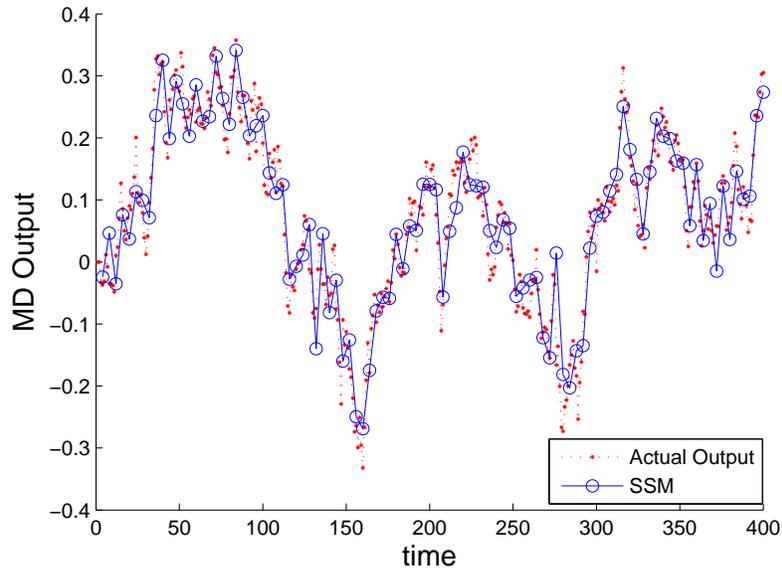
increases! The reason for the lack of dependability on the ACF of the scanning sensor output towards performance assessment is because the scanning sensor presents a periodic combination of many outputs, and can therefore not be treated as a single output, as is required for depending on the η_{acf} index. Alternatively, the variance of the scanning sensor output can be used as a simple measure for over controller performance. In appendix A, it is shown that the variance of the scanning sensor can be interpreted to be equally composed of variance in the MD and CD directions.

5.4 ACF of estimated MD and CD variations

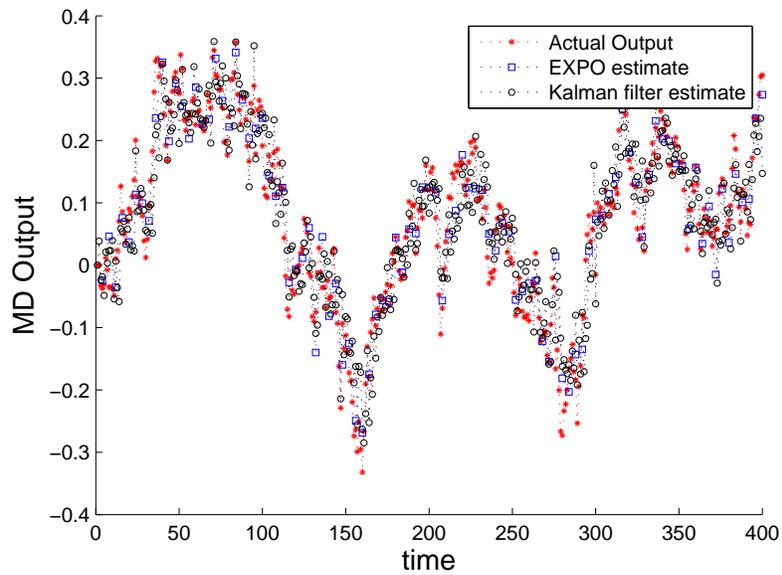
The output of the scanning sensor is typically separated into MD and CD variations for the sake of control. The ACF of these estimated MD and CD variations can be used as a tool for the assessing the performance of the MD and CD controllers, respectively. Prior to discussing the use of the ACF, however, the various techniques for separation of these two components of output variation are discussed and compared below.

5.4.1 Comparison of estimation methods

Three methods are popularly employed for separating MD and CD outputs. The first method is to simply average the output over a scan as an estimate of the MD output and use the deviation of each output from this average as the estimate of the CD output at that location. In this thesis, this method will be called Simple Scan Method (SSM). The second method, described in section 2.7.1, is a modification of the CD output estimation algorithm in SSM. In this method, called exponential multi-scan trending (EXPO), the CD profile is exponentially averaged in order to reduce the effect of MD variations on the CD estimation algorithm. Therefore, identification of CD variations is significantly slower than in SSM. Although the resulting control is sluggish, this method is popularly employed in the industry. The third method, described in section 3.4, employs a time-varying Kalman filter for combined estimation of MD and CD variations.



