

A Vision for MPC Performance Maintenance

Mohammed Tajudeen Jimoh

A THESIS SUBMITTED TO THE COLLEGE OF SCIENCE AND
ENGINEERING, UNIVERSITY OF GLASGOW, FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

©MOHAMMMED JIMOH, NOVEMBER 2013, GLASGOW, SCOTLAND

Acknowledgement

Foremost, I express my profound gratitude to my supervisor, Dr John Howell, whose exceptional knowledge and wisdom serve as my guide throughout the studies.

I must thank Mrs Howell for making myself and my family feel completely at home throughout our stay in Scotland. Mr and Mrs McKenzie, and their son, Mr Bruce treated us like close family members throughout. May God bless all of you.

I thank my main sponsor for this programme, the Petroleum Technology Development Fund (PTDF), Nigeria. The special role played by Engr Muttaqha Rabe Darma, a former Executive Secretary of PTDF, can never be over emphasised.

I am grateful to Bayero University, Kano, for the award of the study fellowship, which contributed immensely to the success of this programme.

To all my colleagues at Bayero University and my friends in Nigeria, I thank you for your supports. Prof A. R. Mohammed, Prof J. S. Enaburekhan, Prof I. S. Diso, Dr A. B. Aliyu, Dr A. B. Ahmed, Dr Aminu Aliyu, Dr Tijanni Darma, Dr N. A. Ademoh, Dr I. A. Yola, Engr Yakubu Momoh, Mall Saidu Abdullahi, Dr Ado Dan-Isa, Dr Akande, Mall Usman Jose, Dr Ibrahim Siraj. I thank you all.

Special thanks to Azeez Oyedele, a friend whose exceptional maturity ensured that I felt very comfortable as he hosted me during the period of the write up.

Thanks to my parents and my brothers and sisters, especially Taju, for their support.

The supports of my wife, Falilat and my children (Nawal, Shafiq, Sabira and Habiba) throughout the programme were invaluable. I am lucky to have all of you in my life.

Dedication

To my wife, Falilat, and my four beautiful children, Nawal, Shafiq, Sabira and Habiba.

Table of Contents

Acknowledgement	i
Dedication	ii
Table of Contents	iii
List of Tables	viii
List of Figures	x
Abstract	xiv
Chapter One	1
Introduction	1
1.1 Background.....	1
1.2 Spirit behind the maintenance tool.....	3
1.3 MPC in the Process Industries.....	6
1.4 Role of Control Performance Assessment and Faults Detection and Diagnosis.....	8
1.5 Outline of the Maintenance Tool.....	9
1.6 Methodology.....	11
1.7 Novel Aspects of this Thesis.....	12
1.8 Summary.....	13
Chapter Two	14
Review of Linear MPC	14
2.1 MPC Principles and Fundamentals.....	14
2.2 Development of MPC.....	16
2.3 MPC Internal Models.....	22
2.3.1 Finite Impulse/Step Response MPC.....	25
2.3.1 Transfer Function MPC.....	25
2.3.2 State-Space MPC.....	27
2.4 MPC Input Objective function.....	28
2.5: Algorithms for MPC Predicted outputs.....	30
2.5.1 Prediction Algorithm for Step Response MPC.....	31
2.5.2 Prediction Algorithm for Transfer Function MPC.....	33
2.5.3 Prediction Algorithm for State Space MPC.....	36
2.6 Solution of Unconstrained MPC.....	39

2.7 MPC and Constraints.....	40
2.8 Numeric Solution of MPC QP Problem: Hildreth's QP Procedure	43
2.9 MPC and Economic Optimisation.....	45
2.10 The MPC and the operators in a Process Plant.....	45
2.11: The trends of a control system	46
2.12 Developments in MPC Performance Assessment and Maintenance.....	48
2.13 Process Model Identification	52
2.14 Conversion of Step Responses to State Space Models.....	57
2.15 Relative gain array.....	60
2.16 Summary.....	61
Chapter Three	63
MPC Models and MPC Control of Selected Nonlinear Processes	63
3.1 MPC Model Representations.....	63
3.2 The Nonlinear Processes Selected for the Case Studies.....	66
3.3 Simulation and control of the CSTR	67
3.3.1 Direct MPC on the CSTR	71
3.3.2 MPC as Supervisory Controller for the CSTR	73
3.3.3 Comparison of direct and supervisory MPC control of the CSTR.....	81
3.4 The Evaporator Process	82
3.4.1: Open Loop Simulation of the Evaporator.....	85
3.4.2 Direct MPC	90
3.4.3 MPC in a supervisory capacity	92
3.3 The FCCU model and properties.....	98
3.3.1 FCCU Simulation	101
3.3.2 FCCU system identification.....	105
3.3.3 MPC Control of the FCCU	107
3.4 Summary.....	113
Chapter Four.....	116
Common MPC Faults and their Isolation.....	116
4.1 What-if Simulations of Common MPC Faults: An Overview	116
4.2 The Relative Weight Array	118
4.3 Case 1: Example Relating to MPC Parameter Tuning	120
4.3.1 The Scene.....	120
4.3.2 The operators' perspective.....	121

4.3.3 Reasoning about the scene	124
4.3.4. Case 1 Conclusion	128
4.4 Case 2: Example relating to MPC design.....	128
4.4.1 The Scene.....	129
4.4.2 The operator perspective.....	129
4.4.3 Reasoning about the scene	131
4.4.4 Possible outcomes: experts perspective	133
4.4.5 Case 2 Conclusion	137
4.5 Case 3: Example relating to sensor/actuator degradation.....	137
4.5.1 The Scene.....	138
4.5.2 The operators Perspective	138
4.5.3 Reasoning about the scene	141
4.5.4 Case 3 Conclusion	142
4.6 Case 4: Example relating to MPC constraints	142
4.6.1 The Scene.....	143
4.6.2 The operators Perspective	144
4.6.3 Reasoning about the scene	144
4.6.4 Case 4 Conclusion	146
4.7 Case 5: Example Relating to Variables Selection	147
4.7.1 The Scene.....	147
4.7.2 The Operator's Perspective.....	150
4.7.3 Reasoning about the scene	150
4.7.4. Case 5 Conclusion	152
4.8 Case 6: Example relating to absence of output constraints in zone control	153
4.8.1 The Scene.....	153
4.8.2 The Operators' Perspective.....	155
4.8.3 Reasoning about the scene	155
4.8.4. Case 6 Conclusion	157
4.9 Case 7: Example relating to model plant mismatch	157
4.9.1 The Scene.....	157
4.9.2 The Operators' Perspective.....	158
4.1.3 Reasoning about the scene	159
4.9.4. Case 7 Conclusion	162
4.10 Case 8: Example relating to PID degradation	162

4.10.1 The Scene.....	163
4.10.2 The operators’ perspective.....	164
4.10.3 Reasoning about the scene.....	165
4.10.4. Case 8 Conclusion.....	167
4.11 Summary.....	168
Chapter Five.....	169
The Maintenance Tool Development.....	169
5.1 Perceptions of the Maintenance Tool and the operators.....	169
5.2 Proposed Scope and Components of the Maintenance tool.....	170
5.3 Trends Comparison Assessment Group.....	171
5.3.1 Reference Graphical Performance Window.....	173
5.3.2 Actual Graphical Performance Window.....	177
5.4 Preliminary diagnostic questions window.....	178
5.5 Suspected Faults Window.....	180
5.6 The Background Information Group.....	182
5.6.1 Virtual plant Without MPC Window.....	182
5.6.2 Transfer Function Matrix.....	184
5.6.3 Steady state gain, RGA and RWA Window.....	185
5.6.4 Virtual plant With MPC Window.....	186
5.7 MPC Investigation.....	187
5.7.1 Investigating inappropriate variables selection.....	188
5.7.2 Investigating model/plant mismatch.....	189
5.7.3 Investigating improper constraints specifications.....	189
5.7.4 Investigating PID degradation.....	190
5.7.5 Investigating actuator degradation.....	191
5.7.6 Investigating Poor MPC tuning.....	191
5.7.7 Investigating Poor MPC design.....	193
5.8 Summary.....	193
Chapter Six.....	194
Conclusions and Recommendations.....	194
6.1 Progress to date.....	194
6.2. Critique.....	199
6.3 Suggestions for further work.....	200
References.....	202

Appendix A.....	212
Low Order Approximation of Processes by Direct Method.....	212
A1: The Direct Method	212
A1-1: First order transfer function approximation	213
A1-2 Second order transfer function approximation	215
A1-3 Second order with inverse response transfer function approximation	219

List of Tables

Table 2.1: Summary of major MPC products	23
Table 3.1: CSTR Input variables	68
Table 3.2: CSTR Output Variables	68
Table 3.3: CSTR Parameters	68
Table 3.4: MPC parameters for control of open loop CSTR	72
Table 3.5: PI settings for the stabilizing controller	74
Table 3.6: Actual steady state gain values of closed loop CSTR	74
Table 3.7: Transfer function matrix for CSTR	76
Table 3.8: MPC simulation parameters for CSTR control	77
Table 3.9: Input variables of the evaporator and the equilibrium values	84
Table 3.10: Output variables of the evaporator and the equilibrium values	85
Table 3.11 Evaporator parameters	85
Table 3.12: Steady state MV-CV gains of the evaporator in open loop	87
Table 3.13: Approximate transfer function matrix for open loop evaporator	88
Table 3.14: Steady state MV-CV gains of the linear model of the evaporator open loop	88
Table 3.15: MPC parameters for direct MPC control of evaporator	90
Table 3.16 Steady state gains of the evaporator MV-CV loops for closed loop	93
Table 3.17: Reduced transfer function matrix for evaporator under regulatory control.	94
Table 3.18: Steady state gains for the state space model	94
Table 3.19: Parameters of the regulatory controllers	102
Table 3.20: Manipulated input variables	102
Table 3.21: Disturbance input variables	102
Table 3.22: Output variables	103
Table 3.23: Initial values of state variables	104
Table 3.24: Equilibrium valve openings	105
Table 3.25: Steady gain values for the FCCU	106
Table 3.26: Identified Transfer Function Matrix for the FCCU	108
Table 3.27: Constraints on outputs (soft, operating constraints)	110
Table 3.28: Constraints on inputs (hard, equipment constraints)	110

Table 4.1 Relative gain array of the CSTR.....	124
Table 4.2: Relative weight array of the CSTR.....	125
Table 4.3: RWA for open loop evaporator	131
Table 4.4 RGA for open loop evaporator	133
Table 4.5 RGA for closed loop evaporator.....	134
Table 4.6: MPC parameters for case 5.....	147
Table 4.7: MPC parameters for case 8.....	163
Table 5.1: Preliminary diagnostic questions.....	178
Table 5.2:Lists of causes associated with each underlying abnormality	181
Table 5.3: Approximate transfer function matrix for process A.....	185

List of Figures

Figure 2.1: Basic Concept of Model Predictive Control	15
Figure 2.2: Signal Flow of a MPC	16
Figure 2.3: Typical process plant control hierarchy	47
Figure 2.4: Main variables of MPC as supervisory control	47
Figure 2.5: Excitation signal on open loop plant for identification	53
Figure 2.6: Excitation signal on closed loop plant for identification.....	56
Figure 3.1: Flow diagram of the CSTR unit (open loop).....	67
Figure 3.2: Open loop simulation of the CSTR.....	69
Figure 3.3: Step response plots of the open loop model of the CSTR.....	71
Figure 3.4: Trends of the controlled variables for direct MPC on unstable CSTR	72
Figure 3.5: Trends of the manipulated variables for direct MPC on unstable CSTR.....	73
Figure 3.6: CSTR with two PI controllers	74
Figure 3.7: Step response plot of the closed loop CSTR	75
Figure 3.8: Step response plots of the close loop CSTR models.....	77
Figure 3.9: Controlled variable trends MPC as supervisory control on the CSTR	79
Figure 3.10: Manipulated variable (MPC output) trends MPC as supervisory control on the CSTR	79
Figure 3.11: Manipulated variable (PI output) trends MPC as supervisory control on the CSTR	80
Figure 3.12: Controlled variable comparison: direct vs supervisory	81
Figure 3.13: Manipulated variable comparison: direct vs supervisory.....	82
Figure 3.14: The Evaporator	83
Figure 3.15 Plant input controlled by a local servo actuated valve	86
Figure 3.16: Step response of nonlinear plant to unit step changes in MVs in open loop	86
Figure 3.17: Step response plots of the models of the closed loop evaporator.....	89
Figure 3.18: CV trends of direct MPC controlled nonlinear evaporator step with changes in set points	91
Figure 3.19 MV trends of direct MPC controlled nonlinear evaporator with step changes in set points	91

Figure 3.20: Step response under regulatory control	92
Figure 3.21: Step response plots for the models for the closed loop evaporator	94
Figure 3.22: Trends of the controlled variables under supervisory MPC.....	95
Figure 3.23: Trends of the manipulated variables (MPC outputs) under supervisory MPC.....	96
Figure 3.24 PI outputs under supervisory MPC control	96
Figure 3.25: Controlled variables trends under direct control and supervisory MPC	97
Figure 3.26 Manipulated variables trends under direct control and supervisory MPC ..	98
Figure 3.27: Flow diagram of the FCCU model	99
Figure 3.28: Output trends of the FCCU under regulatory control	104
Figure 3.29: Input trends of the FCCU under regulatory control	105
Figure 3.30: Step response plots for the FCCU	109
Figure 3.31: Plot of the first 30 singular values of the Hankel matrix for the FCCU step response model	112
Figure 3.32: Comparison of reduced model response with the original response	114
Figure 3.33: Trends of the controlled variables and their nominal values under MPC	115
Figure 3.34: Trends of the manipulated variables (MPC outputs) under MPC	115
Figure 4.1: Controlled variables of the CSTR showing degradation.....	121
Figure 4.2: CSTR manipulated variables: MPC outputs	122
Figure 4.3: CSTR manipulated variables: PI outputs	122
Figure 4.4 Controlled variables of the virtual plant showing degradation	123
Figure 4.5 Controlled variables of the CSTR after recovery	126
Figure 4.6: Manipulated variables of the CSTR (MPC outputs) after recovery	126
Figure 4.7 Manipulated variables of the CSTR (PI outputs) after recovery.....	127
Figure 4.8: Controlled variables of the evaporator showing degradation.....	130
Figure 4.9: Manipulated variables of the evaporator	130
Figure 4.10 Controlled variables of the virtual plant showing degradation	131
Figure 4.11 Manipulated variables of the virtual plant.....	132
Figure 4.12: Closed loop evaporator showing improved MPC performance	135
Figure 4.13: Controlled variables of evaporator showing improved MPC performance in closed loop.....	135
Figure 4.14 Manipulated variables of evaporator showing improved MPC performance in closed loop.....	136
Figure 4.15: Controlled variables of the evaporator showing oscillations	138

Figure 4.16 Manipulated variables of the evaporator showing oscillations	139
Figure 4.17: Limit cycle plots for valve on F2	140
Figure 4.18: Limit cycle plots for valve on F200	140
Figure 4.19: Limit cycle plots for valve on P100	141
Figure 4.20: Trends of evaporator controlled variables showing degradation	143
Figure 4.21: Trends of evaporator manipulated variables	144
Figure 4.22: Trends of evaporator manipulated variables after constraint is relaxed...	145
Figure 4.23: Controlled output trends with set-points on Tr and C _{O₂,sg} no degradation	148
Figure 4.24: Manipulated inputs trends with set-points on Tr and C _{O₂,sg} , no degradation	148
Figure 4.25: Controlled outputs trends with set-points on Tr and C _{O₂,sg} , with degradation	149
Figure 4.26: Inputs trends with set-points on Tr and C _{O₂,sg} with degradation	149
Figure 4.27: Controlled outputs trends from virtual plant with set-points on Tr and C _{O₂,sg} , and with ramp increase in T1	150
Figure 4.28: Controlled outputs from the virtual plant when set-point on Tr only, and with ramp increase in T1	151
Figure 4.29: Controlled output trends with set-points on Tr and T3, no output constraints, no disturbances	154
Figure 4.30: Manipulated input trends with set-points on Tr and T3 , no output constraints, no disturbances	154
Figure 4.31: Controlled outputs trends with set-points on Tr and T3 , no output constraints, but with input disturbance	155
Figure 4.32: Controlled output trends with set-points on Tr and T3 , with output constraints and input disturbance	156
Figure 4.33: Case 7 CV trends	158
Figure 4.34: Case 7 MVs	159
Figure 4.35: Responses of plant and model due to simultaneous application of step signals showing model-plant mismatch.....	160
Figure 4.36 Responses of plant and model due to application of step signal to V_{set} only	161
Figure 4.37 Responses of plant and model due to application of step signal to T_{jset} only	161

Figure 4.38: Trends of the evaporator controlled variables due to MPC degradation..	163
Figure 4.39: Trends of the evaporator manipulated variables due to MPC degradation	164
Figure 4.40: Controlled variable trends of evaporator with regulatory control only....	165
Figure 4.41: sp and pv for regulatory controllers	166
Figure 4.42: sp and pv plots for the regulatory controllers.....	166
Figure 5.1: Structure of the Maintenance tool	172
Figure 5.2 Arrangement of windows in the trends comparison assessment group.....	173
Figure 5.3: Sample rv and cv trends	174
Figure 5.4: Sample sv (MPC manipulated variable) trends.....	174
Figure 5.5: Sample mv against ov plot	175
Figure 5.6: Sample rvo and cv _o trends (open loop)	176
Figure 5.7: Sample mv _o against ov _o plot (open loop).....	177
Figure 5.8: Virtual plant without MPC setup	183
Figure 5.9: Sample step response plots from Virtual plant without MPC.....	183
Figure 5.10: Step response plots of Process B.....	184
Figure 5.11: Display for steady state gains, RGA and RWA	185
Figure 5.12: Virtual Plant with MPC setup	187
Figure 5.13: Plot of mv against ov for common valve problems:	192

Abstract

Model predictive control (MPC) is an advanced control that has found widespread use in industries, particularly in process industries like oil refining and petrochemicals. Although the basic premise behind MPC is easy to comprehend, its inner workings are still generally viewed or regarded as too advanced for actual plant operator understanding. This lack of understanding is exposed when MPC performance deteriorates sometime after commissioning, as is often the case in some commercially operated process plants. Currently operators might have difficulty over reasoning about MPC performance degradation and formulating steps to investigate the cause.

A tool is described that aims to make MPC more transparent to the operators. Commonly reported causes of MPC performance degradation are discussed and ways in which the operator can recognise them when they occur are outlined. Issues that are addressed include: making the set of controlled variables to be used for a given set of manipulated variables simpler and clearer; ways to recognise when a MPC controller is performing poorly and to identify the source of performance deterioration. An aim is to determine under what instances the operator can return the MPC performance to previous levels or request for specialist support or simply switch the MPC off. A goal is to avoid the kind of often reported situation where the operator gets worried that the controller is deteriorating and ends up taking knee jerk actions that cause further problems in MPC.

At the top of the maintenance tool hierarchy is the trends comparison group, where MPC reference graphical performance trends are compared with actual graphical performance trends counterpart. If any abnormality is observed in these trends, the

operator is encouraged to choose an option from a list of preliminary diagnostic questions contained in a group below trends comparison group, which best describes the observed abnormality. Each abnormality is associated with a list of suspected causes. When a suspected cause is chosen from the associated list, the operator is led to the symptoms investigation window, which contains scripts detailing steps for systematic examination of each symptom, with a view to either rejecting or confirming the suspicion. Assisted in the investigation are four background information windows: the virtual plant without MPC window, the virtual plant with MPC window, the transfer function matrix window and steady state gain, relative gain array (RGA) and relative weight array (RWA) window. The windows contain information and guidance that the operator might refer to from time to time during symptom investigation.

Development of the maintenance tool is still at the design stage. The key components described have been research implementing MPC on three nonlinear process models, a continuous stirred tank reactor (CSTR), an evaporator process and a fluid catalytic cracking unit (FCCU). Case studies representing different MPC degradation scenarios are simulated, followed by a systematic procedure for diagnosing, isolating and recovering from such degradation, based on assumed operator's perspective and expert's technical reasoning. The knowledge gained from the case studies is used to develop an outline of a vision for a data-driven model predictive maintenance tool to help the operator make sensible judgements about performance degradation, the form and direction of diagnosis and fault isolation, and possibly, the recovery procedure.