(a)



(b)

Figure 5.4: Comparison of MD estimation algorithms with actual MD variation data. Estimation algorithms compared: (a) SSM, (b) EXPO and Kalman filter

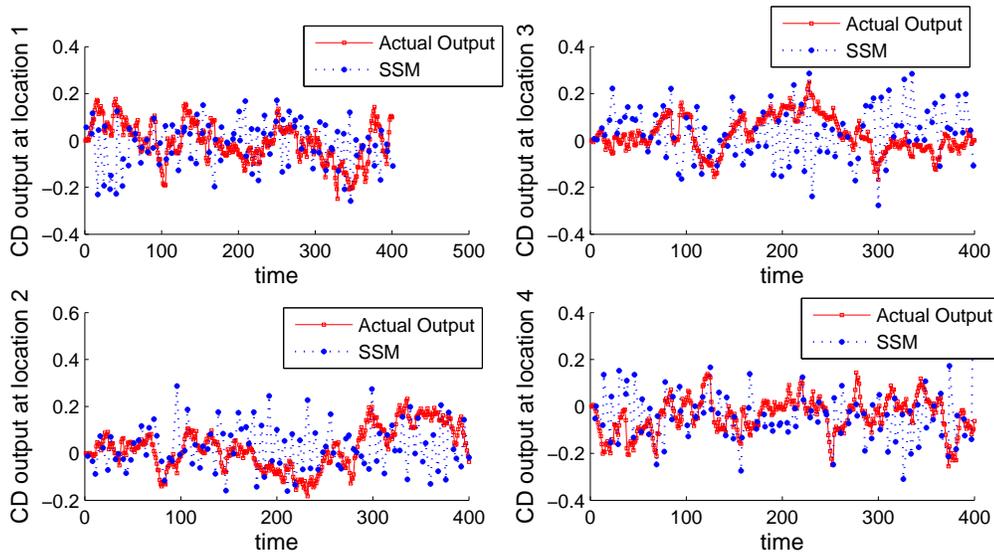
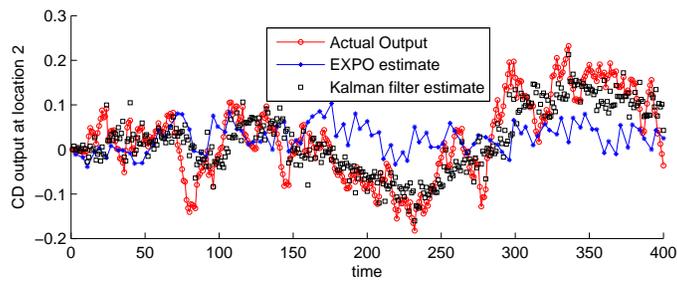
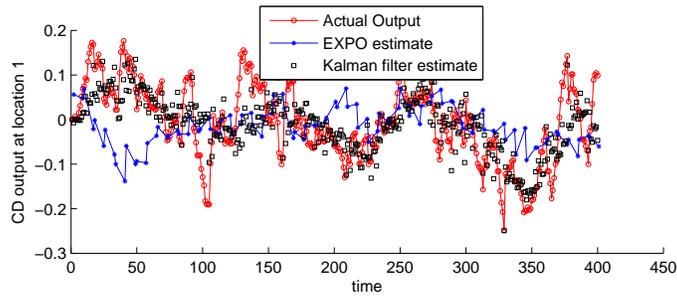
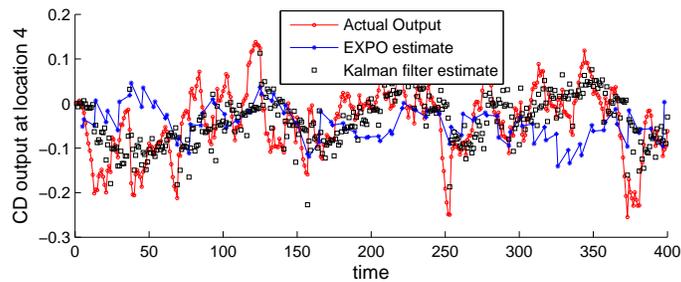
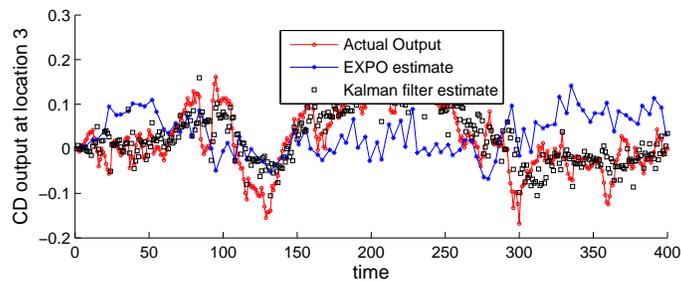


Figure 5.5: Comparison of SSM CD estimation algorithm with actual CD variation data.

Since most industrial processes have slow changes in the CD direction compared to the MD direction (that is ρ is likely to be close to 1), the three algorithms are compared for the case when $\rho = 0.9$ in Figs. 5.4, 5.5 and 5.6. Fig. 5.4 compares the MD estimation algorithms with actual data. All the three algorithms perform satisfactorily, and any of them can be used for performance assessment of MD control. Fig. 5.4 compares the CD estimation algorithm in SSM with actual CD variations. A mismatch is obvious, and the estimated CD variations show a high degree of fluctuations due to corruption from MD variations. The use of CD variations estimated from SSM for performance assessment of CD control is therefore not recommended. Fig. 5.6 compares the EXPO and Kalman filter CD estimation algorithms with actual CD variations. Although the fluctuations caused by MD variations are partially removed in the EXPO algorithm, the resulting estimate is still relatively poor. The Kalman filter algorithm, however, predicts the CD profile quite closely. Hence, it is recommended that only CD variations estimated from the Kalman filter algorithm be used for accessing the performance of CD control. Although this implies that some kind of model is required for estimating CD variations accurately, the actual index for accessing CD performance that is proposed here does not make use of the model and is therefore unaffected by inaccuracies in the model.



(a)



(b)

Figure 5.6: Comparison of EXPO and Kalman filter CD estimation algorithms with actual CD variation data. In EXPO, the historical CD profile is weighted by 0.7 and the current CD output is weighted by 0.3. Location 1 and 2 are shown in (a) while location 3 and 4 are shown in (b).

5.4.2 Accessing MD and CD control

Based on the estimates of the Kalman filter algorithm, the performance of MD and CD control can be accessed using the ACF of the MD and CD variations. Fig. 5.7 (a) shows the ACF of the MD data, and Fig. 5.7 (b) shows the ACF of the CD data from a coupled controller described in section 4.4 with MD and CD weights at $W_{MD} = W_{CD} = 1$, $Q = I$, $r_v = 0.5$ and $r_u = 1.5$. Both ACFs are seen to decay slowly and the η_{acf} values are low, implying that the control is not optimum. This is because the controller used is an LQG controller that does not have integral action, and therefore does not eliminate offsets in the output. The controller settings are changed to $W_{MD} = W_{CD} = 1$, $Q = I$, $r_v = 0.05$ and $r_u = 0.05$ to see the effect of the improved control on the ACFs. The resulting ACFs and η_{acf} values are shown in Fig. 5.8 (a) and 5.8 (b). The ACFs are seen to demonstrate significantly lower values for lags higher than zero, while the η_{acf} values are higher for both MD and CD data, indicating improved control in both directions.

In order to see if the ACF of MD and CD data provide information about the performance of the respective direction only, the simulation is repeated with the controller settings changed to $W_{MD} = W_{CD} = 1$, $Q = I$, $r_v = 0.05$ and $r_u = 1.5$. As explained in section 3.4, the actuator v influences MD control only, and so no improvement is expected in CD control. The actuator u , however, influences both MD and CD control. Since integral action has been employed for actuator v , the expected improvement is marginal. This is confirmed in Fig. 5.9. The η_{acf} values indicate improved MD control with almost no change in CD control. The very minor change in CD η_{acf} values are likely to be caused by minor differences in the estimated CD variations from the Kalman filter.

5.5 Conclusion

In this chapter it is shown that the ACF of the scanning sensor output can be used to identify the relative performance of MD and CD control. Relatively poor performance in the MD direction corresponds to significant correlations for all lags

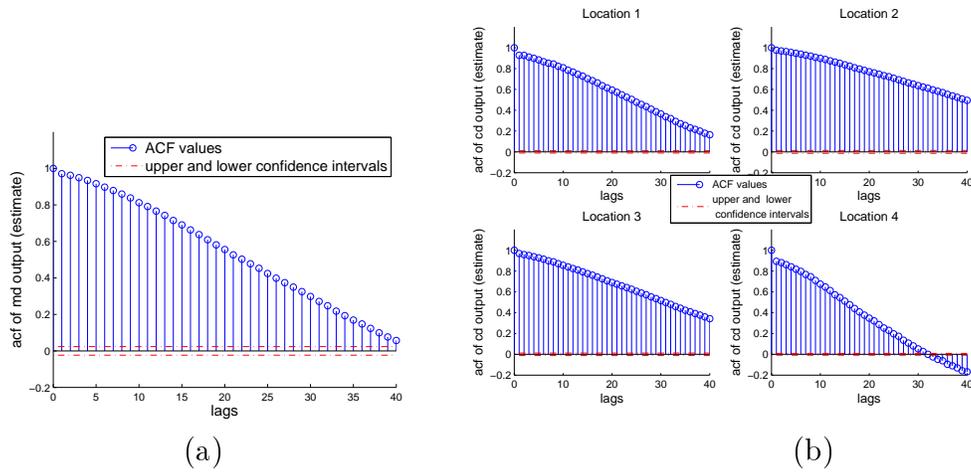


Figure 5.7: ACF of (a) closed-loop estimated MD output ($\eta_{acf} = 0.044439$) and (b) closed-loop estimated CD output ($\eta_{acf, \text{location 1}} = 0.042018$, $\eta_{acf, \text{location 2}} = 0.031967$, $\eta_{acf, \text{location 3}} = 0.035729$, $\eta_{acf, \text{location 4}} = 0.060291$)