Chapter One

Introduction

1.1 Background

Industries now operate in an environment where competition for scarce resources is stiff and where efficiency of operation and production of high quality competitive products can make the difference between economic viability and bankruptcy. This need to improve efficiency of operation has led many industries to adopt advanced control strategies in their operations. Model predictive control (MPC) is one of the advanced controls that has found widespread use in industries, particularly in process industries like oil refining and petrochemicals, from where the MPC technology originated in the 1970s (Qin and Badgwell, 2003). Other industries like chemicals, food processing, pulp and paper, mining and metallurgy, aerospace, defence and automotive have also embraced the technology (Charos et al., 1991, Qin and Badgwell, 2003, Zanolello and Budman, 1999). The MPC is particularly popular because of the ease with which it can be applied to multivariable systems, and its ability to handle constraints almost effortlessly.

Unlike traditional control methods like PID control, MPC might still be viewed or regarded as too advanced for actual plant operator understanding. This is why, unfortunately, industries that have achieved considerable success in the application of MPC technology have been the ones that are usually supported by a large group of specialised MPC control engineers. These groups of engineers are responsible for issues such as the identification of hardware problems that are contributing to poor MPC performance, the detection of significant drift in plant dynamics, and the analysis of

control systems performance. MPC operation might be opaque to the average desk operator, maintenance engineer or even technical operations support. The maintenance tool in this thesis offers a partial solution to those industries that do not have the luxury of large group of support engineers, who are versed in MPC, and for which control actions require attention.

There is anecdotal evidence to suggest that MPC controller performance sometimes deteriorates in commercially operated process plants sometime after commissioning. The MPC is commissioned by specialist engineers who visit the plant to set-up the MPC. Sometime later, the operators start to observe deterioration in performance. In a study by Harris et al. (1999) it was found that as many as 60% of all industrial controllers have some kind of performance problem. This is particularly important for MPC, which relies on the support of experts in control engineering to help assess, detect and diagnose the causes of performance deterioration.

Further evidence suggests that, even where there are significant levels of MPC expertise and experience available, the complexity of the combined process and control system can lead to engineers and operators misdiagnosing control problems. Confusing the effects of an unmeasured disturbance or a hardware problem with issues relating to the MPC system itself can cause operators to interfere with MPC systems in such a way that the performance of the controller is actually reduced further (Huang et al., 2000). In some extreme situations misdiagnosis might even lead to the termination of MPC. Hence, even where high levels of resources and expertise are available, the full potential of MPC might never be realised.

While researchers have devoted considerable attention to the development of robust algorithms, in most cases industry is generally more interested in the “control” of the

MPC, not its algorithm (Darby and Nikolaou, 2012, Huang et al., 2000). So for operators a key question is likely to be “how can MPC be made more transparent to them?” Can the answers to questions like: “which set of manipulated variables (MVs) should be used for a given set of controlled variables (CVs)” be made simpler and clearer to them? How can they realise when a MPC controller is performing poorly and how can they identify the source of performance deterioration? So when performance deterioration occurs, a key decision for operators might be to decide whether they themselves can return the MPC performance to previous levels, whether they require specialist support or whether they should simply switch the MPC off.

MPCs are most commonly employed in a supervisory capacity over lower level regulatory control. Here critical issues include the consideration of how the configuration of lower level regulatory controllers might affect the performance of the MPC, in terms of meeting its control objectives, satisfying constraints and robustness. The issues of variable selection (pairings of CVs and MVs) and the use of cascade control can have a far reaching impact on the performance of a MPC. As Darby and Nikolaou (2012) observed about the configuration of regulatory control: “design decisions have a far bigger impact on the success of an MPC project than the performance of the MPC algorithm itself”.

1.2 Spirit behind the maintenance tool

Over the years there have been many studies on control performance assessment methods covering univariate, multivariate and model based systems (Harris and Seppala, 2002, Harris et al., 1999, Jelali, 2006, Qin, 1998, Qin and Yu, 2007). The studies have led to the development of methods and metrics, which are mainly statistically based, for evaluating control performance. These include the minimum variance benchmark (comprising autocorrelation test, cross correlation test, closed loop

potential, performance index etc.), power spectrum analysis, sensitivity and complimentary sensitivity (frequency domain method) and principal component analysis. Other methods and metrics which were developed to specifically address multivariate systems like the MPC include the historical benchmark, the linear quadratic Gaussian (LQG).

Few actual industrial applications of these methods in monitoring, assessing and isolating faults in MPC have however been reported. Patwardhan et al. (1998) applied multivariate performance measures to isolate the causes of poor controller performance in an industrial plant (propylene splitter DMC MPC) and suggested remedial actions. Gao et al. (2003) used minimum variance benchmark to evaluate the performance of two industrial MPCs. Jiang et al. (2011) used a variety of methods, including model fit index, to analyse the performance of two industrial MPCs.

Huang et al. (2000) investigated the poor performance of MPC on an industrial plant without using any of the assessment methods and metrics mentioned earlier. Their approach was completely data-driven. They went through a process of systematic diagnosis: studying process input and output trends as well as their spectrum plots, and also checking and eliminating suspected causes of poor performance in turn until the root cause of the poor performance of the MPC was identified. Their approach is devoid of complex statistical analysis and might appear more appealing to an average operator.

The fact that MPC is based on time-varying objectives and that the output trends during normal operations are usually affected by operating conditions have made generalizing most of the assessment methods and metrics for MPC difficult. For this reason Patwardhan and Jay (1997) suggest an assessment framework which “allow the user to

combine the operating data and experience (as they are gathered) with the model in a synergistic manner for gradual improvement in the efficiency”.

MPC is normally configured and optimised for a specific plant. This configuration might include an internal model, cost function, constraints, tuning parameters, set-points, manipulated variables, controlled variables, measured and unmeasured disturbances, valves and regulator settings. Over a period of time however the conditions for which the MPC was initially configured might be no longer valid, leading to deterioration in the performance of the MPC.

Anecdotal evidence suggests that when faced with perceived MPC performance degradation, the operator usually takes knee-jerk actions which may further compound the problems. It is easy to understand why this is usually the case: MPC design is often opaque to the operator. Apart from the approach of Jiang et al. (2011) and that of Huang et al. (2000) before it, other works on controller performance assessment, though contributing immensely to the understanding of the issues involved might have alienated a large number of operators who are involved in the day to day operation of MPC implementations. How does the operator avoid the kind of often reported situation where he gets worried that the controller is deteriorating and ends up taking knee jerk actions that cause further problems in MPC? In what ways can the operator be provided with a tool which simplifies the basic MPC principles to him and enables him to make informed interventions for minor cases and be aware of when to seek expert advice? What are the commonly reported causes of MPC performance degradation and how does the operator recognise them when they occur? What steps may be taken by the operator to recover from such degradations?

The above are the more pertinent questions that bother an operator. This thesis seeks to provide answers to the above questions. Using models of nonlinear plants as case studies, the thesis outlines tools to enable the operator to have more informed interactions with the controllers and to facilitate detection of common MPC problems like improper controller configuration and model plant mismatch. The focus of the thesis is on the vision, rather than on the implementation. There has been insufficient time to produce a software tool, so the main contribution here is in the procedures that will be outlined in Chapters 4 and 5. These procedures have been developed by studying MPC performance on a number of process simulations that were found to exhibit particularly problematic behaviour. Chapter 3 describes these simulations before they are used in Chapters 4 and 5.

1.3 MPC in the Process Industries

Model predictive control (MPC) belongs to the class of controllers that use explicit, internal (linear or non-linear) models of a plant to decide how to drive its output variables to their set points. Though it may be desirable to use theoretical models based on the chemistry and physics of a plant, in most applications this is impracticable because of the complexities of the real life processes involved. For most practical purposes internal models are usually obtained empirically through the process of system identification. System identification involves adding special excitation signals to the inputs of a plant at steady state. The resulting deviations of the plant outputs, from their steady state values, are recorded and then relevant model plant models are estimated using direct methods or statistical packages. The excitation signals are of different types, the commonest are step signals and pseudorandom binary signals (PRBS). There are currently many statistical methods for estimating plant models from excitations. They include prediction-error minimisation (PEM) and subspace state space system

identification (n4sid). The MATLAB system identification toolbox contains these packages; it can be used to estimate low order transfer function (TF) models, state space (SS) models, and many types of input-output difference models (AR, ARX, ARMAX, Box–Jenkins etc.).

MPC is basically formulated as an optimisation problem. It uses current plant measurements, the current dynamic state of the process, models, and process variable targets and limits to calculate future changes in the dependent variables, while it simultaneously strives to ensure that constraints on its inputs and outputs are not violated. The calculations are done in a moving or receding horizon manner. The optimum control moves necessary to drive a plant to a reference point within a given prediction horizon are calculated at specific time intervals. Only the first of the resulting control moves, which spans a control horizon, are actually sent to the plant. This process is repeated at every time interval. By far the most convincing analogy for MPC is that given by Wang (2009) where it was stated that “The day begins at 9 o’clock in the morning. We are, as a team, going to complete the tasks of design and implementation of a model predictive control system for a liquid vessel. The rule of the game is that we always plan our activities for the next 8 hours work; however, we only implement the plan for the first hour. This planning activity is repeated for every fresh hour until the tasks are completed.”

The optimisation problem usually incorporates constraints on both inputs and outputs. Usually, plants run more profitably and efficiently when they are operated at or near constraints. Constraints can arise due to the need for profit optimisation (increase in throughput or decrease in inventory and/or operating expense), requirement for high quality products, limits in equipment operation range (valves and other actuators

saturation or slew rates), legislative and regulatory requirements (e.g. pollution control) and safety requirements (e.g. the need not to exceed surge compressors, etc.).

1.4 Role of Control Performance Assessment and Faults Detection and Diagnosis

Model based controllers like the MPC are characterised by many assumptions and approximations in the specification of the model, the measured and unmeasured disturbance dynamics, the choice of input and output variables, and the choice of input and output constraints, that make it almost an impossibility to have perfect MPC. These are the factors that most often lead to performance deterioration or complete breakdown.

When MPC performance deterioration occurs, one of the major tasks is to establish if the deterioration is due to regulatory control, supervisory control or whether it is associated with another cause (Schäfer and Cinar, 2004). Simply observing the trends of controller inputs and outputs may be sufficient to detect degradation and even diagnose the cause of poor performance in some cases. Many control performance assessment methods have been approached from the point of view that the control systems are complex and that the trends of the raw data from the plant often show complicated response patterns resulting from the presence of disturbances, noise, time variant response phenomena and nonlinearities (Schäfer and Cinar, 2004) and as such complex assessment methods are required to handle such situations. The fact remains that MPCs often degrade and some of the diagnostic steps contained in the assessment methods are beyond the comprehension of most operators.

Although extensive research has been carried out into MPC issues such as robustness, improved prediction algorithms, and closed loop stability, very little has been done on

issues that directly address the operator's biggest concern: how do I know what is responsible for a deterioration in my MPC performance, and how do I possibly recover from this situation? The causes of deterioration in industrial applications are many and have been widely reported in literature. Common causes of poor performance reported in industrial contexts include:

- a) controller design (e.g. wrongly specified measured disturbances, improper controlled variable and manipulated variable selection, tuning parameters etc.)
- b) excessive process drift (leading to model/plant mismatch)
- c) large and unmanageable measured and unmeasured disturbances
- d) hardware problems involving regulatory controllers and valve degradation

Chapters 4 and 5 describe procedures that the operator might follow to detect and isolate each of these causes.

1.5 Outline of the Maintenance Tool

This maintenance tool is intended to reduce this opaqueness and improve the operator's confidence and understanding in his interaction with the installed MPC. The MPC envisioned here enables the operator to reason about key features of their MPC and more importantly, about its degradation in performance. The tool allows the operator to interact with offline data, and to carry out what-if simulations of scenarios of possible MPC problems.

The tool includes features that enable the operator to compare actual performance trends with previously defined reference trends. From offline plant data obtained during commissioning or during a period of good performance, the tool plots reference

graphical performance (RGP) trends. Actual performance data are obtained periodically from the plant and are used to generate actual graphical performance (AGP) trends.

This comparison is used in conjunction with a set of diagnostic simulations that relate to possible MPC problems, to isolate causes of poor MPC performance. The diagnostic procedure uses what if simulations performed on the linear model, combined with carefully structured questions, based on knowledge of the behaviour of MPC for some performance issues, to lead the operator to the cause of a problem. From the comparison the operator is expected to select from a structured list of diagnostic questions. When a question is selected, the operator is led to a list containing underlying symptoms associated with the selected question. Again by making a choice from this list of suspected symptoms, the operator is guided through a script containing the procedure for investigating the symptom, with a view to either confirming or rejecting the suspicion. If the suspicion is confirmed, hopefully the recovery procedure should be obvious to the operator.

In carrying out the investigations, the operator is assisted by a number of background information tools which help as refreshers, guides and tutors when there is a need to embark on a certain diagnostic procedure or to explain some observed trends. Foremost among the background information tools is a virtual plant of the real plant on which MPC, similar to the one on the real plant, is implemented. The assumption is that a linear model of the plant under consideration exists, and that this model is accessible to the operator. Assuming that a linear model of the plant exists, the maintenance tool uses this model together with data on the actual MPC settings (prediction horizon, control horizon, constraints etc.) to mirror MPC performance on the assumption that the plant itself is linear.

Other components of the background information tool include steady state gains, relative gain arrays and relative weight arrays. It also includes the transfer function matrix of the linear model, and the step response plots of the linear model and/or of the real plant at commissioning. In cases where the transfer function matrix of the linear model does not exist, Appendix A describes how an approximate transfer function matrix model may be obtained by the operator.

1.6 Methodology

The research work centres round simple benchmark models of nonlinear process plants, which serve as a virtual world to mimic the real world of plant operations. The various internal models of the various MPCs used to control the various nonlinear plants are based on first order and second order transfer functions obtained from the identification of the plant simulations. These linear representations are adopted because they are defined primarily by terms like time constant, damping coefficient and natural frequency; terms to which most plant operators may find it easier to relate, as compared to e.g. terms that relate to state space or input-output representations.

The research accommodates the common industrial practice of employing MPC in a supervisory capacity over regulatory control. This practise is advantageous in that the regulatory controllers can be relied upon to perform in a “fall back” position if the MPC is switched off due to degradation. In so doing, issues surrounding the effect of lower level regulatory controllers on MPC performance can be studied. The MPCs implemented on the plant simulations are based on best performing algorithms sourced and obtained from literature; all implementations are written by the author. One of the reasons for adopting this approach is to avoid opaqueness associated with available implementations of MPC. The other reason is the appreciation of the fact that if one

must diagnose and isolate problems related to MPC, one must understand very well MPC algorithms developments.

The program codes for the three different forms of MPC have been written. The first is the one based on the finite impulse/step response format (dynamic matrix control), the second is the one based on the transfer function format (generalised predictive control) and the last is based on the now very popular state space format.

The MPC in the maintenance tool is unlikely to have all the sophistication of that supplied by a vendor and installed in the real plant: the aim is to provide something that is adequate as a diagnostic tool, i.e. exhibits the correct symptoms when it degrades. The assumption is that this requirement is likely to be independent of the sophistication of the actual MPC implementation.

1.7 Novel Aspects of this Thesis

The original contributions in this thesis are listed below.

- 1) A vision for an operator friendly MPC maintenance tool. The maintenance tool proposed is interactive and flexible: many windows guide and inform.
- 2) The use of process plant nonlinear models in the maintenance tool. Maintenance tool development traversed many stages; an earlier stage resulted in the presentation of a paper entitled “Recovering from a gradual degradation in MPC performance”, at the ADCHEM 2012, in Singapore, in July, 2012
- 3) A transfer function matrix for a FCCU, a complex nonlinear model involving many input-output pairs that exhibit inverse relationships.
- 4) A simple method for obtaining approximate first and second order transfer functions, including those exhibiting inverse relations, for an input-output pair.

This simple method was used to obtain the transfer function matrix for the FCCU.

- 5) The notion of a relative weight array, which might assist operators in MPC diagnosis.

1.8 Summary

This chapter presents a case for the necessity and significance of the research as well as presenting an outline of the whole thesis. Brief background of model predictive control, its relevance in process control as well as its peculiarities in terms of performance maintenance, and the challenges this poses to MPC operators were discussed. The methodology for the research involved adoption and simulation of three nonlinear plant models, writing MPC program codes in MATLAB, implementing MPC on the nonlinear models, inducing MPC faults into the MPC implementations, and making the faults isolation procedure very transparent.

Chapter Two

Review of Linear MPC

2.1 MPC Principles and Fundamentals

MPC has been particularly popular (Maciejowski, 2002) because unlike conventional controllers, it can:

- a) handle multivariable control problems naturally,
- b) account for actuator limitations
- c) allow operation close to constraints

MPC however has certain drawbacks that are not widely reported. Some of these limitations include (Hugo, 2000):

- a) operational difficulty;
- b) high installation and maintenance cost;
- c) lack of flexibility (plant specific);
- d) sluggish disturbance rejection

MPC uses an internal model to predict future process behaviour and in particular a prediction equation for a given prediction horizon is constructed from it. The prediction equation is combined with set-point information to formulate the necessary objective function, or cost function, for a given control horizon. This cost function is then solved to obtain an optimal control law. If the objective function does not include constraints, an explicit solution for the optimal control moves can be obtained using the least squares methods. If constraints are incorporated however, the optimal control moves are obtained as a solution to the resulting quadratic programming (QP) problem. There are

a number of methods available for solving QP problems like the gradient projection method, the interior point method and the active set method. The algorithms for obtaining a numerical solution are also varied. For example Hildreth's quadratic programming procedure (Luenberger, 1984, Wismer and Chattergy, 1978) in Wang (2009) offers a simple algorithm for the solution of the QP problem when it is expressed as a dual problem to the original primal problem.

Internal models may be linear or non-linear. When a linear internal model is used, the MPC is referred to as linear MPC. The form of the internal model determines the nature of the prediction equation, and hence the objective function and the control law. The basic cornerstone of MPC control calculations is to determine, at a sampling instant k , (at regular sampling period T_s) a sequence of M optimal control moves ($[u_{(k)}, u_{(k+1)}, \dots, u_{(k+M-1)}]$) that ensures that P predicted outputs ($[\hat{y}_{(k+1)}, \hat{y}_{(k+2)}, \dots, \hat{y}_{(k+P)}]$) tracks a set-point trajectory ($[s_{(k+1)}, s_{(k+2)}, \dots, s_{(k+P)}]$) optimally (figure 2.1). P is the *prediction horizon* and M is the *control horizon*.

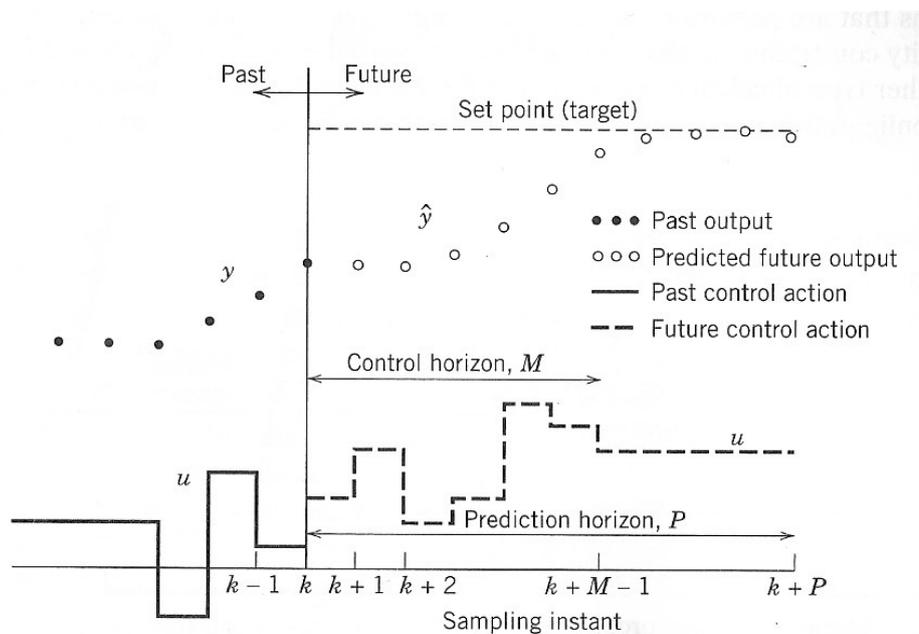


Figure 2.1: Basic Concept of Model Predictive Control (Seborg et al., 2010)

The MPC uses information about measured outputs (y), measured disturbances (v), set-points (s) and possibly constraints at instant k (figure 2.2) to make the decision about the optimal set of control moves. Though M control moves are calculated at every sampling time, only the first control move (u) is sent to the plant as the manipulated variable. It should be noted that the unmeasured disturbance d is not known directly, although it might be inferred from plant outputs. For this reason some model-predictive-controllers incorporate explicit unmeasured disturbance models to account for the effect of the unmeasured disturbances and/or to account for model plant mismatch. In this way the controller is designed to provide feedback compensation for such unmeasured disturbances. In contrast, the measured disturbance (if it exists) is known to the MPC. The MPC can then provide feed-forward compensation for such disturbances to minimize their impact on the outputs.