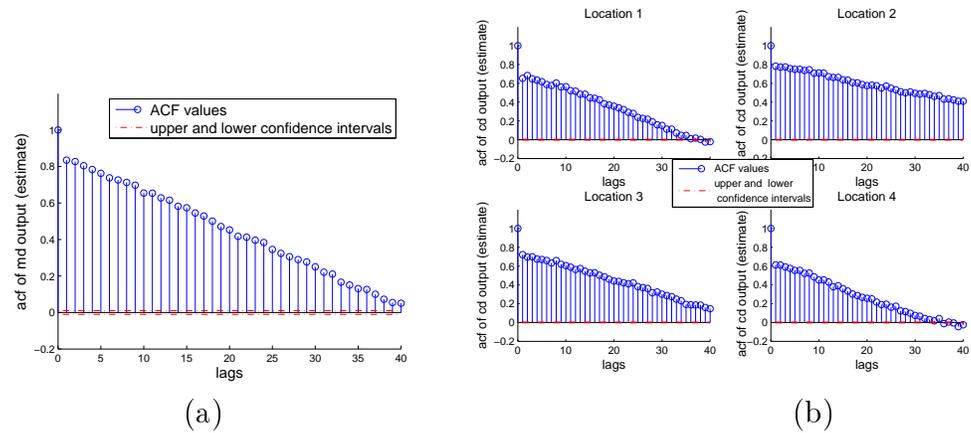


Figure 5.8: ACF of (a) closed-loop estimated MD output ($\eta_{acf} = 0.05325$) and (b) closed-loop estimated CD output ($\eta_{acf, \text{location 1}} = 0.069308$, $\eta_{acf, \text{location 2}} = 0.040495$, $\eta_{acf, \text{location 3}} = 0.053446$, $\eta_{acf, \text{location 4}} = 0.084458$)

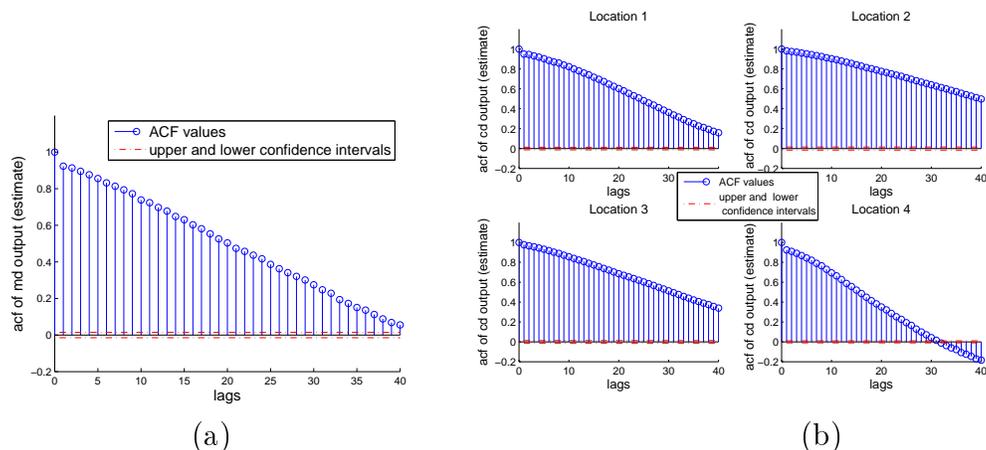


Figure 5.9: ACF of (a) closed-loop estimated MD output ($\eta_{acf} = 0.048107$) and (b) closed-loop estimated CD output ($\eta_{acf, \text{location 1}} = 0.041463$, $\eta_{acf, \text{location 2}} = 0.031762$, $\eta_{acf, \text{location 3}} = 0.03582$, $\eta_{acf, \text{location 4}} = 0.058847$)

except lags that are multiples of the periodicity of the process. Similarly, relatively poor performance in the CD direction corresponds to significant correlation for lags that are multiples of the periodicity of the process, but low correlation for other lags. This information is useful in determining the priority with which controllers may need to be accessed, and in determining the extent of spatial correlation (ρ) in the disturbance. In combination with the variance of the scanning sensor, the ACF of the scanning sensor provides a complete, but approximate, picture of the performance of the controllers.

A comparison of the various techniques for separation of MD and CD variations shows that a Kalman filter is required to access the performance of CD controllers. The performance of MD control, however, may be estimated reasonably well using SSM. From the Kalman filter and SSM estimates, the ACF of the closed-loop output can be used to access the performance of the individual controllers using a performance index that compares the closed loop output with white noise. The resulting indices allow one to monitor changes in controller performance over time and identify the need to retune a controller or reidentify the process model. The techniques could also be used to distinguish between various controller tunings and establish the controller tuning with the best performance.

Chapter 6

Concluding Remarks

From the literature in this area, it is clear that sheet and film forming processes present many interesting and challenging problems for control research. The first problem tackled by researchers was the problem of transforming the measured profile to reduce its dimension to match the dimension of the number of actuators in the machine, in a manner that does not compromise performance. In Gorinevsky et al. (2000), it is shown that the use of a mapping window which matches the actuator response profile is optimum, and results in no loss of performance. The second problem is the problem of estimating and separating machine and cross-machine directional variations optimally from the scanning sensor data. Various types of Kalman filters have been proposed as a solution to this problem (Wang et al. (1993b), Bergh and MacGregor (1987), Chen (1988)). The third problem is the identification of a model suitable for control. While MD models are relatively easy to identify, CD models are more challenging, as illustrated by Featherstone and Braatz (1997). In Featherstone and Braatz (1997), it is shown that incorrectly identifying the sign of the singular values of the interactor matrix could lead to a commonly faced problem in CD control known as picketing, which causes the output profile and actuator setpoints to settle in a ‘jagged’ shape. Therefore, care must be taken to ensure that the sign of the identified singular values are correct. An iterative identification method has been proposed in Featherstone and Braatz (1997) as a way to accurately identify singular values interactor matrix.