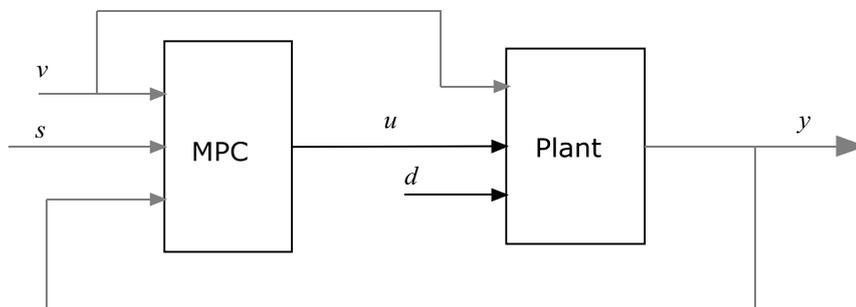


Figure 2.2: Signal Flow of a MPC (adapted from Bemporad et al. (2013))

2.2 Development of MPC

MPC has undergone a great deal of transformation and re-development to arrive at how it is currently perceived in terms of terminology and scope (algorithms formulation, constraints handling, robustness, applications etc.). Earlier model based control methods such as LQR (linear quadratic regulator), MVC (minimum variance control), MAC (model algorithmic control), IMC (internal model control), MOCCA (multivariable, optimal constrained control) and MPHIC (model predictive heuristic

control) have features of what is now generally defined as MPC (Li et al., 1989). Most authors who have reviewed the history of MPC (Froisy, 1994, Froisy, 2006, Maciejowski, 2002, Qin and Badgwell, 1997, Qin and Badgwell, 2003) seem to agree that the first publications on the application of MPC can be attributed to (Richalet et al., 1976, Richalet et al., 1978). The reviewers however mostly concede that the idea of MPC was not completely new at that time. For example Qin and Badgwell (2003) cite previous works of Propoi (1963) and Lee and Markus (1967). The same paper (Qin and Badgwell, 2003) reported that Shell Oil applied a form of MPC in 1973.

Richalet et al. (1976, 1978) referred to their predictive controller as Model Predictive Heuristic Control (MPHC). It is heuristic in the sense that the MPHC techniques were based on the developers' experience and intuition, and the solutions were not always guaranteed to be optimal. Input and output constraints were not defined explicitly. It did have some of the features of the current MPC like a quadratic performance objective, the use of a reference trajectory and a finite prediction horizon. The MPHC algorithm was implemented solution software referred to as Identification and Command (IDCOM). IDCOM used a discrete-time finite response (FIR) model as the MPC internal model.

The work of Cutler and Ramaker (1980) was reported soon after. Cutler and Ramaker (1980) proposed an unconstrained MPC algorithm referred to as Dynamic Matrix Control (DMC) in which the internal model was based on the step response model of the plant. Like in IDCOM, the DMC algorithm used a quadratic performance objective over a finite prediction horizon. DMC incorporated improved algorithms for computing the optimal control signal through the solution of a linear programming (LP) problem, especially for unconstrained problems. Input and output constraints were not explicitly

included in the algorithm, but were still handled on an ad hoc basis. To address this shortcoming, Cutler et al. (1983) published an improved version of DMC, which they termed Quadratic Dynamic Matrix Control (QDMC). Unlike DMC, QDMC had an explicit input and output constraint formulation in its algorithm, which was presented in the form of a quadratic program (QP). The optimum control inputs were computed as the solution of the QP problem. Output constraints were formulated in terms of soft constraints. Disadvantages of the QDMC includes the fact that it did not provide for a means of handling an infeasible solution, lack of robustness, (fault intolerance); neither did it provide a means of ranking the soft constraints in order of importance. Infeasibility might also occur when there were more controlled variables than manipulated variables in a set-point tracking MPC.

The development of MPC, with improved algorithms to overcome practical issues (infeasibility, robustness, soft constraints handling etc.), then surfaced in industrial applications of MPC. MPC applications included IDCOM-M (multivariable IDCOM), HIECON (Hierarchical Constraint Control Technology), SMOC (Shell Multivariable Optimizing Controller), RMPCT (Robust Multivariable Predictive Control) and Connoisseur. IDCOM-M (Froisy and Matsko, 1990, Grosdidier et al., 1988) has a means of handling infeasibility by incorporating a controllability supervisor which screens out ill-conditioned plant subsets. It also incorporates multiple quadratic objective functions: a quadratic output objective function followed by a quadratic input objective function. It explicitly specifies hard and soft constraints in its objective functions, with hard constraints ranked in order of priority. It also employs a new concept that of the use of coincidence points, which are chosen from a subset of the reference trajectory. The SMCA (set-point multivariable control architecture) is the

results of combining the identification, simulation, configuration and control of IDCOM-M related products into a single package (Qin and Badgwell, 2003).

HEICON is a predictive control product, developed about the same time as IDCOM-M, with many of the IDCOM-M features. The HEICON algorithm however has a distinctive feature: it uses a predictive control method called Functional Predictive Control (PFC), where the future input is assumed to be a linear combination of a few *basis functions* (Maciejowski, 2002). The basis function is usually a polynomial of order c , where c is a tuning parameter.

Different types of MPC use different types of internal model: the transfer function or input-output difference equation was reported in the 1980s (Clarke et al., 1987, Peterka, 1984). This class of MPC is referred to broadly as Generalised Predictive Control (GPC). The two popular products in this category are RMPCT and Connoisseur. RMPCT used the auto regressive with exogenous input (ARX) model as internal model, while Connoisseur used the auto regressive moving average with exogenous input (ARMAX). RMPCT implements zone or funnel control, instead of set-point tracking. Like IDCOM-M, it uses the QP objective function to calculate future inputs, and has two levels of control: one for set-point optimization, and the other for optimum input calculation. It allows blocking of predicted moves, and incorporates strategies to handle ill-conditioning. One distinctive feature is that it could adjust the internal model by evaluating ISE (Integrated square error) values of alternatives. Connoisseur also uses quadratic cost function, with an explicit definition of input and output constraints in the objective function. It has a feature for input blocking too.

MPC algorithms incorporating linear state space internal models were introduced by the late 1980s. The first account and description of this form of MPC (Qin and Badgwell,

2003) were contained in Marquis and Broustail (1988) and Yousfi and Tournier (1991). Their publications described a form of MPC they referred to as the Shell Multivariable Optimizing Controller (SMOC). The state space internal model could accommodate problem dynamics easily: integrating, stable and unstable. SMOC incorporated many of the now standard features of MPC: an explicit disturbance model to describe the effect of unmeasured disturbance, a Kalman filter to estimate the states of the plant, an explicit specification of input and output constraints, soft constraints ranking, infeasibility handling, the solution of optimum control moves via solution of QP problem and the ability to handle a wide range of processes (stable, unstable, integrating). Further publications and improvements to the state space MPC algorithms followed as shown in Ricker (1991), Rawlings and Muske (1993), Rawlings (2000), Maciejowski (2002), Wang (2009) and many others.

Over the last decade or thereabouts, there has been a great deal of transformation in the MPC landscape: many MPC product vendors have merged and MPC products have been developed to incorporate many more of the now standard MPC features and terminologies. The trend has been towards the development of a product comprising identification, control, assessment, monitoring and diagnosis as well as robustness. The MPC products now incorporate advanced graphical user interfaces. Some of the current leading developers and vendors of MPC technology are Aspen Technology (Aspentech), Honeywell, Shell and ABB.

Aspentech markets DMCplus® (Aspentech, 2013). As currently marketed the MPC product has many additional capabilities . It can automatically identify and detects performance issues with models and pinpoints the areas of the model needing attention. It can also maintain robust controller behaviour during background step testing, through

the use of adaptive process control. It can collect suitable data for modelling using small amplitude testing while still maintaining optimizing control. It has the capacity to monitor issues like valve saturation, PID mode changes; process upset detection, and bad measurement readings, and can detect and repair co-linearity.

3dMPC is a software suite developed and marketed by ABB Automation (ABB, no date). The latest version boasts of improvements which include on-line components for control and operator interaction as well as off-line components for controller specification, modeling, tuning and analysis. It has the ability to independently design set-points and feedback/ feed-forward parts of the controller, thus giving great design flexibility and robust process handling more accurate controller commands. It incorporates the use of performance variable methodology to enable the controller to sense changes in process conditions that cannot be measured directly, and initiates countermeasures long before normal MPC technology can detect the variations. An on-demand adaptation which permits automatic adaptation of the controller to suit changed process conditions, while still maintaining overall control of the production is part of its improvements.

The MPC suite developed and currently marketed by the Shell Group is called *Pro Technology*. Pro Technology comprises of AIDA^{Pro} (Advanced Identification and Data Analysis), for process identification, MD^{Pro} (monitoring and diagnosis), for MPC performance monitoring and diagnosis, SMOC^{Pro} (Shell Multivariable Optimizing Control), for optimum control calculations, and RQE^{Pro} (Robust quality Control) to handle issues of robustness, fault intolerance etc (Shell, 2004).

Honeywell's markets the integrated control suite is called *Profit Controller* (Honeywell, no date). Profit Controller comprises of the Honeywell Range Control Algorithm

(RCA) and the Robust Multivariable Predictive Control Technology (RMPCT). The RCA minimizes the effects of model changes or uncertainty while determining the smallest process moves required to simultaneously meet advanced process control and optimization objectives. The RMPCT helps to stabilize complex processes to reduce operating upsets and drives processes to their optimal operating level. Its increased robustness enables the controller to stay on-line over a wider range of operating conditions, resulting in higher and more profitable operations.

Table 2.1 below gives a summary of the popular types of MPC that have been developed or are currently being marketed. The summary highlights key features of each MPC in terms of the type of internal model used, the constraints handling capabilities, the types of objective functions and their robustness.

2.3 MPC Internal Models

All the currently available MPC designs are based on one of three general approaches, each of which is determined by the type of internal model used in its formulation. The type of internal model affects the specification of the prediction model, and by extension the objective functions as well as the algorithm for obtaining the optimum control law. The three types of internal model are the finite impulse response/step response model, the transfer function (also called difference equation or input-output) model, and the state space model. The following sub-sections give brief descriptions of the three MPC approaches, and summarises the strength and weaknesses of each formulation.

Table 2.1: Summary of major MPC products (adapted mainly from (Maciejowski, 2002, Qin and Badgwell, 2003))

S/N	Product	Model	Features
1	IDCOM Identification and Command (1976)	Impulse response	<ul style="list-style-type: none"> • Quadratic objective function over a finite prediction horizon • Input and output constraints included in the formulation • Optimal inputs computed using heuristic iterative algorithm • Use of reference trajectory
2	IDCOM-M (1988)	Impulse response	<p>Offshoot of IDCOM, with the following improvements:</p> <ul style="list-style-type: none"> • Controllability supervisor to screen out ii-conditioned plant subsets • Multiple objective function (quadratic output objective function followed by quadratic input objective function) • Provides for hard and soft constraints explicitly in the objective function • Rank outputs in order of priority • Calculates one single move for each input • Use of coincidence points to control subsets of future points in time • Optimal solution computed as quadratic problem. • Provides move suppression
3	DMC Dynamic Matrix Control (1980)	Step response	<ul style="list-style-type: none"> • Developed by Shell Oil • Quadratic performance objective function over a finite prediction horizon • Set-point tracking • Constraints not explicitly specified in the objective function, but done ad hoc • Optimal solution computed as least squares problem. • For stable systems only
3	QDMC Quadratic Dynamic Matrix Control (1983)	Step response	<p>Offshoot of DMC, with the following improvements:</p> <ul style="list-style-type: none"> • Input and output constraints explicitly specified in the objective function • Optimal inputs computed as solution to a quadratic problem.
4	DMC-plus	Step response	<p>Offshoot of DMC and QDMC, with the following improvements</p> <ul style="list-style-type: none"> • Can be extended to accommodate integrators • Control law computed in two steps, both as solutions to LP problems • Constraints are imposed through variable penalty weights approach • Infeasible solution handled ad hoc through the use of “equal concern error” concept • Improved identification technology

S/N	Product	Model	Features
5	RMPCT Robust Multivariable Predictive Control Technology	ARX	<ul style="list-style-type: none"> • Developed by Honeywell • Implements the zone or funnel control, instead of set-point tracking • QP objective function to calculate future inputs • Two levels of control – one for optimization, and the other for control • Allows blocking of predicted moves • Incorporates strategies to handle ill-conditioning • Can adjust the internal model by evaluating ISE values of alternatives
6	Connoisseur	ARMAX	<ul style="list-style-type: none"> • Developed by simulation sciences • Uses a quadratic cost function • Explicit definition of input and output constraints in the objective function • Incorporates input blocking • Optimization of set-points via linear programming
7	HIECON Hierarchical Constraint Control		<ul style="list-style-type: none"> • Based on the principles as PFC (treats future inputs as a linear combination of a few simple basis function, usually polynomial) • Incorporates fully constraint optimization • Multivariable control
8	SMOC Shell Multivariable Optimizing Controller	State space	<ul style="list-style-type: none"> • Developed by shell • Can handle stable, unstable and integrating processes • Use of Kalman filter to estimate plant states • Input and output constraints defined explicitly, and optimal inputs obtained through solution to QP problem • Use of an explicit unmeasured disturbance model
9	3dMPC	State space	<ul style="list-style-type: none"> • Developed by ABB • Quadratic cost function • Hard and soft constraints • Can prioritize output constraints • Incorporated direct identification using subspace method

2.3.1 Finite Impulse/Step Response MPC

Finite Impulse/Step Response MPC use linear discrete time empirical finite impulse/step response models in its formulation. The step response model is given as:

$$y_{(k+1)} = y_0 + \sum_{i=1}^{N-1} S_i \Delta u_{(k-i+1)} + S_N u_{(k-N+1)} \quad \dots(2.1)$$

where $y_{(k+1)}$ is the output variable at the $(k+1)$ -th sampling instant; y_0 is the initial value; $\Delta u_{(k-N+1)}$ denotes the change in the manipulated input from one sampling instant to the next; S_i are the step response coefficients for the sampling instants between 1 and N. N is the number of step response coefficients used for the model. For the purpose of MPC application, the numbers of step response coefficients required are chosen large enough to cover at least 99% of the open-loop settling time of the plant. The normal value of N lies between 30 and 120 sampling instants. This consideration logically affects the choice of sampling time (T_s) for the system. Recent discussions on the development of algorithms for step response MPC formulation are given in Hokanson and Gerstle (1993), Huang and Kadali (2008) and Seborg et al. (2010).

Step response MPC formulation is appealing because the step response based internal model is noted to give transparent descriptions of process time delays, response times and gains in a manner that cannot be accurately described by either of the other approaches. The approach is intuitive and clearly reflects the effects of each manipulated variable. Its major disadvantage is that it is limited to stable processes and processes without integrators.

2.3.1 Transfer Function MPC

The transfer function MPC formulation is based on the transfer function model. The transfer function models used in MPC are the linear polynomial input-output parametric difference equations of the form:

$$A(z^{-1})y(k) = \frac{z^{-d}B(z^{-1})}{F(z^{-1})}u_k + \frac{C(z^{-1})}{D(z^{-1})}e_k \quad \dots(2.2)$$

where

e_k is the white noise sequence (the stochastic disturbance input).

d is the pulled out input-output delay.

$A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$ and $F(z^{-1})$ are polynomials with the back shift operator z^{-1} defined as follows:

$$A(z^{-1}) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{na}z^{-na}$$

$$B(z^{-1}) = b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_{nb}z^{-nb}$$

$$C(z^{-1}) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{nc}z^{-nc}$$

$$D(z^{-1}) = 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_{nd}z^{-nd}$$

$$F(z^{-1}) = 1 + f_1z^{-1} + f_2z^{-2} + \dots + f_{nf}z^{-nf}$$

In equation 2.2, if $C(z^{-1})$, $D(z^{-1})$ and $F(z^{-1})$ are each equal to unity, the resulting difference equation is referred to as the autoregressive with exogenous input (ARX) model. When both $F(z^{-1})$ and $D(z^{-1})$ are each equal to unity and all the other terms are present, the resulting different equation is called autoregressive moving average with exogenous input (ARMAX) model. An autoregressive (AR) model is obtained when $C(z^{-1})$, $D(z^{-1})$ and $F(z^{-1})$ are each equal to one and $B(z^{-1})$ is equal to zero.

MPC based on a transfer function formulation is suitable for stable and unstable processes, and typically requires fewer past inputs as compared to step response MPC. Transfer function MPC requires n past output data and $(n+d)$ past input data (where n is the order of the model, and d is the input delay) for its prediction. Clearly the choice of n will influence model fidelity, so some thought needs to be applied when choosing the order. For algorithms for MPC based on transfer functions Camacho (1993), Maciejowski (2002) and Camacho and Bordons (2004) give derivations.

2.3.2 State-Space MPC

The state space MPC formulation is based on a state-space model, usually in the discrete time format. Discrete time state space models are of the form:

$$x_{k+1} = Ax_k + Bu_k \quad \dots(2.3)$$

$$y_k = Cx_k + Du_k \quad \dots(2.4)$$

where x is a column vector of dimension n (number of states), u is a column vector of dimension q (number of inputs), and y is a column vector of dimension m (number of measured outputs). A is referred to as the state (or system) matrix, B as the input matrix, C as the output matrix, and D as the feed-through (or feed-forward) matrix.

Using state space models, multivariable processes can be represented in a straight forward manner and a large collection of modern control theory and analysis method can be applied easily to state space MPC algorithm development. Despite its widespread use, state space MPC has attracted some criticisms: one obvious criticism is that certain processes may require very large state, input or output matrices to fully describe them, and that this may lead to numerical problems during computations. Another criticism is that a state space model is not very transparent: it is most times difficult to see what the figures are saying, except for very experienced engineers. Another criticism is that the state space MPC require more complicated prediction and optimal control calculations, with the additional necessity of including an observer when the states are not accessible (Huang and Kadali, 2008). Some researchers are now showing that some of these criticisms are not completely justified. For example Maciejowski (2002) showed that complicated multivariable step response model with input delays can be converted to equivalent state space model.

State space MPC formulations are well documented. See for instance Garcia et al. (1989), Muske and Rawlings (1993), Kothare et al. (1996), Camacho and Bordons (2004) and Borrelli et al. (2005).

2.4 MPC Input Objective function

Fundamental to MPC is the calculation of a set of optimum control moves at every sampling interval. This set of optimum control moves is the solution of an appropriately formulated quadratic objective function. The quadratic objective function, when written in the vector-matrix notation, is of the form:

$$\min_{\Delta \mathbf{U}_{(k)}} J = (\mathbf{s}_{(k+1)} - \hat{\mathbf{Y}}_{(k+1)})^T \mathbf{Q} (\mathbf{s}_{(k+1)} - \hat{\mathbf{Y}}_{(k+1)}) + (\Delta \mathbf{U}_{(k)})^T \mathbf{R} (\Delta \mathbf{U}_{(k)}) \quad \dots(2.5)$$

where:

$\mathbf{s}_{(k+1)}$ is a vector of set points for all outputs within prediction horizon

$\hat{\mathbf{Y}}_{(k+1)}$ is a vector of all predicted outputs within the prediction horizon

\mathbf{Q} and \mathbf{R} are diagonal matrices containing the output and input weights respectively.