The fourth, and perhaps most important, problem is the design of an optimum controller that can handle the constraints, process uncertainties and the large size of the problem. While robust control methodologies that can handle the large size of the problem have been proposed, these algorithms do not handle the constraints

explicitly and their optimality cannot be guaranteed. Optimal control algorithms such as LQG or MPC controllers can handle the constraints as either soft or hard constraints, but are not designed to handle the uncertainties in the process. In any case, using such an optimal control methodology, coupled control of both MD and CD directions has been previously proposed for processes with relatively fast changes in the CD direction. Such a controller has the advantage of overcoming the effect of time-delay in the MD actuators and provides improved estimates of MD and CD variations as their interactions are taken into consideration during the estimation process. For processes with relatively slow changes in the CD direction, however, coupled control has so far not been recommended as it may result in poor CD control compared to MD control. For such processes, it is typical to separate the control into two separate problems, leading to a suboptimal performance. However, these processes are the most common type among industrial sheet and film forming processes.

Chapter 3.3 provides a solution to this aspect of the fourth problem. In this chapter, it is shown that minimizing the variance of the output, which is the objective of coupled control, can be interpreted as minimizing the MD and CD variances with equal weighting. As this is not the desired objective in practice, a coupled control scheme that allows a trade-off between the two control objectives has been proposed. The implementation of the trade-off using a standard LQG framework is also discussed. By exploiting the trade-off between the objectives, coupled control can now be implemented on processes with relatively slow changes in the CD direction.

A fifth problem in sheet and film forming processes is the assessment of controller performance. One of the major challenges here is the time-varying nature of the process created by the nature of the measurement process. As a result of the measurement process, an estimation algorithm is sometimes employed to estimate the unmeasured states, and this can cause an increase in the output variance. The effect of this estimation process has not been considered in previous works on performance assessment of sheet and film forming processes (Taylor and Duncan (2005), Duncan et al. (2000)).

Chapter 4 provides a model-based solution to the fifth problem. In section 4.2, the LQG trade-off curve proposed by Huang and Shah (1999) is extended to time-varying MIMO processes, of which sheet and film forming processes are an

example due to the periodic nature of their measurement equation. The approach assumes a state-space model, but a transfer-function model can also be used by converting it to the state-space model. The LQG controller gain and the Kalman filter gains are allowed to vary at each time-step, as this is typically the case for sheet and film forming processes. The proposed approach is applied to the coupled control scheme proposed in chapter 3 and the trade-off between all three variables (MD, CD and input variance) is illustrated. Simulation results illustrate a good match with theoretical results. In addition to the practicing control engineering, the work in chapter 4 is also beneficial to vendors of scanning or full-scan sensors as the proposed method allows one to access the benefits of implementing additional sensor or altering the scanning pattern

While model based solutions for the performance assessment problem are useful, data-driven approach present an important advantage over model-based approaches in terms ease of implementation, as no *a priori* model is required. Chapter 5 proposes some data driven techniques for the assessment of controller performance for sheet and film forming processes. First, the ACF of the scanning sensor is shown to be a useful tool for qualitatively accessing the relative performance of MD and CD control. This tool has at least two uses: determining the priority with which controllers may need to be accessed, and identifying the extent of spatial correlation in the disturbance (ρ). Following this tool, an index based on the ACF of the estimated MD and CD outputs is proposed to judge the performance of the individual controllers. The proposed index compares the ACF of the closed-loop output with the ACF of white noise, as an output of white noise indicates no further scope for improvement from control. This tool can be used to identify changes in controller performance over a period of time, and help detect the need to retune the controller or reidentify the model. However, a comparison of the methods for separating MD and CD variations indicate that a Kalman filter is required to sufficiently capture information on both MD and CD variations. The use of the index is illustrated with simulation examples.

6.1 Suggestions for Future Research

While many problems pertaining to sheet and film forming processes are addressed in this thesis, there are a few directions in which further research can be carried out. For example, the design of a robust coupled controller which has the ability

to tradeoff between the control objectives is of great importance. Due to the robustness of the controller, integral action can be employed in the CD direction, leading to substantial improvements in CD control. In addition, the fact that the controller is coupled allows for improved MD control. The primary issue with the design of such a controller is that the inclusion of the MD actuator gains in the interactor matrix invariably results in a singular interactor matrix.

For the performance assessment problem, one possible direction for further research is to establish a LTV MIMO minimum variance benchmark for sheet and film forming processes. The development of such a benchmark requires the extension of LTI MIMO minimum variance theory to LTV processes. Such a benchmark is useful in determining the possible benefits of relaxing the physical constraints using process data only.

Bibliography

- Adamy, J., 1997. Adaptation of cross-direction basis-weight control in paper machines using fuzzy decision logic. *International Journal of Approximate Reasoning* 16, 25 – 42.
- Anson, S. J. I., Constantino, R. P. A., Hoole, S. M., Sampson, W. W., 2008. Estimation of the profile of cross-machine shrinkage of paper. *Measurement Science and Technology* 19, 1–11.
- Astrom, K., 1967. Computer control of a paper machine - an application of linear stochastic control theory. *IBM Journal of Research and Development* 11, 389 – 405.
- Baki, H., Wang, H., Soylemez, M., Munro, N., 2001. Implementing machine-directional basis weight control for a pilot paper machine. *Control Engineering Practice* 9, 621 – 630.
- Balderud, J., Wilson, D. I., 2001. Decoupling basis-weight measurements in paper manufacture. In: *Proceedings of the American Control Conference*, Arlington, VA, USA. Vol. 3. pp. 2222–2224.
- Bergh, L. G., MacGregor, J. F., 1987. Spatial control of sheet and film forming processes. *Canadian Journal of Chemical Engineering* 65, 148 – 155.
- Bialkowski, W. L., 1983. Application of kalman filters to the regulation of dead time processes. *IEEE Transactions on Automatic Control* 28, 400 – 406.
- Bishop, A. B., 1963. Techniques for real-time determination of the components of variance for control of continuous-sheet processes. *IEEE Transactions on Automatic Control* 8 (4), 321–326.

- Borisson, V., Wittenmark, B., 1974. An industrial application of a self-tuning regulator. In: 4th IFAC Symposium on Digital Computer Applications to Process Control, Zurich, Switzerland.
- Boyd, S., Barratt, C., 1991. Linear Control Design. Prentice Hall, Englewood Cliffs, New Jersey.
- Boyle, T. J., 1977. Control of cross-machine variations in web forming machines. Canadian Journal of Chemical Engineering 55, 457–461.
- Boyle, T. J., 1978. Practical algorithms for cross-directional control. Tappi 61(1), 77–80.
- Breecher, A. E., Bareiss, R. A., 1970. Theory and practice of automatic control of basis weight profiles. Tappi 53(5), 847–852.
- Brewster, D. B., Bjerring, A. K., 1970. Computer control in pulp and paper 1961-1969. Proceedings of the IEEE 58, 49–69.
- Chang, D. M., Yu, C., Chien, I.-L., 2001. Arrangement of multi-sensor for spatio-temporal systems, application to sheet-forming processes. Chemical Engineering Science 56, 5709 – 5717.
- Chen, S. C., 1988. Kalman filtering applied to sheet measurement. In: Proceedings of the American Control Conference, Piscataway, NJ. pp. 643 – 647.
- Chen, S.-C., Subbarayan, R., 1999. Cross-machine direction (cd) response modeling with two-dimensional sheet variation measurements. In: Proceedings of the American Control Conference, San Diego, CA, USA. pp. 3082–3086.
- Chen, S.-C., Wilhelm Jr., R. G., 1986. Optimal control of cross-machine direction web profile with constraints on the control effort. In: American Control Conference Proceedings.
- Chen, Z., Li, X., Liu, Z., Yuan, Z., 2008. A novel neural-net-based nonlinear adaptive control and application to the cross-direction deviations control of a polymer film spread line. Chaos, Solitons and Fractals 35, 808–813.
- Corscadden, K. W., Duncan, S. R., 2000. Multivariable disturbance modelling for web processes. International Journal of Systems Science 31, 97–106.
- Cutshall, 1990. The nature of paper variation. Tappi Journal 73 (6), 81 – 90.

- Dahlin, E. B., 1970. Computational methods in a dedicated computer system for measurement and control on paper machines. *Tappi* 53 (6), 1100–1105.
- Doyle, J. C., 1987. A review of μ for case studies in robust control. In: IFAC World Congress, Munich, Germany.
- Dumont, G. A., 1986. Application of advanced control methods in the pulp paper industry. *Automatica* 22, 143–153.
- Dumont, G. A., Davies, M., Natarajan, K., 1993. Estimation of moisture variations on paper machines. *IEEE Transactions on Control Systems Technology* 1 (2), 101 – 113.
- Duncan, S. R., 1989. The cross-directional control of web forming processes. Ph.D. thesis, University of London.
- Duncan, S. R., 1994. The design of robust cross-directional control systems for paper making. Control Systems Centre Report 807, UMIST, Manchester, UK.
- Duncan, S. R., Dumont, G. A., Gorinevsky, D. M., 1999. Evaluating the performance of cross-directional control systems. In: Proceedings of the American Control Conference, San Diego, CA, USA.
- Duncan, S. R., Dumont, G. A., Gorinevsky, D. M., 2000. Performance monitoring for cross-directional control systems. In: Proceedings of the PAPTAC Control Systems Conference, Victoria, BC, Canada. pp. 173 – 179.
- Fan, J., 2003. Model predictive control for multiple cross-directional processes: Analysis, tuning, and implementation. Ph.D. thesis, The University of British Columbia.
- Featherstone, A. P., Braatz, R. D., 1997. Control-oriented modeling of sheet and film processes. *Automatica* 43, 1989–2001.
- Flickr, September 2009. Papermaking machine at a paper mill near pensacola. URL <http://www.flickr.com/photos/floridamemory/3325350513/>
- Gacon, D. M., Zarrop, M. B., Kulhavy, R., 1996. Em estimation and control of two-dimensional systems. UKACC International Conference on Control '96, Sept., 2–5.