The solution of the quadratic objective function J is the set of $\Delta \mathbf{U}_{(k)}$ (within a given control horizon) which ensures that the error between the set-point $\mathbf{s}_{(k+1)}$ and predicted outputs $\hat{\mathbf{Y}}_{(k+1)}$ is a minimum.

In the quadratic objective function of equation 2.5, a reference trajectory vector $\mathbf{r}_{(k+1)}$ is commonly used in place of the set-point trajectory $\mathbf{s}_{(k+1)}$ vector. This ensures that the predictions make a gradual transition from the current output $y_{(k)}$ to the set-point $s_{(k+1)}$. The path of the reference trajectory can be defined as an exponential approach (Maciejowski, 2002), though there are many other possible definitions (Qin and Badgwell, 2003, Seborg et al., 2010).

Using the exponential approach, the reference trajectory for multiple input, multiple output (MIMO) system with m outputs and q inputs, and prediction horizon P is defined as:

$$r_{i,(k+j)} = \alpha_i^j y_{i,(k)} + (1 - \alpha_i^j) s_{i,(k)} \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, P \quad \dots(2.6)$$

where α_i is a filter term for each output and it gives a measure of the speed of response of the reference trajectory towards the set-point. Generally, $0 \leq \alpha_i < 1$.

The reference trajectory equation 2.6 can be compactly written in vector-matrix notation as:

$$\mathbf{r}_{(k+1)} = \mathbf{A}\mathbf{y}_{(k)} + \mathbf{B}\mathbf{s}_{(k)} \quad \dots(2.7)$$

where

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_P \end{bmatrix}; \quad \text{and} \quad \mathbf{r}_i = \begin{bmatrix} r_{1,(k+i)} \\ r_{2,(k+i)} \\ \vdots \\ r_{m,(k+i)} \end{bmatrix}; \quad i = 1, 2, \dots, P$$

$$\mathbf{y}_{(k)} = \begin{bmatrix} y_{1,(k)} \\ y_{2,(k)} \\ \vdots \\ y_{m,(k)} \end{bmatrix}; \quad \mathbf{s}_{(k)} = \begin{bmatrix} s_{1,(k)} \\ s_{2,(k)} \\ \vdots \\ s_{m,(k)} \end{bmatrix};$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_P \end{bmatrix}; \quad \text{and} \quad \mathbf{A}_i = \begin{bmatrix} \alpha_1^i & 0 & \dots & 0 \\ 0 & \alpha_2^i & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_m^i \end{bmatrix}; \quad i = 1, 2, \dots, P$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_P \end{bmatrix}; \quad \text{and} \quad \mathbf{B}_i = \begin{bmatrix} 1 - \alpha_1^i & 0 & \dots & 0 \\ 0 & 1 - \alpha_2^i & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 - \alpha_m^i \end{bmatrix}; \quad i = 1, 2, \dots, P$$

For the general case of a MIMO system with m outputs and q inputs and with individual output and input weights $(qq_1, qq_2, \dots, qq_m)$ and $(rr_1, rr_2, \dots, rr_q)$ respectively, the output and input weight diagonal matrices of equation 2.5 are defined as:

$$\mathbf{Q} = \overbrace{\begin{bmatrix} Q & 0 & \cdots & 0 \\ 0 & Q & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q \end{bmatrix}}^{P \text{ times}}; \quad \mathbf{Q} = \begin{bmatrix} qq_1 & 0 & \cdots & 0 \\ 0 & qq_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & qq_q \end{bmatrix}$$

$$\mathbf{R} = \overbrace{\begin{bmatrix} R & 0 & \cdots & 0 \\ 0 & R & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & R \end{bmatrix}}^{M \text{ times}}; \quad \mathbf{R} = \begin{bmatrix} rr_1 & 0 & \cdots & 0 \\ 0 & rr_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & rr_q \end{bmatrix}$$

Matrix \mathbf{Q} is a *weighing matrix* used to create preferential tracking of the outputs according to their relative importance. Relatively high qq_i values imply that the controller should be very concerned about the deviation of the predicted output y_i from its set-points, and should try as much as possible to make the output track its set-point. Matrix \mathbf{R} is a weighting matrix that is used to preferentially suppress aggressive behaviours of manipulated inputs. The higher the rr_i value, the more input u_i is suppressed.

All the three types of MPC (step response, transfer function and state space) use the quadratic objective function of equation 2.5. The major difference between the MPCs is in the derivation of the algorithm to calculate the predicted outputs $\hat{y}_{(k+1)}$ over a given prediction horizon P and a control horizon M. Brief discussions on the derivation of the algorithms for calculating the predicted output vector for each type of MPC is given in Section 2.5.

2.5: Algorithms for MPC Predicted outputs

The MPC prediction equation uses parameters M and P, with known reference trajectory set $(r_{(k+1)}, r_{(k+2)}, \dots, r_{(k+P)})$, input and output weight matrices R and Q , to predict a set of future outputs $(\hat{y}_{(k+1)}, \hat{y}_{(k+2)}, \dots, \hat{y}_{(k+P)})$.

The predicted output sequence consists of two terms: the free (or unforced) response (y^{fo}), and the forced response (y^{fr}). The free response depends only on the effects of past control actions (inputs and/or outputs). It corresponds to the response prediction of the output variable if the inputs are kept at their current values. The forced response accounts for the effect of current and future control actions. The predicted output ($\hat{y}_{(k+1)}$) is generally written as:

$$\hat{y}_{(k+1)} = y^{fo} + y^{fr} \quad \dots(2.8)$$

Both the prediction horizon (P) and control horizon (M) are also very important MPC tuning parameters. In cases where constraints are imposed, a high value of P enables the controller to see the activation of any of these constraints far ahead, so that early action can be taken to minimise their impacts, or avoid them completely. The selection of values for P and M are significant in situations where plants have time delays. If a plant has a time delay d, the controller's current move $u_{(k)}$ has no effect until $y_{(k+d+1)}$. So it is essential that $P \gg d$ and $M \ll P - d$ as this forces the controller to consider the full effect of each move (Bemporad et al., 2013).

2.5.1 Prediction Algorithm for Step Response MPC

Consider a finite step response model with N step response coefficients as given in equation (2.1), the set of predicted outputs over the prediction horizon P for a SISO system is given as:

$$\begin{aligned} \hat{y}_{k+1} &= S_1 \Delta u_k + \sum_{i=2}^{N-1} S_i \Delta u_{k-i+1} + S_N u_{k-N+1} \\ \hat{y}_{k+2} &= S_1 \Delta u_{k+1} + S_2 \Delta u_k + \sum_{i=3}^{N-1} S_i \Delta u_{k-i+2} + S_N u_{k-N+2} \\ &\vdots \\ \hat{y}_{k+P} &= \underbrace{\sum_{i=1}^P S_i \Delta u_{k+P-i}}_{\text{forced prediction}} + \underbrace{\sum_{i=P+1}^{N-1} S_i \Delta u_{k-i+P} + S_N u_{k-N+P}}_{\text{free prediction}} \end{aligned} \quad \dots(2.9)$$

This can be written in vector-matrix notation for a general case (MIMO system with m outputs and q inputs) as:

$$\hat{Y} = G\Delta U + (E\Delta U_p + FU_p) = Y^{fo} + Y^{fr} \quad \dots(2.10)$$

where

$$\hat{Y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_p \end{bmatrix}; \quad \text{and} \quad \hat{y}_i = \begin{bmatrix} \hat{y}_{1,(k+i)} \\ \hat{y}_{2,(k+i)} \\ \vdots \\ \hat{y}_{m,(k+i)} \end{bmatrix}; \quad i = 1, 2, \dots, P$$

$$\Delta U = \begin{bmatrix} \Delta u_1 \\ \Delta u_2 \\ \vdots \\ \Delta u_M \end{bmatrix}; \quad \text{and} \quad \Delta u_i = \begin{bmatrix} \Delta u_{1,(k+i-1)} \\ \Delta u_{2,(k+i-1)} \\ \vdots \\ \Delta u_{q,(k+i-1)} \end{bmatrix}; \quad i = 1, 2, \dots, M$$

$$\Delta U_p = \begin{bmatrix} \Delta u_{p,1} \\ \Delta u_{p,2} \\ \vdots \\ \Delta u_{p,N-p-1} \end{bmatrix}; \quad \text{and} \quad \Delta u_{p,i} = \begin{bmatrix} \Delta u_{1(k-i)} \\ \Delta u_{2(k-i)} \\ \vdots \\ \Delta u_{q(k-i)} \end{bmatrix}; \quad i = 1, 2, \dots, N - P - 1$$

$$U_p = \begin{bmatrix} u_{p,1} \\ u_{p,2} \\ \vdots \\ u_{p,P} \end{bmatrix}; \quad \text{and} \quad u_{p,i} = \begin{bmatrix} u_{1(k-j)} \\ u_{2(k-j)} \\ \vdots \\ u_{q(k-j)} \end{bmatrix}; \quad i = 1, 2, \dots, P; \quad j = N - i$$

$$G = \begin{bmatrix} G_1 & 0 & \dots & 0 \\ G_2 & G_1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ G_{p-1} & G_{p-2} & \dots & \dots \\ G_p & G_{p-1} & \dots & G_{p-M+1} \end{bmatrix}; \quad \text{and} \quad G_i = \begin{bmatrix} S_{11,i} & S_{12,i} & \dots & S_{1q,i} \\ S_{21,i} & \dots & \dots & S_{2q,i} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1,i} & \dots & \dots & S_{mq,i} \end{bmatrix};$$

$i = 1, 2, \dots, P$

$$E = \begin{bmatrix} E_2 & E_3 & \dots & E_{N-1} \\ E_3 & E_4 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ E_p & E_{p+1} & \ddots & \vdots \\ E_{p+1} & E_{p+2} & \dots & 0 \end{bmatrix}; \quad \text{and} \quad E_i = \begin{bmatrix} S_{11,i} & S_{12,i} & \dots & S_{1q,i} \\ S_{21,i} & S_{22,i} & \dots & S_{2q,i} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1,i} & S_{m2,i} & \dots & S_{mq,i} \end{bmatrix};$$

$i = 2, 3, \dots, N - 1$

$$F = \begin{bmatrix} \overbrace{F_N}^{P \text{ times}} & 0 & \dots & 0 \\ 0 & F_N & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & F_N \end{bmatrix}; \quad \text{and} \quad F_N = \begin{bmatrix} S_{11,N} & S_{12,N} & \dots & S_{1q,N} \\ S_{21,N} & S_{22,N} & \dots & S_{2q,N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1,N} & S_{m2,N} & \dots & S_{mq,N} \end{bmatrix}$$

Bias Correction:

If the prediction equation 2.10 is used as is, cumulative effects of model inaccuracy and unmeasured disturbances can lead to inaccurate predictions. The prediction accuracy of the equation is improved by utilizing the latest measurement in the predictions in a strategy, called output feedback (Qin and Badgwell, 2003).

For a SISO system, a residual signal $b_{(k)}$ can be defined as the difference between the latest measurement $y_{(k)}$ and the predicted output $\hat{y}_{(k)}$. That is:

$$b_{(k)} = y_{(k)} - \hat{y}_{(k)} \quad \dots(2.11)$$

The corrected free prediction \tilde{y}^f is then given as:

$$\tilde{y}^{fr} = y^{fr} + b_k \quad \dots(2.12)$$

In vector-matrix notation, eqn. (2.12) is written as:

$$\tilde{\mathbf{Y}}^{fr} = \mathbf{Y}^{fr} + I_{mP} b_k \quad \dots(2.13)$$

where

$$I_{mP} = \underbrace{[I_m \quad I_m \quad \dots \quad I_m]^T}_{P \text{ times}}, \text{ and } I_m \text{ is an identity matrix of size } m$$

The corrected predicted output equation now becomes:

$$\tilde{\mathbf{Y}} = \mathbf{G}\Delta\mathbf{U} + \tilde{\mathbf{Y}}^{fr} \quad \dots(2.14)$$

2.5.2 Prediction Algorithm for Transfer Function MPC

The derivation of the prediction algorithm for Transfer function MPC depends on a variant of the difference equation 2.2. The auto regressive with exogenous input (ARX) variant of equation 2.2 without the white noise (e_k) term for SISO system results in a simple difference equation as:

$$A(z^{-1})y_{(k)} = z^{-d}B(z^{-1})u_k \quad \dots(2.15)$$

For a MIMO system, each input-output channel may have a different input delay and different poles and zeros. This makes it slightly complicated. The difference equation of a typical MIMO system of m outputs and q inputs can be written as:

$$\begin{bmatrix} y_{1(k+1)} \\ y_{2(k+1)} \\ \vdots \\ y_{m(k+1)} \end{bmatrix} = \begin{bmatrix} z^{-d_{11}} \frac{B_{11}(z^{-1})}{A_{11}(z^{-1})} & z^{-d_{12}} \frac{B_{12}(z^{-1})}{A_{12}(z^{-1})} & \dots & z^{-d_{1q}} \frac{B_{1q}(z^{-1})}{A_{1q}(z^{-1})} \\ z^{-d_{21}} \frac{B_{21}(z^{-1})}{A_{21}(z^{-1})} & z^{-d_{22}} \frac{B_{22}(z^{-1})}{A_{22}(z^{-1})} & \dots & z^{-d_{2q}} \frac{B_{2q}(z^{-1})}{A_{2q}(z^{-1})} \\ \vdots & \vdots & \ddots & \vdots \\ z^{-d_{m1}} \frac{B_{m1}(z^{-1})}{A_{m1}(z^{-1})} & z^{-d_{m2}} \frac{B_{m2}(z^{-1})}{A_{m2}(z^{-1})} & \dots & z^{-d_{mq}} \frac{B_{mq}(z^{-1})}{A_{mq}(z^{-1})} \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_q(k) \end{bmatrix} \quad \dots(2.16)$$

where

$$A_{ij}(z^{-1}) = 1 + a_{1,ij}z^{-1} + a_{2,ij}z^{-2} + \dots + a_{na,ij}z^{-na,ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, q$$

$$B_{ij}(z^{-1}) = b_{0,ij} + b_{1,ij}z^{-1} + b_{2,ij}z^{-2} + \dots + b_{nb,ij}z^{-nb,ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, q$$

d_{ij} are the factored out delays in the concerned input-output channel

The compact vector-matrix form of difference equation 2.16 can be written as:

$$\mathbf{Y}_{(k+1)} = \mathbf{\Phi} \mathbf{\Psi} \quad \dots(2.17)$$

where

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_1 & 0 & \dots & 0 \\ 0 & \mathbf{\Phi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{\Phi}_m \end{bmatrix}$$

$$\mathbf{\Psi} = [\mathbf{\Psi}_1 \quad \mathbf{\Psi}_2 \quad \dots \quad \mathbf{\Psi}_m]^T$$

$$\mathbf{\Phi}_i = [\varphi_{i1}, \varphi_{i2}, \dots, \varphi_{iq}]$$

$$\mathbf{\Psi}_i = [\psi_{i1}, \psi_{i2}, \dots, \psi_{iq}]$$

$$\varphi_{ij} = [-a_{1,ij}, -a_{2,ij}, \dots, -a_{na,ij}, n_{z1,ij}, n_{z2,ij}, b_{0,ij}, b_{1,ij}, \dots, b_{nb,ij}, n_{z3,ij}]$$

$$\psi_{ij} = [y_{i(k)}, y_{i(k-1)}, \dots, y_{i(k-n_{ym})}, u_{j(k)}, u_{j(k-1)}, \dots, u_{j(k-n_{um})}]$$

$$n_{ym} = \max(a_{na,ij}), \quad i = 1, 2, \dots, m, j = 1, 2, \dots, q$$

$$n_{um} = n_{ym} + \max(d_{ij}), \quad i = 1, 2, \dots, m, j = 1, 2, \dots, q$$

$$c_{1,ij} = n_{ym} - na, ij$$

$$c_{2,ij} = na, ij + d_{ij} - nb, ij - 1$$

$$c_{3,ij} = n_{um} - (na, ij + d_{ij})$$

$$n_{z1,ij} = \overbrace{[0, 0, \dots, 0]}^{c_{1,ij} \text{ times}}$$

$$n_{z2,ij} = \overbrace{[0, 0, \dots, 0]}^{c_{2,ij} \text{ times}}$$

$$n_{z3,ij} = \overbrace{[0, 0, \dots, 0]}^{c_{3,ij} \text{ times}}$$

The difference equation 2.17 can be used to calculate the free response, the step response and the model response over a given horizon. The difference in the calculation (and the results) is the nature of past inputs that are used in the equation. In calculating the free response vector \mathbf{Y}^{fr} for example, all future inputs remain constant at the most previous value within the prediction horizon. That is $u_{(k-i)} = u_{(k)}$ for $i = 1, 2, \dots, (k - n_u - 1)$. Using equation 2.17 to calculate unit step response simply requires that all future inputs remain constant at the most previous value, which is 1. That is $u_{(k-i)} = u_{(k)} = 1$ for $i = 1, 2, \dots, (k - n_u - 1)$

The prediction equation for Transfer function MPC within a prediction horizon P and control horizon M can therefore be written as:

$$\hat{\mathbf{Y}} = \mathbf{G}\Delta\mathbf{U} + \underbrace{\Phi\Psi}_{\mathbf{Y}^{fr}} \quad \dots(2.18)$$

The terms \mathbf{G} and $\Delta\mathbf{U}$ are as defined in the case of the step response MPC in equation 2.10 above.

After the predicted output vector \hat{Y} is calculated from equation 2.18, the calculation of the corrected output prediction \tilde{Y} is based on similar equations as those described for step response MPC.

2.5.3 Prediction Algorithm for State Space MPC

The derivation of the output prediction algorithm is much easier for state space MPC.

Performing a difference operation on both sides of the state space equations above

(equation 2.3) with $D = 0$ gives:

$$\Delta x_{(k+1)} = A\Delta x_{(k)} + B\Delta u_{(k)} \quad \dots(2.19)$$

where

$$\Delta x_{(k+1)} = x_{(k+1)} - x_{(k)} \quad \dots(2.20)$$

$$\Delta u_{(k)} = u_{(k)} - u_{(k-1)} \quad \dots(2.21)$$

Similarly, equation (2.4) can be written as:

$$\Delta y_{(k+1)} = C\Delta x_{(k+1)} \quad \dots(2.22)$$

Combining equations (2.20), (2.21) and (2.22) and rearranging, we have:

$$y_{(k+1)} - y_{(k)} = CA\Delta x_{(k)} + CB\Delta u_{(k)} \quad \dots(2.23)$$

For a generalised system of m outputs, q inputs and for which the dimension of the state

vector x is n , equations 2.23 can be compactly written in state-space form as (Wang,

2009):

$$\overbrace{\begin{bmatrix} \Delta x_{(k+1)} \\ y_{(k+1)} \end{bmatrix}}^{x_{g(k+1)}} = \overbrace{\begin{bmatrix} A & 0_{n,m} \\ CA & I_m \end{bmatrix}}^{A_g} \overbrace{\begin{bmatrix} \Delta x_{(k)} \\ y_{(k)} \end{bmatrix}}^{x_{g(k)}} + \overbrace{\begin{bmatrix} B \\ CB \end{bmatrix}}^{B_g} \Delta u_{(k)} \quad \dots(2.24)$$

$$y_{(k)} = \overbrace{\begin{bmatrix} 0_{n,m} & I_m \end{bmatrix}}^{C_g} \begin{bmatrix} \Delta x_{(k)} \\ y_{(k)} \end{bmatrix} \quad \dots(2.25)$$

where:

$0_{n,m}$ is an all zero elements matrix of n rows and m columns. I_m is an identity matrix of dimension m .

The three matrices (A_g, B_g, C_g) combine to form the augmented model which is used in the design of state-space MPC. The augmented model is used with a new state matrix $x_{g(k+1)}$.