- Gorinevsky, D., Vyse, R., Heaven, M., 2000. Performance analysis of cross-direction process control using multivariable and spectral models. *IEEE Transactions on Control Systems Technology* 8 (4), 589 – 600.
- Gorinevsky, D. M., Gheorghe, C., 2003. Identification tools for cross directional processes. *IEEE Transactions on Control Systems Technology* 11 (5), 629–640.
- Harris, T. J., 1989. Assessment of closed loop performance. *Canadian Journal of Chemical Engineering* 67, 856–861.
- Heath, W. P., 1993. Orthogonal functions for cross-directional control of paper making machines and plastic film extruders. Control Systems Centre Report, UMIST.
- Heath, W. P., 1996. Orthogonal functions for cross-directional control of web forming processes. *Automatica* 32, 183 – 198.
- Heaven, E., Kean, T., Jonsson, I., Manness, M., Vu, K., Vyse, R. N., 1993. Applications of system identification to paper machine model development and controller design. In: *Second IEEE Conference on Control Applications*, Vancouver, BC, Canada. Vol. September.
- Heaven, E. M., Jonsson, I. M., Kean, T. M., Manness, M., Vyse, R. N., 1994. Recent advances in cross machine profile control. *IEEE Control Systems* October, 35–46.
- Holik, H., 2006. *Handbook of Paper and Board*. Wiley.
- Huang, B., 2002. Minimum variance control and performance assessment of time-variant processes. *Journal of Process Control* 12, 707–719.
- Huang, B., Shah, S., 1999. *Performance Assessment and Control of Multivariable Processes*. Springer.
- Ismail, A., Dumont, G. A., 2003. Dual adaptive control of paper coating. *IEEE Transactions on Control Systems Technology* 11, 289–308.
- Jelali, M., 2006. An overview of control performance assessment technology and industrial applications. *Control Engineering Practice* 14, 441–466.

- Kjaer, A. P., Heath, W. P., Wellstead, P. E., 1995. Identification of cross-directional behaviour in web production: Techniques and experience. *Control Engineering Practice* 3 (1), 21–29.
- Kristinsson, K., Dumont, G. A., 1993. Paper machine cross directional basis weight control using gram polynomials. In: *Second IEEE Conference on Control Applications*, Vancouver, BC, Canada.
- Kwakernaak, H., Sivan, R., 1972. *Linear Optimal Control Systems*. Wiley-Interscience.
- Laughlin, D. L., Jordan, K., Morari, M., 1986. Internal model control and process uncertainty: mapping uncertainty regions for siso controller design. *International Journal of Control* 44, 1675 – 1698.
- Laughlin, D. L., Morari, M., Braatz, R. D., 1993. Robust performance of cross-directional basis-weight control in paper machines. *Automatica* 29 (6), pp. 1395–1410.
- Maenpaa, T., 2006. Robust model predictive control for cross-directional processes. Ph.D. thesis, Helsinki University of Technology.
- Norbury Jr, M. A., 1996. Measurement of instantaneous profile: a new gauging system for hot and cold rolling. *Nondestructive Testing & Evaluation* 12, 263–272.
- Rawlings, J. B., Chien, I.-L., 1996. Gage control of film and sheet-forming processes. *AIChE Journal* 42, 753–766.
- Ringwood, J. V., Grimble, M. J., 1990. Shape control in sendzimir mills using crown and intermediate roll actuators. *IEEE Transactions of Automatic Control* 35 (4), 453 – 459.
- Rossiter, J. A., Kouvaritakis, B., 1993. Constrained stable generalized predictive control. *IEEE Proceedings - D, Control Theory and Applications* 140, 243 – 254.
- Skogestad, S., 2003. Simple analytic rules for model reduction and pid controller tuning. *Journal of Process Control* 13, 291 – 309.

- Skogestad, S., Morari, M., Doyle, J. C., 1988. Robust control of ill-conditioned plants: high purity distillation. *IEEE Transactions of Automatic Control* 33, 1092–1105.
- Smook, G. A., 1992. *Handbook for Pulp & Paper technologists*. Angus Wilde Publications, Bellingham, Washington.
- Stewart, G. E., 2000. Two dimensional loop shaping controller design for paper machine cross directional process. Ph.D. thesis, The University of British Columbia.
- Stewart, G. E., Gorinevsky, D. M., Dumont, G. A., 1998. Robust gmvcross directional control of paper machines. In: *Proceedings of the American Control Conference*, Philadelphia, PA, USA.
- Taylor, A., Duncan, S., 2005. Bayesian methods for control loop performance assessment in cross-directional control. In: *16th IFAC Triennial World Congress*, Prague, Czech Republic. pp. 434–439.
- Thake, A. J., 1997. Design of filter based controllers for cd control on paper machines. Master's thesis, University of Toronto.
- Tyler, M. L., Morari, M., 1995. Estimation of cross directional properties: Scanning versus stationary sensors. *AIChE Journal* 41, 846–854.
- VanAntwerp, J. G., Braatz, R. D., 2000. Fast model predictive control of sheet and film processes. *IEEE Transactions on Control Systems Technology* 8, 408–417.
- VanAntwerp, J. G., Featherstone, A. P., Braatz, R. D., 2000. *Identification and control of sheet and film forming processes*. Springer.
- VanAntwerp, J. G., Featherstone, A. P., Braatz, R. D., 2001. Robust cross-directional control of large scale sheet and film processes. *Journal of Process Control* 11, 149 – 177.
- VanAntwerp, J. G., Featherstone, A. P., Braatz, R. D., Ogunnaike, B., 2007. Cross-directional control of sheet and film processes. *Automatica* 43, 191 – 211.
- Vyse, R., King, J., Heaven, M., Pantaleo, S., 1996. Consistency profiling - a new technique for cd basis weight control. *Pulp & Paper Canada* 97, 62–66.

- Wang, X. G., Dumont, G. A., Davies, M. S., 1993a. Adaptive basis weight control in paper machines. In: Second IEEE Conference on Control Applications, Vancouver, BC, Canada.
- Wang, X. G., Dumont, G. A., Davies, M. S., 1993b. Estimation in paper machine control. *IEEE Control Systems* 13 (4), 34–43.
- Wellstead, P. E., Heath, W. P., Kjaer, A. P., 1998. Identification and control of web processes: polymer film extrusion. *Control Engineering Practice* 6, 321–331.
- Wellstead, P. E., Zarrop, M. B., Duncan, S. R., 2000. Signal processing and control paradigms for industrial web and sheet manufacturing. *International Journal of Adaptive Control and Signal Processing* 14, 51–76.
- Wellstead, P. E., Zarrop, M. B., Heath, W. P., Kjaer, A. P., Troyas, X., 1996. Two dimensional methods for machine direction and cross direction estimation and control. In: Control systems 96 preprints, Montreal, QC, Canada. Vol. CPPA. pp. 5–8.
- Wilson, D. I., Balderud, J., 2000. Experiences building and validating a large-scale dynamic paper machine model. *Model Validation for Plant Control and Condition Monitoring*, IEE Seminar on.
- Zheng, A., 1999. Reducing on-line computational demands in model predictive control by approximating qp constraints. *Journal of Process Control* 9, 279 – 290.

Appendix A

Relationship between scanning sensor variance and MD variance and CD variance

The scanning sensor variance is defined as:

$$\gamma = \sum_{k=1}^T \frac{(y(k) - \bar{y})^T (y(k) - \bar{y})}{mT - 1} \quad (\text{A.1})$$

where $y(k)$ is the vector of measurements from the k^{th} scan with elements $y(j, k)$ and

$$\bar{y} = \underbrace{\left[y_{avg}, y_{avg}, \dots, y_{avg} \right]}_m, \text{ where } y_{avg} = \sum_{k=1}^T \sum_{j=1}^m \frac{y(j, k)}{m.T}$$

Therefore, equation A.1 is equivalent to:

$$\begin{aligned} \gamma &= \sum_{k=1}^T \sum_{j=1}^m \frac{(y(j, k) - y_{avg})^2}{mT - 1} \\ &= \sum_{k=1}^T \sum_{j=1}^m \frac{(y(j, k) - \bar{y}(k) + \bar{y}(k) - y_{avg})^2}{mT - 1} \end{aligned} \quad (\text{A.2})$$

where $\bar{y}(k) = \frac{1}{m} \sum_{j=1}^m y(j, k)$ is the average output over the k^{th} scan, and can be treated as the MD output of a scan, or MD_{scan} . Note that the overall average, y_{avg} is simply the average of MD_{scan} over T scans: $y_{avg} = \sum_{k=1}^T \frac{\bar{y}(k)}{T}$. Rearranging

the definition of γ gives:

$$\gamma = \sum_{k=1}^T \sum_{j=1}^m \frac{(y(j, k) - y_{avg})^2}{mT - 1} \quad (\text{A.3})$$

$$= \frac{1}{mT - 1} \sum_{k=1}^T \sum_{j=1}^m [(y(j, k) - \bar{y}(k))^2 + (\bar{y}(k) - y_{avg})^2] \quad (\text{A.4})$$

$$+ 2(y(j, k) - \bar{y}(k))(\bar{y}(k) - y_{avg})^2] \quad (\text{A.5})$$

$$= \frac{1}{mT - 1} \sum_{k=1}^T \sum_{j=1}^m [(y(j, k) - \bar{y}(k))^2 + (\bar{y}(k) - y_{avg})^2] \quad (\text{A.6})$$

$$= \frac{1}{mT - 1} \sum_{k=1}^T \sum_{j=1}^m (y(j, k) - \bar{y}(k))^2 + \frac{1}{mT - 1} \sum_{k=1}^T \sum_{j=1}^m (\bar{y}(k) - y_{avg})^2 \quad (\text{A.7})$$

$$= \text{Var}(CD_{scan}) + \text{Var}(MD_{scan}) \quad (\text{A.8})$$

where $\text{Var}(CD_{scan})$ is the variance of the CD output of the k^{th} scan, and is defined as: $\frac{1}{mT-1} \sum_{k=1}^T \sum_{j=1}^m (y(j, k) - \bar{y}(k))^2$. The definition of $\bar{y}(k)$ is used to eliminate the third term in equation A.4. Hence, γ can be thought of as the sum of CD and MD variances taken over a scan.

Assuming that the CD variance calculated over a scan is representative of the steady-state CD variances taken over a single time instant ($\text{Var}(CD)$), and similarly assuming that the MD variance calculated over a scan is representative of the steady-state MD variances taken over a single time instant ($\text{Var}(MD)$), we may say

$$\gamma = \text{Var}(CD) + \text{Var}(MD)$$

That is, the scanning sensor variance is the sum of steady-state MD and CD variances.