Using the augmented model, the future state variables for all sampling intervals within a prediction horizon can be obtained as:

$$x_{g(k+1)} = A_g x_{g(k)} + B_g \Delta u_{(k)}$$

$$x_{g(k+2)} = A_g^2 x_{g(k)} + A_g B_g \Delta u_{(k)} + B_g \Delta u_{(k+1)}$$

⋮

$$x_{g(k+P)} = A_g^P x_{g(k)} + A_g^{P-1} B_g \Delta u_{(k)} + A_g^{P-2} B_g \Delta u_{(k+1)} + \dots + A_g^{P-M} B_g \Delta u_{(k+M-1)}$$

Using the predicted state variables, the predicted output is given as:

$$y_{(k+1)} = C_g A_g x_{g(k)} + C_g B_g \Delta u_{(k)}$$

$$y_{(k+2)} = C_g A_g^2 x_{g(k)} + C_g A_g B_g \Delta u_{(k)} + C_g B_g \Delta u_{(k+1)}$$

⋮

$$y_{(k+P)} = C_g A_g^P x_{g(k)} + C_g A_g^{P-1} B_g \Delta u_{(k)} + C_g A_g^{P-2} B_g \Delta u_{(k+1)} + \dots + C_g A_g^{P-M} B_g \Delta u_{(k+M-1)}$$

The prediction equation for state space MPC can therefore be written in vector-matrix form as:

$$\hat{Y} = \Theta \Delta U + \underbrace{\Gamma X_g}_{yfr} \quad \dots(2.26)$$

where

$$\Theta = \begin{bmatrix} C_g B_g & 0 & 0 & \dots & 0 \\ C_g A_g B_g & C_g B_g & 0 & \dots & 0 \\ C_g A_g^2 B_g & C_g A_g B_g & C_g B_g & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_g A_g^{P-1} B_g & C_g A_g^{P-2} B_g & C_g A_g^{P-3} B_g & \dots & C_g A_g^{P-M} B_g \end{bmatrix} \quad \dots(2.27)$$

$$\Gamma = \begin{bmatrix} CA \\ CA^2 \\ CA^3 \\ \vdots \\ CA^P \end{bmatrix} \quad \dots(2.28)$$

As can be seen from equation 2.26 the predicted free response vector \hat{Y}^{fr} can be calculated very easily compared to the other two MPC formulations.

State Space MPC and State Estimation: Observer Design

MPC State Space formulation usually includes an observer to estimate the unknown states from process measurements. This is because not all the state variables are measurable in reality due to the fact that the states may be inaccessible or that the measured outputs consist of some linear combinations of states. A state observer can also act like a noise filter to reduce the effect of noise on the measurement.

In the augmented state space MPC design, the observer estimates of state vector $\hat{x}_{g(k+1)}$ is used instead of $x_{g(k+1)}$. The estimated state vector is obtained by using an observer gain matrix, L , and output variable feedback as (Wang, 2009)

$$\hat{x}_{gk+1} = \overbrace{A_g \hat{x}_{g(k)} + B_g \Delta u_{(k)}}^{\text{model}} + \overbrace{L(y_k - C_m \hat{x}_{g(k)})}^{\text{corrected term}} \quad \dots(2.29)$$

The error state $\tilde{x}_{g(k+1)}$ is given as:

$$\tilde{x}_{gk+1} = A_g \tilde{x}_{g(k)} - LC_g \tilde{x}_{g(k)} \quad \dots(2.30)$$

The observer gain matrix L that makes $\tilde{x}_{gk+1} \rightarrow 0$ as $k \rightarrow \infty$ in equation 2.30 is used in the state estimation equation 2.29. Apart from ensuring that $\tilde{x}_{gk+1} \rightarrow 0$, L must be chosen such that the observer dynamics is much faster than the system. The observer can be designed independently and offline. There are a few methods for designing the observer. A common one is the pole placement method. In the pole placement method the Eigenvalues (poles_a) of the augmented state matrix A_g are found. Then a new set of poles (poles_b) is derived such that the poles are located father to the left of the system dominant pole(s), to ensure an observer with very fast dynamics.

The corrected predicted outputs equation for state space MPC then becomes:

$$\tilde{Y} = \Theta \Delta U + \Gamma \hat{X}_g \quad \dots(2.31)$$

2.6 Solution of Unconstrained MPC

If there are no input and output constraints, the set of optimal control actions that minimises the MPC objective function of equation 2.5 can be obtained by finding the least squares solution of the objective function.

By using the reference trajectory in equation 2.5 instead of a set-point, the new quadratic objective function becomes:

$$\min_{\Delta U} J = (\mathbf{r} - \tilde{Y})^T \mathbf{Q} (\mathbf{r} - \tilde{Y}) + (\Delta U)^T \mathbf{R} (\Delta U) \quad \dots(2.32)$$

By substituting the corrected predicted outputs equation 2.14 (for step response MPC, or its equivalent for transfer function MPC, or its state space equivalent, which is equation 2.31), into the objective function of equation 2.32, we have:

$$\min_{\Delta U} J = -2\Delta U^T \mathbf{G}^T \mathbf{Q} \mathbf{E} + \Delta U^T (\mathbf{G}^T \mathbf{Q} \mathbf{G} + \mathbf{R}) \Delta U + \mathbf{E}^T \mathbf{E} \quad \dots(2.33)$$

$$\text{where } \mathbf{E} = (\mathbf{r} - \tilde{Y}^r)$$

The first derivative of the cost function J (equation 2.33) gives:

$$\frac{\partial J}{\partial \Delta U} = -2\mathbf{G}^T \mathbf{Q} \mathbf{E} + 2(\mathbf{G}^T \mathbf{Q} \mathbf{G} + \mathbf{R}) \Delta U = 0 \quad \dots(2.34)$$

Then the optimum control sequence for unconstrained MPC then becomes:

$$\Delta U = (\mathbf{G}^T \mathbf{Q} \mathbf{G} + \mathbf{R})^{-1} \mathbf{G}^T \mathbf{Q} \mathbf{E} = \mathbf{H}^{-1} \mathbf{V} \quad \dots(2.35)$$

where

$$\mathbf{H} = (\mathbf{G}^T \mathbf{Q} \mathbf{G} + \mathbf{R})$$

$$\mathbf{V} = \mathbf{G}^T \mathbf{Q} \mathbf{E}$$

The matrix \mathbf{H}^{-1} is called the Hessian matrix.

2.7 MPC and Constraints

When the MPC is constrained, the optimum control moves can no longer be obtained as a least squares solution. Instead the optimum control moves are obtained by quadratic programming (QP). Constraints can be imposed on the amplitude of the input signal $u_{(k)}$, on the rate of change of input signal $\Delta u_{(k)}$ or on the output $y_{(k)}$. Input signal constraints are most common (Wang, 2009), because they represent the saturation characteristics of actuators (valve opening, flowrate, voltage etc.), and hence are usually hard constraints. Apart from restricting the rate at which an input signal can vary, a constraint on the rate of change of input signal can also be used to limit the direction of movement. Again they are usually specified as hard constraints. Output constraints, are usually softened, to allow for violation if and when necessary, because un-softened (hard) output constraints can result in stability issues. One way of implementing soft constraints is to add slack variables to convert constraints inequalities into equalities.

Inequality constraints can be written generally as:

$$\mathbf{W} \Delta U \leq \eta \quad \dots(2.36)$$

where

W is a matrix which, when multiplied by rate of change of input ΔU , gives the rate of change for every input and output element within the prediction and control horizons.

η is the vector containing the upper and lower boundaries (constraints) of the rate of change of input, input and output.

Matrix W and vector η have three terms each:

$$W = \begin{bmatrix} W_{du} \\ W_u \\ W_y \end{bmatrix} \quad \dots(2.37)$$

$$\eta = \begin{bmatrix} \eta_{du} \\ \eta_u \\ \eta_y \end{bmatrix} \quad \dots(2.38)$$

The terms in equations 2.37 are defined as:

$$W_{du} = \begin{bmatrix} I_{Mq} \\ -I_{Mq} \end{bmatrix} \quad \dots(2.39)$$

$$W_u = \begin{bmatrix} W_{uu} \\ -W_{uu} \end{bmatrix} \quad \dots(2.40)$$

$$W_y = \begin{bmatrix} \phi \\ -\phi \end{bmatrix} \quad \dots(2.41)$$

where

W_{du} is the multiplying matrix for rate of change of input.

W_u is the multiplying matrix for input

ϕ is the multiplying matrix for each output defined in the prediction equations

above ($\phi = G$ as in step response and transfer function MPC, and $\phi = \Theta$ as in state space MPC).

I_{Mq} is identity matrix of dimension Mq

$$W_{uu} = \begin{matrix} M \text{ (with } Mq \text{ elements)} \\ \left[\begin{array}{ccccc} I_q & 0 & 0 & \cdots & 0 \\ I_q & I_q & 0 & 0 & 0 \\ I_q & I_q & I_q & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_q & I_q & \cdots & I_q & I_q \end{array} \right] \end{matrix}$$

I_q is identity matrix of dimension q

The terms in equations 2.38 are defined as:

$$\eta_{du} = Z_{du}\omega_{du} \quad \dots(2.42)$$

$$\eta_u = Z_u\omega_u + T_u\gamma_u \quad \dots(2.43)$$

$$\eta_y = Z_y\omega_y + T_y\gamma_y \quad \dots(2.44)$$

where:

η_{du} is the vector containing the upper and lower bounds of rate of change of input

η_u is the vector containing the upper and lower bounds of input

η_y is the vector containing the upper and lower bounds of output

$$Z_{du} = \begin{bmatrix} Z_{uu} & 0_{Mq,q} \\ 0_{Mq,q} & -Z_{uu} \end{bmatrix}$$

$$Z_u = \begin{bmatrix} Z_{uu} & 0_{Mq,q} \\ 0_{Mq,q} & -Z_{uu} \end{bmatrix}$$

$$Z_y = \begin{bmatrix} Z_{yy} & 0_{Pm,m} \\ 0_{Pm,m} & -Z_{yy} \end{bmatrix}$$

$$Z_{uu} = \overbrace{[I_q, I_q, \dots, I_q]}^{M \text{ times}}^T$$

$$Z_{yy} = \overbrace{[I_m, I_m, \dots, I_m]}^{P \text{ times}}^T$$

$$\omega_{du} = [\Delta U_{1,max}, \Delta U_{2,max}, \dots, \Delta U_{q,max}, \Delta U_{1,min}, \Delta U_{2,min}, \dots, \Delta U_{q,min}]^T$$

$$\omega_u = [U_{1,max}, U_{2,max}, \dots, U_{q,max}, U_{1,min}, U_{2,min}, \dots, U_{q,min}]^T$$

$$\omega_y = \omega_{y1} - \omega_{y2}$$

$$\omega_{y1} = [Y_{1,max}, Y_{2,max}, \dots, Y_{m,max}, Y_{1,min}, Y_{2,min}, \dots, Y_{m,min}]^T$$

$$\omega_{y2} = -[Y_{1,k+1}^S, Y_{2,k+1}^S, \dots, Y_{m,k+1}^S, Y_{1,k+1}^S, Y_{2,k+1}^S, \dots, Y_{m,k+1}^S]^T$$

$$T_u = \begin{bmatrix} -Z_{uu} \\ Z_{uu} \end{bmatrix}$$

$$T_y = \begin{bmatrix} -\psi \\ \psi \end{bmatrix}$$

$$\gamma_u = [U_{1,k-1}, \Delta U_{2,k-1}, \dots, \Delta U_{q,k-1}]^T$$

$$\gamma_y = x_{g,(k+1)}$$

$$Z_{uu} = \overbrace{[I_q, I_q, \dots, I_q]^T}^{M \text{ times}}$$

I_m is identity matrix of dimension m

$0_{Mq,q}$ is all elements zero matrix of dimension $Mq \times q$

$0_{Pm,m}$ is all elements zero matrix of dimension $Pm \times m$

2.8 Numeric Solution of MPC QP Problem: Hildreth's QP Procedure

The objective function (equation 2.33), ignoring the last term on the right hand side of the equation, is generally written as:

$$J = \frac{1}{2} \Delta \mathbf{U}^T H \Delta \mathbf{U} + \Delta \mathbf{U}^T V \quad \dots(2.45)$$

which is subject to the inequality constraint equation 2.37

Equations (2.45) can be written in Lagrangian form (Wang, 2009) as:

$$J = \frac{1}{2} \Delta \mathbf{U}^T H \Delta \mathbf{U} + \Delta \mathbf{U}^T V + \lambda^T (W \Delta \mathbf{U} - \eta) \quad \dots(2.46)$$

The elements of the vector λ are called *Lagrange multipliers*. The optimal λ and ΔU are defined by:

$$\lambda = -(WH^{-1}W^T)^{-1}(\eta + WH^{-1}V) \quad \dots(2.47)$$

$$\Delta \mathbf{U} = -H^{-1}(W^T \lambda + V) = -H^{-1}V - H^{-1}W^T \lambda \quad \dots(2.48)$$

The global solution that gives the optimal $\Delta \mathbf{U}$ without constraints is the first term of equation (2.48), i.e.

$$\Delta \mathbf{U}_{global} = -H^{-1}V \quad \dots(2.49)$$

The correction term for the optimal ΔU due to the equality constraints is the second term of equation (2.48) and is given by:

$$\Delta \mathbf{U}_{corrected} = -H^{-1}W^T \lambda \quad \dots(2.50)$$

Using the *Primal-Dual* method, it was also shown that the equivalent dual quadratic programming problem of the original objective function is given as (Wang, 2009):

$$\min_{\lambda \geq 0} \left(\frac{1}{2} \lambda^T \tau \lambda + \lambda^T \mu + \frac{1}{2} \eta^T H^{-1} \eta \right) \quad \dots(2.51)$$

where:

$$\tau = WH^{-1}W^T \quad \dots(2.52)$$

$$\mu = \eta + WH^{-1}V \quad \dots(2.53)$$

The set of optimal Lagrange multipliers that minimize the Dual objective function subject to $\lambda \geq 0$ are denoted as λ_{act} , and the corresponding constraints are described by W_{act} . With the values of λ_{act} and W_{act} the primal variable vector is:

$$\Delta \mathbf{U} = -H^{-1}V - H^{-1}W_{act}^T \lambda_{act} \quad \dots(2.54)$$

Hildreth's quadratic programming method (Luenberger, 1984, Wismer and Chattergy, 1978) for solving the dual problem (2.51) is expressed explicitly as:

$$\lambda_i^{m+1} = \max(0, w_i^{m+1}) \quad \dots(2.55)$$

with

$$w_i^{m+1} = -\frac{i}{h_{ii}} \left[k_i + \sum_{j=1}^{i-1} h_{ij} \lambda_j^{m+1} + \sum_{j=i+1}^n h_{ij} \lambda_j^m \right] \quad \dots(2.56)$$

where the scalar h_{ij} is the *ij-th* element in the matrix $\tau = WH^{-1}W^T$ and k_i is the *i-th* element in the vector $\mu = \eta + WH^{-1}V$. Also note that in (2.56) there are two sets of λ values in the computation: one involves λ^m and the other involves the updated λ^{m+1}

Because the converged λ^* vector contains either zero or positive values of Lagrange multipliers, we have

$$\lambda_{act}^* = -(W_{act}H^{-1}W_{act}^T)^{-1}(\eta_{act} + W_{act}H^{-1}V) \quad \dots(2.57)$$

Hence,

$$\Delta U = -H^{-1}V - H^{-1}W_{act}^T\lambda_{act}^* \quad \dots(2.58)$$

2.9 MPC and Economic Optimisation

A unique steady state unconstrained solution exists when MPC is implemented on a square system (number of manipulated variables is equal to the number of controlled variables). Such a unique steady state solution does not exist when there are more manipulated variables than controlled variables (non-square systems), because it results in a system with extra degrees of freedom. An economic optimisation problem is used in the presence of extra degrees of freedom to specify the economic compromises among the process variables. The economic optimisation calculates the ideal resting values of the MVs, or alternatively forces the system to be square by defining “pseudo CVs that are identical to MVs”. For the case where there are too few degrees of freedom, the economic optimisation either allows only some CVs (equal to the number of available MVs) to be controlled to set points or for all CVs to be “out of control” (that is with offsets) (Froisy, 1994).

2.10 The MPC and the operators in a Process Plant

Though MPC may be used to directly manipulate process valve position (Darby and Nikolaou, 2012), the standard practice in modern processing plants is to employ MPC as a supervisory controller, as a part of multilayer hierarchy of process control (Darby and Nikolaou, 2012, Qin and Badgwell, 2003, Tatjewski, 2010). A simple version of this hierarchical arrangement is shown in figure 2.3 below. The MPC computes the set-points for lower level regulatory control. The MPC requires information about the

optimal steady state settings for the processes, which is supplied by an optimizer that performs linear programming. The plant operators interact with the MPC (and the process, DCS, optimizer) through a station in the control room. Communications between layers are channelled through input/output devices, which can be integral with the controllers, or located remotely.

The human interface station provides windows for operation and monitoring of processes. It has pre-programmed process plan and process flow diagrams (PFD) to ensure easy visualisation of the configurations and operations occurring in the process. Through the interfaces, which have systems of alarms, trends, messages, faceplates and tuning windows, the operators can tune and monitor MPC performance.

2.11: The trends of a control system

The MPC variables shown in figure 2.2 are still used in the MPC prediction algorithm, since the prediction neither discriminates between direct or supervisory MPC nor accounts for the variables between regulatory controllers and valves. However in a real industrial situation, where MPC is used as supervisory controller like figure 2.3, where many process trends have to be monitored, a more expanded definition to figure 2.2 is preferred. Figure 2.4 below shows the important variables in MPC as supervisory control.

In figure 2.4, rv is the MPC reference signal, while sv is the set-point to the regulatory controller (the output from the MPC). The signal pv is the process controlled output which is fed back to the regulatory controller. The valve is actuated by signal ov , while mv is the actual output of the valve. In set-point control, the process output variable cv is designed to be controlled to the reference signal rv .

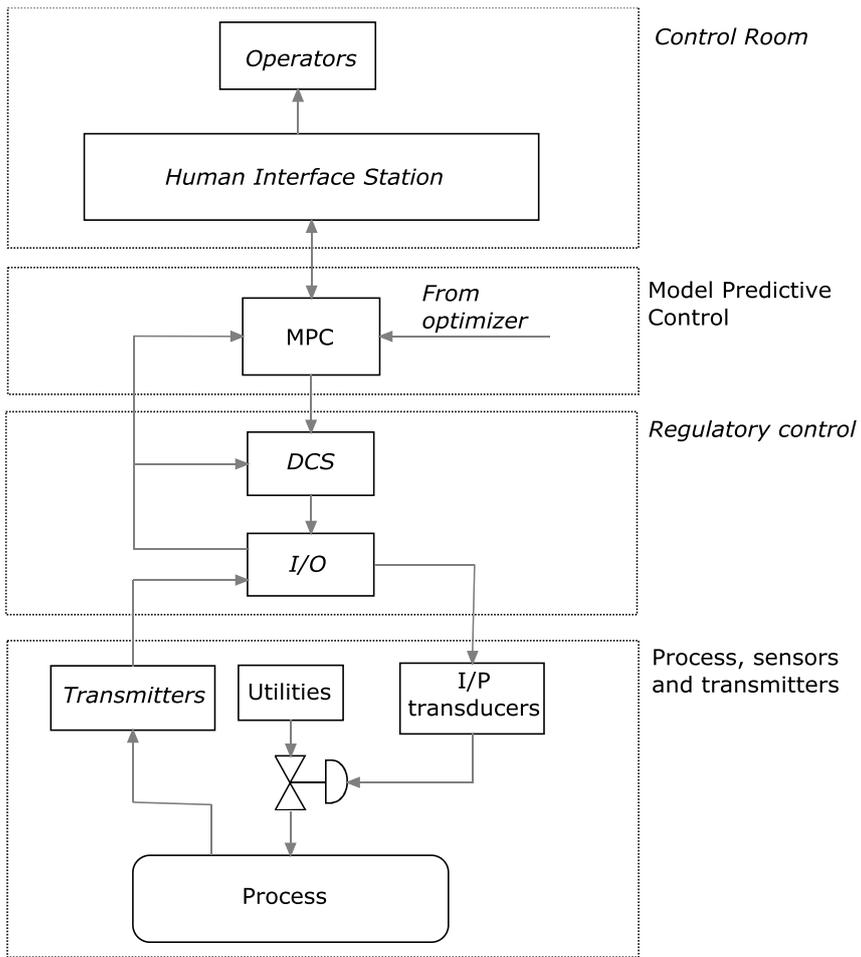


Figure 2.3: Typical process plant control hierarchy

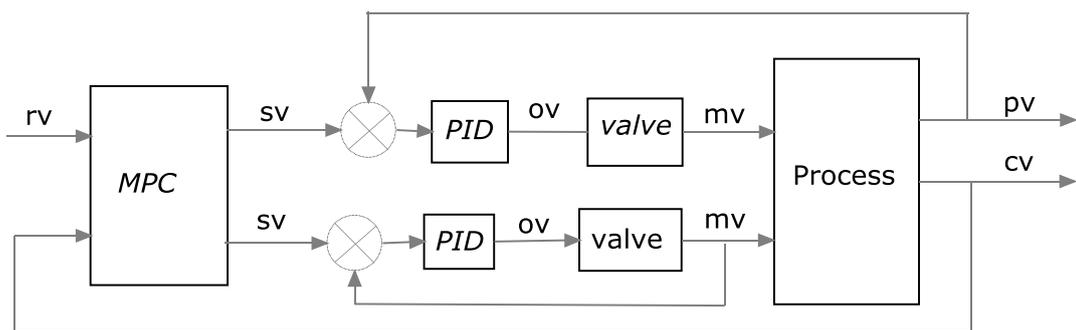


Figure 2.4: Main variables of MPC as supervisory control

2.12 Developments in MPC Performance Assessment and Maintenance

The pioneering works of Harris (1989) and Desborough and Harris (1992) in the development of methods and metrics for univariate control systems performance assessment and the extension of the methods to multivariate systems by Ender (1993), (Kozub and Garcia, 1993), Harris et al. (1996), Huang et al. (1996), Tyler and Morari (1996), Huang et al. (1997), Huang and Shah (1999), Yu and Qin (2008a), Yu and Qin (2008b) and Ko and Edgar (2000), have contributed immensely to understanding of control assessment generally, and have also aroused the interests of many more researchers employing various additional approaches to control assessment. Some of the performance assessment metrics resulting from these studies include: Harris index, historical benchmark, extended horizon performance index, minimum variance benchmark, generalised minimum variance, and linear quadratic Gaussian (LQG).

All of the above methods use real process data and one or more statistical methods to obtain indices through which performance of controllers are assessed. Statistical methods like principal component analysis (PCA), partial least squares (PLS), independent component analysis (ICA), Fisher discriminant analysis (FDA) are commonly used. The use of subspace identification approach in control performance analysis by Bezergianni and Georgakis (2003) and Ding et al. (2009) have also been reported. Yin et al. (2012) employed the above statistics, in addition to dynamic principal component analysis (DPCA), modified partial least squares (MPLS), modified independent component analysis (MICA) and total projection to latent structure (TPLS), to obtain two indices each: the false alarm rate (FAR) and the fault detection rate (FDR) which were used to assess the performance the Tennessee Eastman process.

A number of publications have emerged which are devoted specifically to MPC systems. These include the works of Patwardhan and Jay (1997), Ko and Edgar (2001), Patwardhan et al. (2002), Patwardhan et al. (2002), Loquasto III and Seborg (2003), Julien et al. (2004), Lennox (2005), Zhang and Li (2006), Agarwal et al. (2007a), , Agarwal et al. (2007b), Xu et al. (2007), and AlGhazzawi and Lennox (2009). As with the previous assessment methods these also employ statistical methods in the analysis of process data to obtain performance measures.

The diagnostic tool by Patwardhan and Jay (1997) requires information about the dynamic model of the MPC, in addition to process data (past and present), to estimate statistical measure for detecting degradation in MPC. In Patwardhan et al. (2002), the proposed performance measure is a ratio of the achieved objective function to that of the design value of the MPC. The proposed method was demonstrated on an industrial QDMC application on a recycle surge drum level control. This method also requires the knowledge of the internal model of the MPC. Loquasto III and Seborg (2003) use principal component analysis (PCA) and distance similarity factors to monitor MPC performance, where several PCA pattern classifiers were developed to monitor the control system, and to identify abnormal behaviour. An MPC controlled Wood–Berry distillation column model was employed to demonstrate the application of the developed technique.

Xu et al. (2007) employed variance based performance assessment method, and using process data and steady state gains and linear matrix inequalities, to develop indices for evaluating the economic performance of MPC. The performance assessment of MPC through the study of the relationship among process variability, constraints, and probabilistic economic performance of MPC was the focus of the work of Agarwal

(2007a, 2007b). They used both probabilistic and Bayesian calculations to evaluate the economic performance of MPC.

Lennox (2005) showed that by using partial least square (PLS) to identify the dynamic model of a process, the T2 and SPE statistics obtained from such model can be used as fault detection and isolation tool for MPC. The use of multivariate statistical process control (MSPC) techniques as MPC condition monitoring tool was demonstrated in AlGhazzawi and Lennox (2009). The PCA and PLS models obtained from the study were used to identify abnormalities and their causes in an application on condensate fractionation process.

The use of the MPC assessment and monitoring methods cited above and the computation of their assessment metrics involve heavy statistical analysis. The works of Harrison and Qin (2009), Lee et al. (2008) and Gabriele et al. (2013) are further examples of the huge computational and statistical involvement of these methods. Interpretations of the statistics from many of these methods require considerable knowledge of process and MPC, which mainly senior process or control engineers with considerable experience in MPC applications possess. It may be argued that the average operator may be more interested in practical issues like what does a trend suggest without the heavy statistics. Arguably, many MPC operators (the front-line users such as control room operators and junior engineers) would prefer that can explain abnormalities and possible recovery methods being bothered with heavy statistics.

One of the earliest reported cases of the application of some of the methods and metrics cited above to the performance analysis and troubleshooting of an industrial model predictive control system is by Kadali et al. (1999). In the work, by analysing the variance of routine process operating data, problems of model mismatch and

identification of disturbances having significant effects on process variability were identified. Huang et al. (2000) adopted a completely data-driven approach in successfully trouble shooting an industrial MPC system. By adopting simple methods of process trends comparison devoid of complex statistics, and using none of the above metrics, the source of poor performance of the MPC system was identified. The study reflects very well how operators might prefer to address diagnosis of degradation in their control systems.

Another practical paper is that by Gao et al. (2003). Process data from two industrial MPCs (before and after the MPC implementation) were analysed to obtain and compare several different measures of multivariate controller performance. The comparison was used to diagnose model-plant-mismatch as source of poor performance in one of the controllers. Apart from the minimal multivariate controller performance metrics used in this study, the need for careful observation of actual process trends was emphasised.

The work of Schäfer and Cinar (2004) was not based on actual industrial process data. They proposed two simple performance assessment indices, the ratio of historical and achieved performance, for monitoring, and the ratio of design and achieved performance, for diagnosis. They used these indices to diagnose poor performances in simulated case studies. The diagnoses at best could only point to a group of possible causes of poor performance, and not the actual cause isolation. The work of Badwe et al. (2009) focused solely on detection of model-plant mismatch in MPC applications. The significance of this study here is that it was demonstrated on data from an industrial process.

Jiang et al. (2011) used a combination of models quality evaluation and actual process data trends comparison to diagnose and identify the root causes of sub-optimal

performance industrial model predictive controllers. The work, which is devoid of complex statistics, has certain similarities to the methods of Huang et al. (2000) and Gao et al. (2003)

2.13 Process Model Identification

The existence of an accurate process model can be critical to the successful design and implementation of MPC. For real processes the mathematical models that describe their dynamics are usually obtained empirically, rather than through mathematical modelling that make use of physical and chemical laws. This is so because the processes involved are usually too complex for such mathematical modelling. Also system parameters may not remain constant throughout the lifetime of a plant. So the need to model can arise, not only at the design stage of the MPC but sometimes later, especially when the MPC begins to degrade and significant model-plant mismatch is suspected.

Many analytical and statistical techniques have been developed to identify process models from experimental data. Many methods are described in Eykhoff (1974), Goodwin and Payne (1977), Söderström and Stoica (1988) and Ljung (1999). A new form of process identification referred to as subspace identification has become popular more recently (Van Overschee and De Moor, 1997), (Katayama, 2005) and (Favoreel et al., 2000).

Plant identification begins by superimposing an appropriate excitation signal onto the input of a process that is currently running steadily, and then subjecting the resulting process output to appropriate analysis to obtain the process model. In the simplest form, plant identification is as described in figure 2.5 below. The excitation signal u_e is superimposed on the steady state input signal u_s to the plant running at its nominal operating point. For linear model identification, the plant nominal output y_s is

subtracted from the recorded experimental output y to obtain a deviation signal y_d that is processed to obtain the appropriate model. Then the transfer function of the plant is given as:

$$G_p(S) = \frac{y_d(s)}{u_d(s)} \quad \dots(2.59)$$

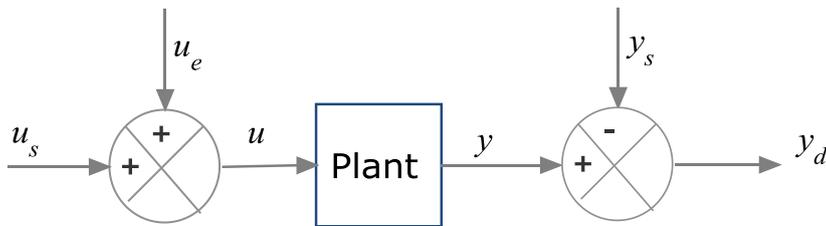


Figure 2.5: Excitation signal on open loop plant for identification

Common excitation signals are the step response signal and the pseudo-random binary sequence (PRBS) signal. PRBS is a periodic, deterministic signal with white noise like properties. Random signals are preferred as input sequences in many cases, because of their superb noise reduction properties when compared with a standard step response signal. Also some processes are better described by parametric models and responses generated from PRBS are more suitable for such analysis. More importantly using PRBS multiple excitation signals can be applied to a MIMO system simultaneously, instead of one input at a time. This can help to reduce significantly the total time required for the identification process.

The excitation signal must meet certain requirements, for the input-output signal obtained to be useful for identification purposes, because the excitation signal is subject to various constraints: the physical limits of the signal generator, the physical constraints of the process equipment and the dynamic range of the sensors. Guidelines like those given below need to be followed (Annus et al., 2012, Söderström and Stoica, 1988):

- 1) For a PRBS excitation signal, at least one of the pulses must be larger than the rising time of the plant to be identified. Physically this is realised by providing the estimates of the fastest (lowest rise time tr_L or time constant tc_L) and the slowest (highest rise time tr_H or time constant tc_H) of the process. These values are used to define the bandwidth for the input signal excitation as:

$$bw = [\omega_{min} \ \omega_{max}]$$

where

$$\omega_{min} = \frac{1}{\beta_s tc_H} \leq \omega \leq \frac{\alpha_s}{tc_L} = \omega_{max} \quad \dots(2.60)$$

Typical values of α_s and β_s are 2 and 3 respectively.

- 2) For a PRBS, the magnitude should not exceed a few percentage of the steady state control signal amplitude. If the amplitude is so large as to make the output perturb too far from equilibrium, it may result in nonlinearities, detrimental to the process and to the integrity of the resulting model.
- 3) For multivariable PRBS, there should be adequate lack of correlation between the signals applied at each channel.
- 4) The input signal should be selected such that the valve does not become saturated. Though this should be done bearing in mind that the larger the input magnitude, the smaller the asymptotic variance of the output signal, which also helps to reduce the signal to noise ratio.

Step response excitation is simple to apply. The resulting output deviation variable y_d can be fitted to appropriate and approximate first order or second order transfer functions by using a direct method. First order transfer function model ($G_{1p}(s)$) are defined by their delay time (Td), time constant (T) and steady state gain (K_p); second

order transfer function models ($G_{2p}(s)$) can be described in terms of their delay time (T_d), steady state gain (K_p), damping coefficient (ζ), and natural frequency (ω_n):

$$G_{1p}(s) = \frac{K_p e^{-T_d s}}{T_s s + 1} \quad \dots(2.61)$$

$$G_{2p}(s) = \frac{K_p e^{-T_d s} \omega_n^2}{s^2 + 2\zeta \omega_n s + \omega_n^2} \quad \dots(2.62)$$

It is often easy to observe the nature of response from step response test and estimate the approximate parameters that satisfy either a first or second order system. As Luyben (1989) pointed out, “probably 80 percent of all chemical engineering open loop processes can be model by a gain, dead time (delay time), and one lag”. A tutorial is given in Appendix A on how an approximate first order or second order transfer function model of a process may be obtained by applying a direct method to step response data.

Process model identification using input output data obtained when an excitation signal is applied to a plant is more appropriate when the plant is in open loop. Such data represents the plant’s true dynamics, because there is no correlation between the excitation and the output signals. However not all systems can be operated in open loop: some plants are unstable, contain integrators, or are just unsafe to be operated in open loop. In such situations the plant must be operated in closed loop. But input output data obtained when the plant is in closed loop is usually less informative, because the feedback signal causes a correlation between the excitation and the output signals, which tends to ‘mask’ the plant’s dynamics. There are at least three ways to identify a process model from closed loop data, depending on the information about the feedback signal and the controller dynamics. The three methods are described below.

- 1) **Direct identification:** in the direct method the excitation signal u_e is superimposed on the controller output (figure 2.6). Then the input u into the process and the output y from the process are measured, ignoring the effect of the feedback signal. Effectively the closed loop system is treated as an open loop system, which removes the need to know anything about the controller dynamics.

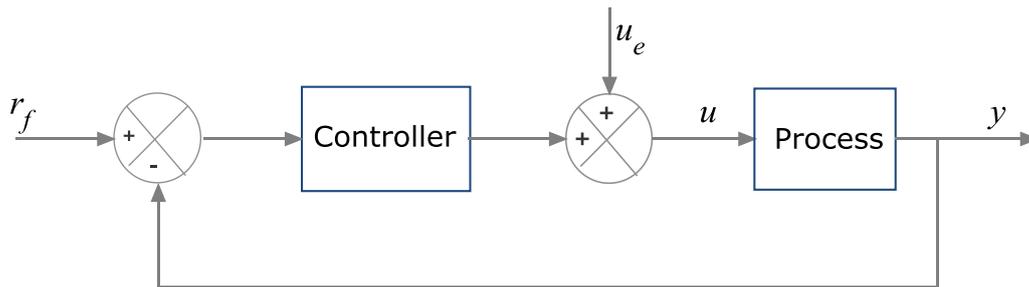


Figure 2.6: Excitation signal on closed loop plant for identification

Therefore the transfer function of the plant has the same expression as that of equation 2.59 above. Models obtained using this method are only approximate, although accuracy can be high if the signal to noise ratio is high, and the feedback contribution to the input spectrum is small.

- 2) **Indirect identification:** this method can only be used if the structure of the controller is known. The excitation signal u_e is applied to the controller output as before. The closed loop model is then obtained from measurements of the plant input u and the plant output y (figure 2.6). If the closed loop model transfer function obtained from identification is G_{CL} (using y and r_f), and the controller transfer function is known as G_C , the plant (open loop) transfer function G_p can be retrieved by using the expression:

$$G_p = \frac{G_{CL}}{1 - G_C G_{CL}} \quad \dots(2.63)$$

where

$$G_{CL} = \frac{y(s)}{r_f(s)}$$

3) **Joint input output identification:** in this method three measurements are required, the reference input r_f , the plant input u and the plant output y (figure 2.6). By using r_f and y to obtain the overall closed loop transfer function G_{CL} and then estimating the transfer function model G_{RU} relating r_f to u , the model of the plant can be estimated from:

$$G_P = \frac{G_{CL}}{G_{RU}} \quad \dots(2.64)$$

where

$$G_{RU} = \frac{u(s)}{r_f(s)}$$

Process model identification from closed loop data can be very demanding, especially for multivariable process. There are a number of publications which describe how process models can be estimated directly from closed loop data even when the information about the dynamics of the controllers used are not available. This is particularly attractive in the cases of plants under MPC. Most of those available are based on the subspace identification method and are most suited to identifying state space models in innovation form (Ljung and McKelvey, 1996) and (Van den Hof, 1998).

2.14 Conversion of Step Responses to State Space Models

Almost all current publications and training software are based entirely on state space MPC. Yet some practitioners still contend that MPC based on step response design is indispensable, as far as process industries are concerned. The major reason for this assertion is the belief that plants in process industries have dead times (delays), and that

although response patterns can be complicated, step response models can capture underlying behaviours. While it is obvious that step response models represent process time delay, response time and gain transparently (Wang, 2009), it is also true that equivalent state space model can be derived for step responses which replicate the complicated patterns of the step responses, including time delays (Maciejowski, 2002). This is so if the step responses have finite settling time (that is the step responses are bounded). The algorithm described below shows how this can be done.

Consider a multivariable system of m outputs, q inputs and represented by N step response coefficients. The step response coefficients matrix can be represented as:

$$\mathbf{S} = [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \cdots \quad \mathbf{S}_N]^T \quad \dots(2.65)$$

where

$$\mathbf{S}_k = \begin{bmatrix} \mathbf{S}_{k1,1} & \mathbf{S}_{k1,2} & \cdots & \mathbf{S}_{k1,q} \\ \mathbf{S}_{k2,1} & \mathbf{S}_{k2,2} & \cdots & \mathbf{S}_{k2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{km,1} & \mathbf{S}_{km,2} & \vdots & \mathbf{S}_{km,q} \end{bmatrix} \quad k = 1, 2, \dots, N$$

$\mathbf{S}_{ki,j}$ is the step response coefficient k of output i due to input j . The corresponding pulse response coefficients are represented as:

$$\mathbf{H} = [\mathbf{H}_1 \quad \mathbf{H}_2 \quad \cdots \quad \mathbf{H}_N]^T \quad \dots(2.66)$$

where

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{H}_{k1,1} & \mathbf{H}_{k1,2} & \cdots & \mathbf{H}_{k1,q} \\ \mathbf{H}_{k2,1} & \mathbf{H}_{k2,2} & \cdots & \mathbf{H}_{k2,q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{km,1} & \mathbf{H}_{km,2} & \vdots & \mathbf{H}_{km,q} \end{bmatrix} \quad k = 1, 2, \dots, N$$

$$\mathbf{H}_k = \mathbf{S}_k - \mathbf{S}_{k-1}$$

Using the pulse response coefficients, the block-Hankel matrix \mathbf{M} is obtained as:

$$\mathbf{M} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \cdots & \mathbf{H}_N \\ \mathbf{H}_2 & \mathbf{H}_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_N & \mathbf{0} & \vdots & \mathbf{0} \end{bmatrix} \quad \dots(2.67)$$

The dimension of \mathbf{M} is mN by qN . Using singular value decomposition (SVD), \mathbf{M} can be decomposed into three matrices \mathbf{U} , $\mathbf{\Sigma}$ and \mathbf{V} such that:

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad \dots(2.68)$$

Both \mathbf{U} and \mathbf{V} are unitary (orthogonal) matrices. The matrix $\mathbf{\Sigma}$ is a diagonal matrix, the same dimension as \mathbf{M} , and with nonnegative and nonzero leading diagonal elements $\sigma_1, \sigma_2, \sigma_3 \dots \sigma_r$ in decreasing order. These elements are the singular values of the matrix \mathbf{M} . The total number of the singular values (r) is the rank of the matrix \mathbf{M} . The rank r is equal to $\min(mN, qN)$.

The \mathbf{A} , \mathbf{B} and \mathbf{C} matrices of the state space model that match the original step responses exactly are obtained as:

$$\mathbf{B} = \text{1st } r \text{ rows and 1st } q \text{ columns of } \mathbf{\Sigma}^{1/2}\mathbf{V}^T$$

$$\mathbf{C} = \text{1st } m \text{ rows and 1st } r \text{ columns of } \mathbf{U}\mathbf{\Sigma}^{1/2}$$

$$\mathbf{A} = \text{1st } r \text{ rows and 1st } r \text{ columns of } \mathbf{\Sigma}^{-1/2}\mathbf{U}^T\mathbf{U}^s\mathbf{\Sigma}^{1/2}$$

where \mathbf{U}^s is the matrix \mathbf{U} shifted upwards by m rows.

The dimension of the state matrix \mathbf{A} is r by r

Once the matrices \mathbf{A} , \mathbf{B} and \mathbf{C} are obtained, a state space model of smaller dimension n ($n < r$) can be obtained by truncating the matrices \mathbf{A} , \mathbf{B} and \mathbf{C} into \mathbf{A}_n , \mathbf{B}_n , and \mathbf{C}_n as (Maciejowski, 2002):

$$B_n = \text{1st } n \text{ rows of } B$$

$$C_n = \text{1st } n \text{ columns of } C$$

$$A_n = \text{1st } n \text{ rows and 1st } n \text{ columns of } A$$

An idea of how small n might be so that the resulting state space model still approximates the exact state space model of order r can be deduced from the following analysis. The diagonal matrix Σ can be replaced with a matrix Σ_n in which the singular values $\sigma_{n+1}, \sigma_{n+2}, \sigma_{n+3} \dots \sigma_r$ are set to zero. The new block Hankel matrix is given by:

$$M_n = U\Sigma_n V^T \quad \dots(2.69)$$

The rank of the matrix M_n is n .

The matrix M_n is the closest as possible to M “in both operator and Frobenius norms” (Patwardhan and Shah, 2002) . The Frobenius norm of the error between M and M_n is defined as:

$$\|M - M_n\|_F = \|E\|_F = \sqrt{\text{trace}(E^T E)} = \sigma_{n+1} \quad \dots(2.70)$$

Therefore a good approximation is to choose n for which σ_{n+1} is as small as possible. Using the above equation, it is obvious that an appropriate state space of low order n is obtained when $\sigma_{n+1} - \sigma_n$ approaches zero.

2.15 Relative gain array

The relative gain array (RGA) provides a measure of interaction for multivariable control systems, and a recommendation of the most effective pairing of controlled and manipulated variables for the system. The measure, which is based on a system’s steady gain matrix, was originally developed by Bristol (1966) for square systems. (Chang and

Yu, 1990) extended the RGA to non-square systems, in particular systems having more outputs than input.

The RGA for a square system is given as (Bristol, 1966):

$$RGA = K \otimes (K^{-1})^T \quad \dots(2.71)$$

where

K = steady gain array of the system

\otimes = element –by – element multiplication

The subscript T denotes transpose

The non-square array (NRGA) is given as (Chang and Yu, 1990):

$$NRGA = K \otimes (K^+)^T \quad \dots(2.72)$$

where

K^+ is the Moore-Penrose pseudo-inverse of K , defined as $K^+ = (K^T K)^{-1} K^T$

For square, both RGA and NRGA give the same answer. This means that they both have consistent definition under this condition. The sum of elements in each column of RGA or NGRA is equal to unity. While the sum of elements in each row of RGA is unity, that of NRGA falls between zero and unity. The RGA or NRGA value for an input-output pair indicates the degree of interaction between the pair. A value of unity indicates that the output depends entirely on the input, while a value of zero indicates that there is no interaction at all between the pair. A negative value indicates that the input has adverse effect on the output. In general the closer to unity a positive RGA (or NRGA) value is the greater the influence of the input on the output.

2.16 Summary

This chapter reviews literature relevant to the research. The review includes the history and principles of MPC (the transformations it has undergone from the time it was first

reported in the 1970s, to the present time, the development of its prediction algorithms, and the solution of its quadratic objective function (Hildreth's QP procedure). The contributions of many researchers to control performance assessment in general and MPC in particular are also reviewed. Both the model based approach and the data driven approach to MPC performance assessment are found to have made significant contributions. The data driven approach employing little statistics is considered in this research to be more relevant to the needs of an average MPC operator. The chapter also includes a review of literature relevant to process identification, conversion of step response models to state space models, and presentation of algorithm for relative gain array, which may assist in MPC performance assessment.

Chapter Three

MPC Models and MPC Control of Selected Nonlinear Processes

Chapter 3 describes a number of process control simulations, which act as vehicles for case studies in Chapter 4. Each simulation was chosen, because it was relatively difficult to control.

3.1 MPC Model Representations

Each MPC simulation utilises an internal model that is based on low order (maximum order of two) transfer function matrix format, to reveal the dynamics of the plant transparently. With this format it is easier to reason about dynamic behaviour for instance before the application of MPC, or to assist in diagnosis. This choice of format also emphasizes that approximate low order transfer function models are in many cases adequate for MPC applications in very complex and high order processes. During MPC design however, the transfer function matrix is converted to state space format and used in state space MPC.

The three widely used formulations of MPC differ in the type of internal algorithms that they use

- (i) for predictions within specified prediction horizon and
- (ii) for the calculation of optimum control moves for a given control horizon.

State space MPC has prediction algorithms based on the state space formulation. Transfer function MPC has prediction algorithm based on the input-output difference

equations. Step response MPC has internal prediction algorithms based on step response input-output equations using step response coefficients.

Once the MPC algorithms have been formulated however, the actual internal model can be in any acceptable format, because models can be converted. For example although it is expected that a state space MPC internal model would be presented in state space format, the available internal model might have been identified in a different format, for example as step responses or as transfer function models. The model can still be used for prediction in the state space MPC provided a means is available for converting it to state space format. The reason for preferring one form of MPC over the other should therefore be about transparency and performance, and not about the format in which the internal model is presented.

Computer programs that implement MPC in each of the three formulations (state space, step response and transfer function) were developed to compare the ease with which an MPC can be constructed and control performance evaluated. The state space MPC algorithm was easier to program, even for multivariable systems, and the MPC could handle both stable and unstable plants. One of its major components, the observer, used for process states prediction, can be designed offline for a specific plant. The state space representation of an internal model has the advantage of being able to represent the dynamics of processes (especially multivariable processes, even with delays) more compactly.

The coding for transfer function MPC would be most difficult, if it had to accommodate wide range of input-output difference equation formats (AR, ARX, ARMAX etc.). It would be challenging to have one program for all. In this thesis transfer function MPC was implemented using the ARX format difference equation. Transfer function MPC

can work on both stable and unstable systems. Unstable plant implementations that ensure offset-free tracking require modifications to the basic formulation which introduce additional complications to program development (e.g. model realignment, the incorporation of integral action into the controller (Maciejowski, 2002)

Step response MPC is moderately easy to program. Its major drawback is that it cannot be used for unstable plants, since the step response model on which it is based is predicated on the fact that the rate of change of the process must be bounded. Despite this shortcoming, step response MPC might be adopted, because its internal model structure gives a transparent description of plant dynamics (gains, time constants), and more importantly easily handle plants with internal delays. This is advantageous to the average person who is unlikely to relate too well to a state space formats.

The programming and implementation of the three different MPC formulations demonstrate that state space MPC has a clear advantage in terms of its wide range of application (stable and unstable plants), and the ease of design. For these reasons state space MPC is the choice here in the case studies. The implementations demonstrate that transparency is not a good enough reason to adopt step response MPC. The case studies demonstrated that systems defined by step responses with time delays can be converted to equivalent low order state space representations without really compromising model transparency.

A number of the case studies involve nonlinear processes: here state space models are created at specified operating points. A discussion is given for each case study on the derivation of appropriate low order transfer function matrices for the processes. Transfer function matrices are obtained in one of two ways: (i) by applying steps to the appropriate manipulated inputs and measured disturbances in turn, recording the step

responses, and then using simple techniques to identify appropriate low order transfer function models; (ii) by linearizing the process about a specified operating point, applying steps to the linearized model and then obtaining appropriate low order transfer function models for each input-output model using simple techniques. In all cases the step response plot obtained directly from the plant is compared with the step response of the linearized model and the identified low order transfer function model to ensure that an acceptable low order model has been identified. The exception to this is when an unstable plant is involved: the only model used is that obtained from linearization about the operating point, since it is impossible to obtain bounded step response.

3.2 The Nonlinear Processes Selected for the Case Studies

Three nonlinear process models were selected: a non-isothermal Continuously Stirred Thermal Reactor (CSTR), an evaporator process and a Fluid Catalytic Cracking Unit (FCCU). The particular CSTR model was chosen because of its properties: it is integrating; it is open loop unstable and thus normally requires regulatory controllers for stabilisation. With this the effect of degeneration of regulatory controllers and even comparison of direct MPC versus MPC as supervisory control could be studied. The evaporator is open loop stable, but has inputs which can be used as measured or unmeasured disturbances. It also has integrating loops. This process is used to study MPC configurations and measured disturbances that can shift the operating point to a zone that exhibits very different dynamics. The FCCU was chosen because of its complexity: many inputs and outputs, and incorporating many regulatory controllers. Its control also requires that many of the outputs have output constraints. Its outputs are also very highly coupled, and this facilitated a study into zone control as compared to set-point control. The complexity of the FCCU model lends itself to a study into effects caused by hardware problems, and by disturbance caused constraints violation.

The following sections introduce the processes and their model predictive control under normal operating conditions.

3.3 Simulation and control of the CSTR

The non-isothermal CSTR model published in Luyben (1989) is widely used by many authors in the study of process control. In this CSTR, the feed is introduced to the reaction tank at a flowrate F_0 , with temperature T_0 and concentration C_{A0} . An impeller continuously stirs the reactants in the tank to ensure perfect mixing (figure 3.1). Through chemical reaction, the reactants are transformed to another substance at temperature T with final concentration C_A , irreversibly and exothermally. The volume V of reactant in the tank is affected by the flowrate F of the product from the tank. The heat of reaction is removed with the aid of a cooling jacket that surrounds the reactor. Cooling water is added to the jacket at a volumetric flowrate F_j and with inlet temperature of T_{j0} . The volume of water in the jacket is held constant at V_j . This multivariable plant model consists of four non-linear ordinary differential equations (equations 3.1 to 3.4)

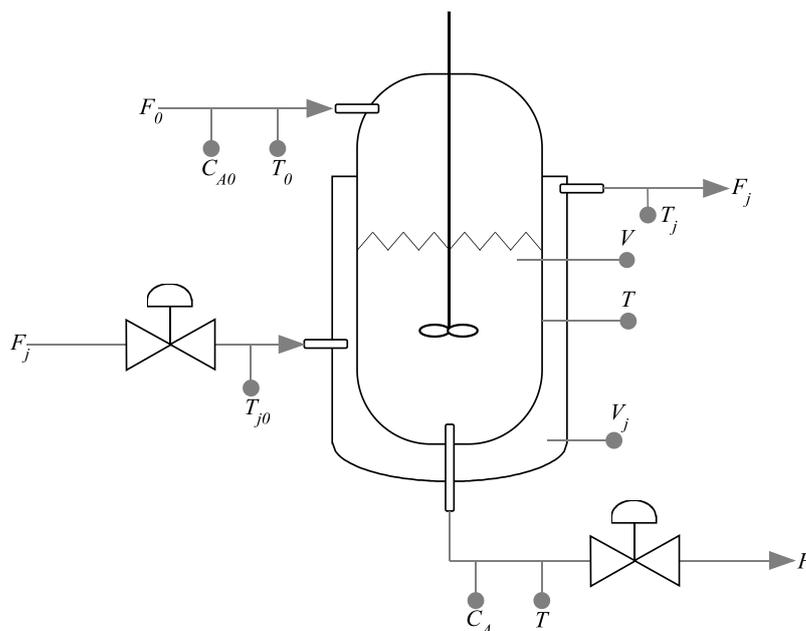


Figure 3.1: Flow diagram of the CSTR unit (open loop)

Nominal steady-state values and parameter values are shown in table 3.1 to table 3.3.

The equations and the values are taken from Luyben (1989)

$$\frac{dV}{dt} = F_0 - F \quad \dots(3.1)$$

$$\frac{d(VC_A)}{dt} = F_0C_{A0} - FC_A - VC_A\alpha e^{-E/RT} \quad \dots(3.2)$$

$$\frac{d(VT)}{dt} = F_0T_0 - FT - \frac{\lambda VC_A\alpha e^{-E/RT}}{\rho C_p} - \frac{UA_h}{\rho C_p}(T - T_j) \quad \dots(3.3)$$

$$\frac{dT_j}{dt} = \frac{F_j(T_{j0} - T_j)}{V_j} + \frac{UA_h}{\rho_j V_j C_j}(T - T_j) \quad \dots(3.4)$$

Table 3.1: CSTR Input variables

Variable	Description and units	Nominal values	Units
F	Reaction Product flowrate	40.0	ft ³ /h
F_j	Cooling water flow rate	49.9	ft ³ /hr
F_0	Feed flow rate	40.0	ft ³ /hr
C_{A0}	Feed concentration	0.5	ib.mol A/ft ³
T_0	Feed Temperature	530.0	°R

Table 3.2: CSTR Output Variables

Variable	Description and units	Nominal values	Units
V	Reactor holdup volume	48.0	ft ³
C_A	Reaction product concentration	0.245	ib.mol A/ft ³
T	Reactor absolute Temperature	600.0	°R
T_j	Cooling water temperature	594.6	°R

Table 3.3: CSTR Parameters

Variable	Description	Nominal values	Units
T_{j0}	Cooling water temperature in	530.00	°R
V_j	Cooling water volume	3.85	ft ³
E	Activation Energy	30,000.00	Btu/ib.mol
U	Overall heat transfer coefficient	150.00	Btu/h ft ² °R
C_p	Heat capacity of process liquid	0.75	Btu/ibm °R
ρ	Density of process liquid	50.00	ibm/ft ³
A	Frequency factor	7.08×10^{10}	h ⁻¹
R	Universal gas constant	1.99	Btu/ib mol
A_h	Heat transfer area	250.00	ft ²
λ	Heat of reaction -	30,000.00	Btu/ib mol
C_j	Heat capacity of cooling liquid	1.00	Btu/ibm °R
ρ_j	Density of cooling water	62.30	ib/ft ³

Of the five inputs variables (F , F_j , F_0 , C_{A0} and T_0), two of them (C_{A0} and T_0) can be measured but cannot be manipulated, so they can only be used as disturbance variables. The feed flowrate F_0 usually depends on an upstream process and would normally not be manipulated. Therefore only two input variables (F and F_j) may be used as manipulated variables. There are four possible controlled variables in V , C_A , T and T_j . The model of the nonlinear CSTR is simulated with MALAB and Simulink around the steady state operating point of table 3.1. The open-loop trends of the outputs are shown in figure 3.2 below for the situation where the simulation is initialized at values given in tables 3.1, 3.2 and 3.3.

The plots of figure 3.2 suggest that the open loop CSTR is unstable without control, and is also nonlinear, like the case of an *inverted pendulum*. Different perturbations of the inputs give the same outputs always. The exothermic reaction stops causing the reactor outlet temperature to tend towards the feed temperature of 530⁰R. The operating point shown in tables 3.1 to 3 can only be achieved with the use of controllers.

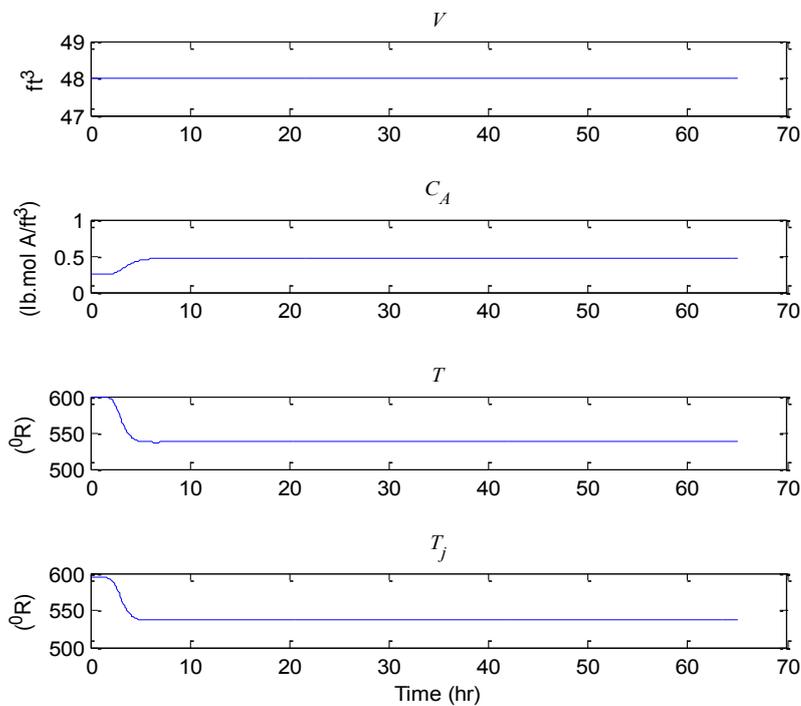


Figure 3.2: Open loop simulation of the CSTR

Open loop instability means that it is not appropriate to perform a step test on it for purposes of obtaining the open loop dynamic model. Fortunately the nonlinear model of the CSTR is available, so that a linear approximation can be obtained about the operating point specified in tables 3.1 to 3.3 by using the first-order Taylor series expansion:

$$\text{if } \frac{dx}{dt} = f(x, u) \quad \dots(3.5)$$

then,

$$\frac{dx'}{dt} = \left. \frac{\partial f}{\partial x} \right|_{\bar{x}, \bar{u}} (x - \bar{x}) + \left. \frac{\partial f}{\partial u} \right|_{\bar{x}, \bar{u}} (u - \bar{u}) \quad \dots(3.6)$$

The variable x represents the states of the differential equation while the variable u represents the input variables. The terms $(x - \bar{x})$ and $(u - \bar{u})$ serve as deviation variables.

The state space linear model obtained by linearizing the nonlinear model about the values of table 3.1 to 3.3 is shown in equation 3.7. Step response plots obtained from this linear model are shown in figure 3.3 below. The plots show that the loop between F and V is an integrating one (liquid level control). It shows that V is unaffected by changes in F_j ; that the response of C_A , T and T_j due to unit step changes in F and F_j are unbounded. This confirms that the open loop nonlinear plant is unstable.

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ -0.004427 & -1.701 & -0.008897 & 0 \\ 1.215 & 693.7 & -14.55 & 20.83 \\ 17.59 & 0 & 156.3 & -169.3 \end{bmatrix} \\ B &= \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -16.78 \end{bmatrix} \\ C &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad \dots(3.7)$$

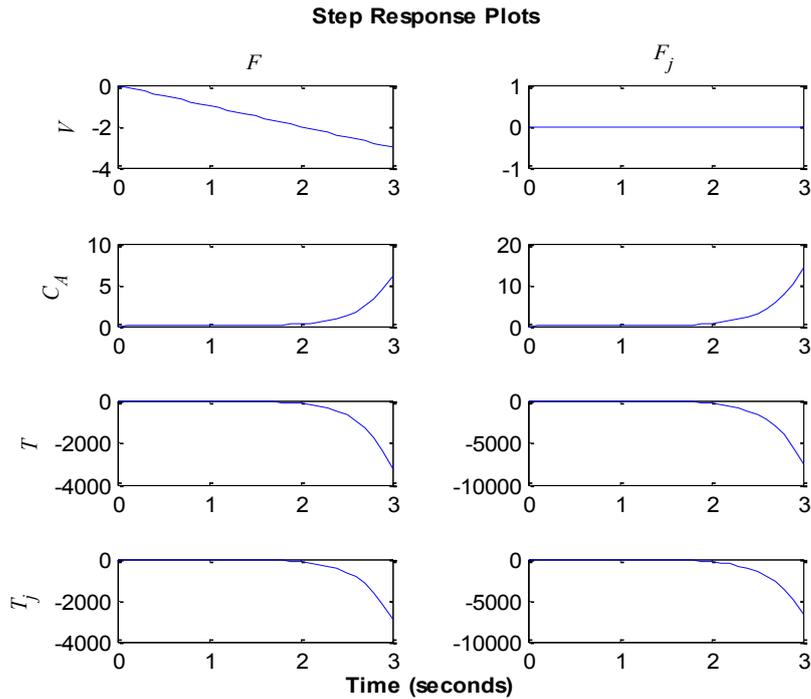


Figure 3.3: Step response plots of the open loop model of the CSTR

Since the open loop step response of the CSTR is unbounded, it is impossible to obtain a finite number of step response coefficients to represent the true dynamics of the system. Therefore, MPC based on the step response formulation of equation 2.1, cannot be used to control the plant directly. To use MPC based on the step response formulation, the plant must be stabilised first using regulatory controllers, and the step response MPC can then be implemented as a supervisory controller to the lower level regulatory controllers. State space equation 3.7 represents the dynamics of the CSTR completely about the operating point, so MPC based on this formulation can be used to control the plant directly. A main requirement is that the model must be observable, because state space MPC incorporates an observer to estimate the state of the plant.

3.3.1 Direct MPC on the CSTR

State space MPC is applied directly using two MVs (F and F_j) and two CVs (V and T), that is as a square system. The MPC parameters used are given in table 3.4. Figures 3.4 and 3.5 show the trends obtained when the controller is asked to keep the two CVs at

their nominal operating points for the first 40 hrs when steps are applied to the set-points of the two controlled variables simultaneously. The MPC is well able to ensure set-point tracking of the controlled variables before and after steps are applied to the set-points. The control variable V exhibits a nonminimum-phase phenomenon when the steps are applied: F_j and F increase (momentarily) to compensate for the demanded change in temperature T ; V responds because it is solely dependent on F . After a while equilibrium is reached where F_j attains a value sufficient to maintain T at the desired set-point; F reverts to its initial value, and hence V levels off. The two MVs (F and F_j) did not saturate nor were they too aggressive in their control actions.

Table 3.4: MPC parameters for control of open loop CSTR

Parameter	Symbol	Value
Sampling Interval	Ts	0.1
Prediction Horizon	P	30
Control Horizon	M	3
Input weights	iw	[0.1, 0.1]
Output weights	ow	[1, 1]
Constraint on F		$0 \leq F \leq 80$
Constraint on F_j		$0 \leq F_j \leq 100$

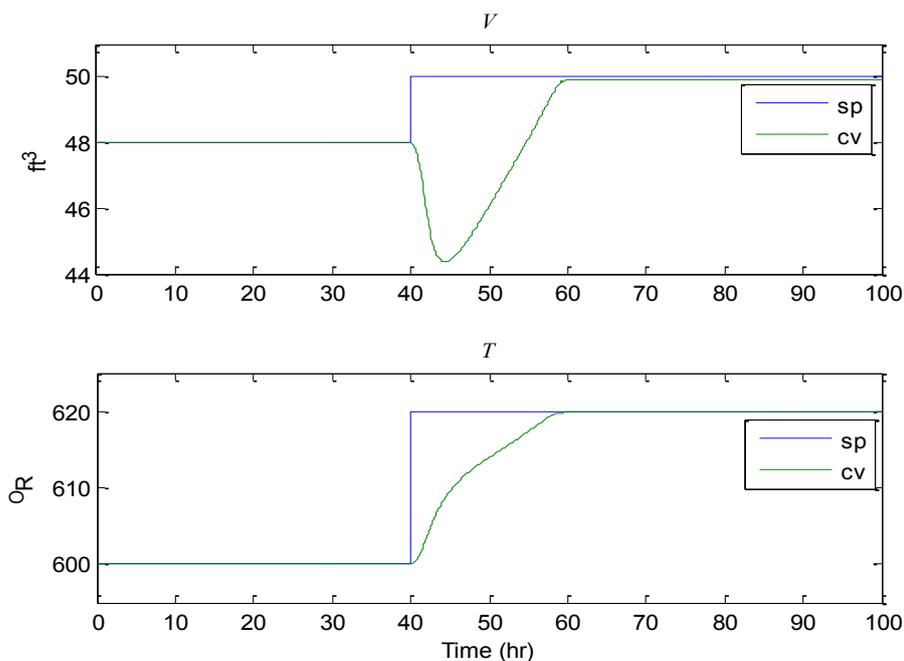


Figure 3.4: Trends of the controlled variables for direct MPC on unstable CSTR (sp = set point. cv = control variable)

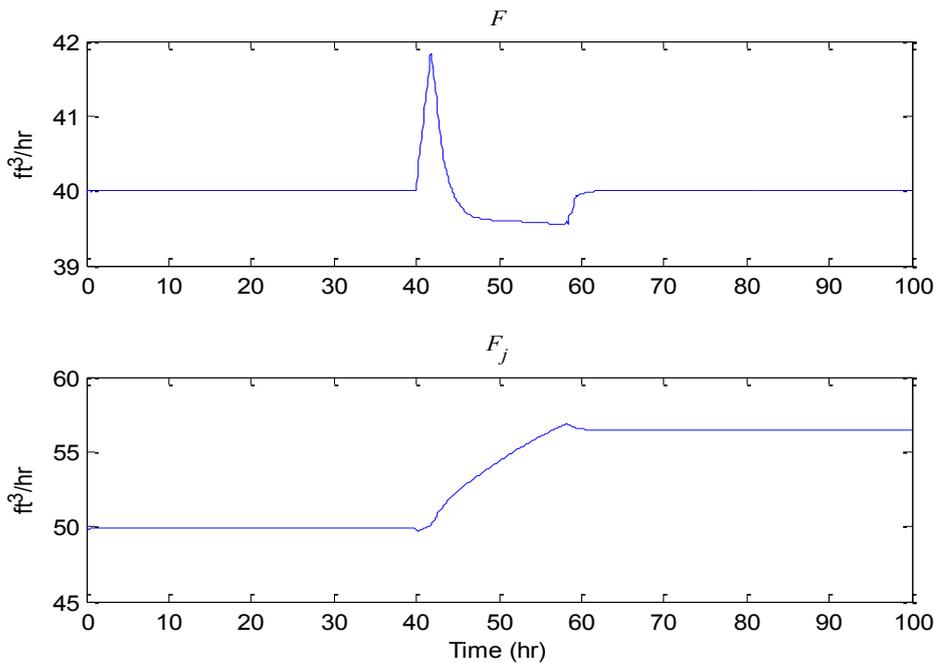


Figure 3.5: Trends of the manipulated variables for direct MPC on unstable CSTR

3.3.2 MPC as Supervisory Controller for the CSTR

This case study reflects many industrial applications where MPC is used in a supervisory capacity with to lower level regulatory controllers. Luyben (1989) already showed that the plant can be stabilised using two proportional controllers. One for loop $F-V$ and the other for loop F_j-T . From the open loop step response plots of figure 3.3, it is seen that F_j has a similar effect on T_j as on T . Therefore closing loop F_j-T_j has an identical effect as closing loop F_j-T . The loop F_j-T_j is chosen for regulatory control and T is used as a controlled variable in the supervisory controller (see figure 3.6). Such a scheme would be sensible from a commissioning perspective because the two regulatory control loops could be tuned before any reaction took place. For the purpose of offset free tracking, integral actions are added to each of the proportional controllers. Table 3.5 below gives PI controller settings for the two loops.

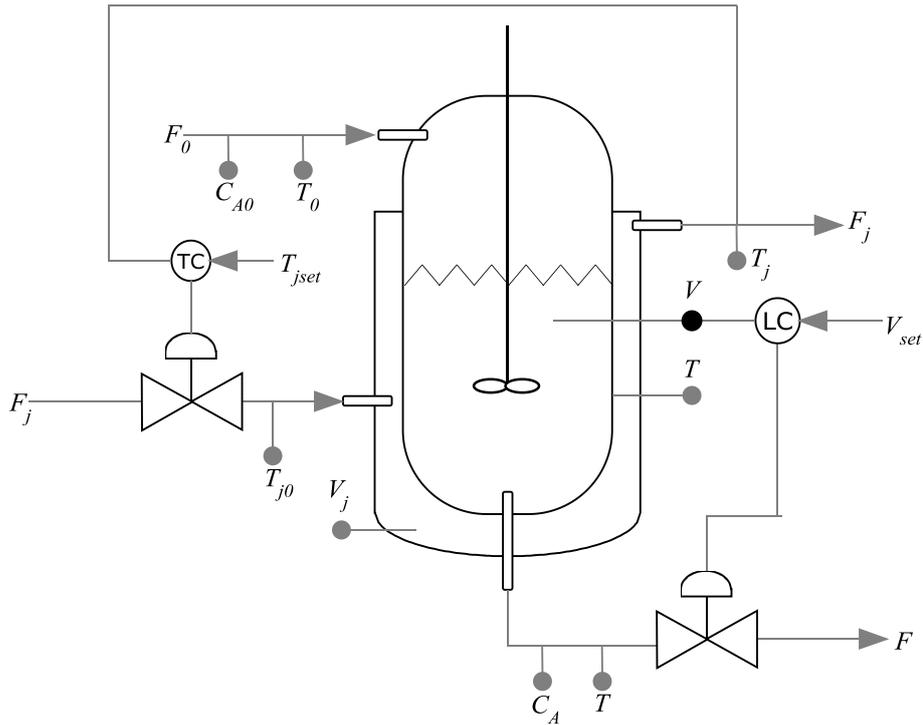


Figure 3.6: CSTR with two PI controllers

Table 3.5: PI settings for the stabilizing controller

Loop	Proportional Gain P	Integral Time Constant Ti
$F-V$	-10	2
F_j-T_j	-5	2

Figure 3.7 shows step response plots obtained for this scheme. The associated steady state gain values are given in table 3.6. The two closed loops have unit steady gain values each, as expected. The zero steady state gain values in the table indicate that any change in V_{set} and T_{jset} has no effect on the corresponding outputs at steady state.

Table 3.6: Actual steady state gain values of closed loop CSTR

	V_{set}	T_{jset}
V	1.0000	0.000
C_A	-0.0024	-0.006
T	-0.0325	1.146
T_j	0.0000	1.000

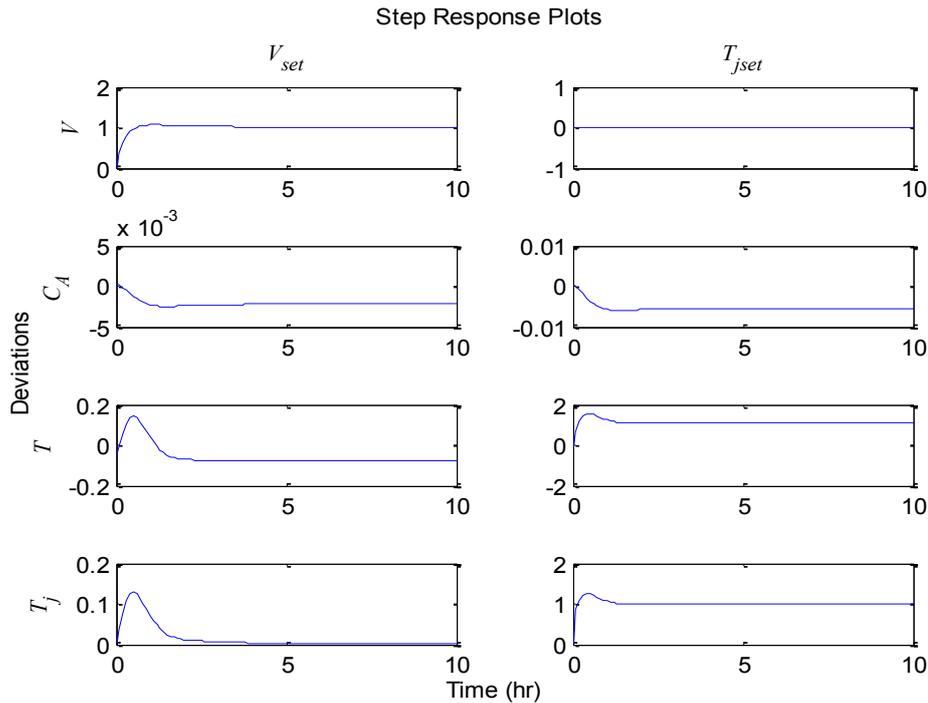


Figure 3.7: Step response plot of the closed loop CSTR

The linear state space model of the system incorporating regulatory control is shown in equation 3.8 below. The linear state space model of the closed loop CSTR can be obtained by linearizing at the operating point specified in tables 3.1 and 3.2. The linear state space model is of order six, which includes the states of the PI controllers.

$$A = \begin{bmatrix} -10 & 0 & 0 & 0 & -1 & 0 \\ -0.004427 & -1.701 & -0.008897 & 0 & 0 & 0 \\ 1.215 & 693.7 & -14.55 & 20.83 & 0 & 0 \\ 17.59 & 0 & 156.3 & -240.3 & 0 & -16.78 \\ 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.5 & 0 & 0 \end{bmatrix};$$

$$B = \begin{bmatrix} 10 & 0 \\ 0 & 0 \\ -0 & 0 \\ 0 & 83.9 \\ -5 & 0 \\ 0 & -2.5 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix};$$

...(3.8)

Appropriate low order transfer function models for each input-output pair can be identified from the step response data obtained from the closed loop plant. Using low

order model better represents real industrial practice. The transfer function models are given in table 3.7 below.

Table 3.7: Transfer function matrix for CSTR

	V_{set}	T_{jset}
V	$\frac{1 + 2s}{1 + 2s + 0.4s^2}$	0
C_A	$\frac{-0.0021(1 + 2.1s)}{1 + 2s + 0.69s^2}$	$\frac{-0.006(1 + 0.8s)}{1 + 1.03s + 0.215s^2}$
T	$\frac{-0.08(1 - 2s)}{1 + 0.64s + 0.16s^2}$	$\frac{1.1(1 + 0.61s)}{(1 + 0.432s + 0.065s^2)}$
T_j	$\frac{-0.001(1 - 7.5s)}{1 + 0.81s + 0.2025s^2}$	$\frac{1 + 0.48s}{(1 + 0.414s + 0.053s^2)}$

Actual step response plots (s_a), linear state space model step response plots (s_l), and the reduced order model step response plots (s_m) are all shown in figure 3.8 below. The trends of the step response plots are very similar so any of these models could be used as the internal model of the MPC. Note that changes in T_{jset} has no effect on V . The absence of the s_l plot for loop $V_{set}-T_j$ indicate that the linear state space model could not capture the very insignificant dynamics, thus reinforcing our understanding from the very small steady state gain as shown in table 3.7.

For comparison, two different MPCs (MPC_{TF} and MPC_{SS}) are implemented as supervisory controllers to the plant. The low order transfer function matrix shown in table 3.7 above is used as the internal model to MPC_{TF}. MPC_{TF} manipulates the PI regulatory controllers set point variables V_{set} and T_{jset} . The PI regulatory controllers manipulate the variables F and F_j to the plant; the controlled variables (CVs) are V and T ; the MPC parameters values are given in table 3.8 below. The constraints on V_{set} and T_{jset} are hard and for safety reasons are not allowed to drift far from their nominal values.

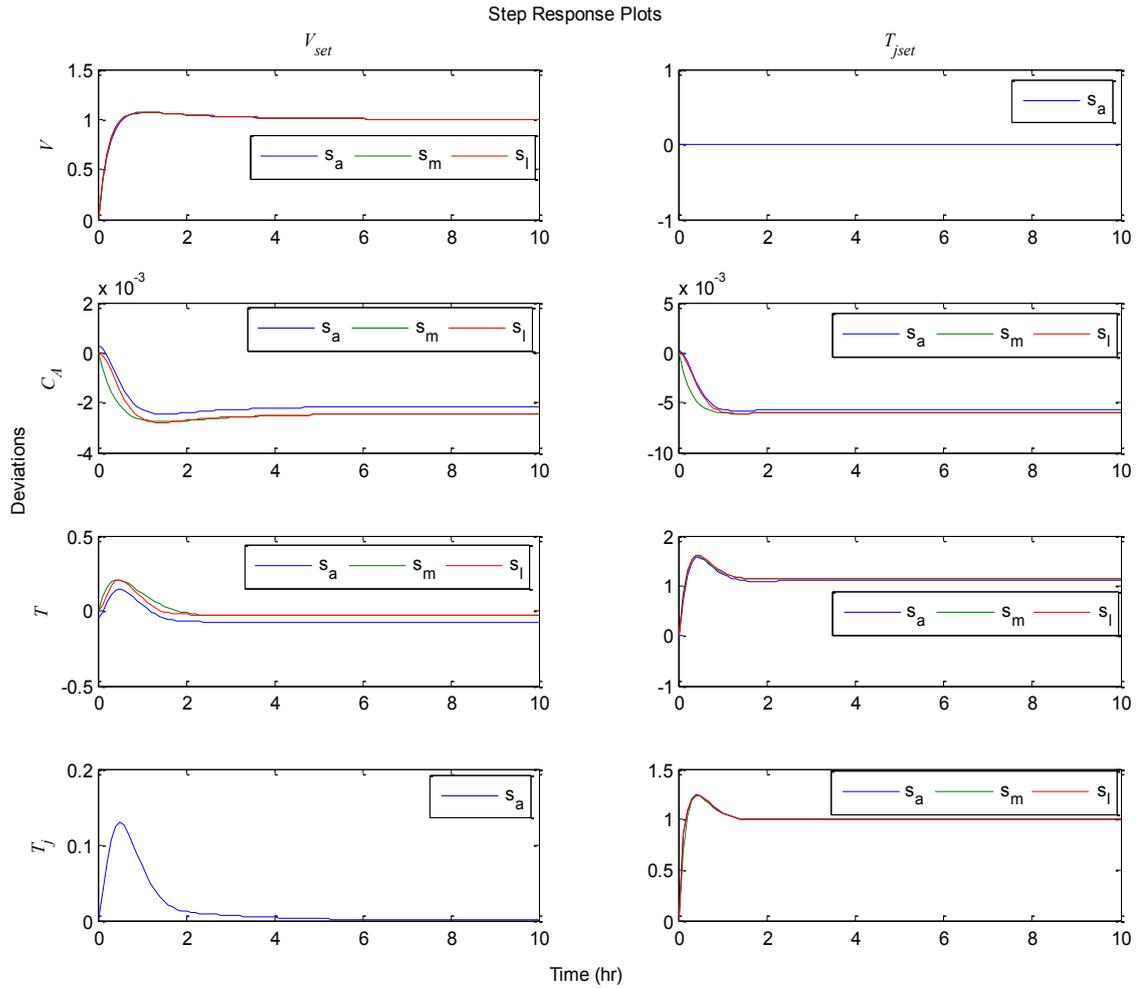


Figure 3.8: Step response plots of the close loop CSTR models (s_a = actual. s_l = linear. s_m = reduced order)

Table 3.8: MPC simulation parameters for CSTR control

Name	Symbol	Value
Sampling Interval	T_s	0.1 hr
Prediction Horizon	P	30
Control Horizon	M	3
Input weights	iw	[0.1, 0.1]
Output weights	ow	[1, 1]
Constraint on V_{set}		$38 \leq V_{set} \leq 48$
Constraint on T_{jset}		$550 \leq T_{jset} \leq 650$

The same MPC settings of table 3.8 are used for the second MPC, MPC_{SS}. The internal model of this MPC is the linear state space of equation 3.8.

The two MPCs are each simulated for an equivalent of 100 hrs: the plant is operated at the nominal values of tables 3.1 to 3.3 for the first 40 hrs; at the 40 hr mark, simultaneous steps are applied to the set points of the two controlled variables; the magnitudes of the two steps are 2 ft^3 and $20 \text{ }^0\text{R}$ respectively. The first part is to show that the two MPCs can keep the plant running at the nominal operating point. The second part is to demonstrate the set point tracking capabilities of the two MPCs when the plant is operated at points slightly away from the nominal operating. The results are shown in the plots of figures 3.9 to 3.11. For the two MPCs, the two controlled variables V and T (figure 3.9) track their set-points during normal operation and when the set-points are stepped from their normal operating point. The figure also shows that both the linear state space model and the derived low order transfer function matrix give largely very adequate performance. Both have almost the same settling time, though their transient characteristics for the control of reactor volume V are different. The trends of the manipulated variables from the MPCs (V_{set} and T_{jset} in figure 3.9) and the manipulated variables from the PI controllers (F and F_j in figure 3.10) show that the control action for the MPC using low order transfer function matrix is less aggressive compared to the one using linear state space model. A noticeable effect of the MPC as supervisory control is that the non-minimum-phase phenomenon observed when a step is increases in the control variable T , is no longer present, having been absolved by the PI controller (see figure 3.11).

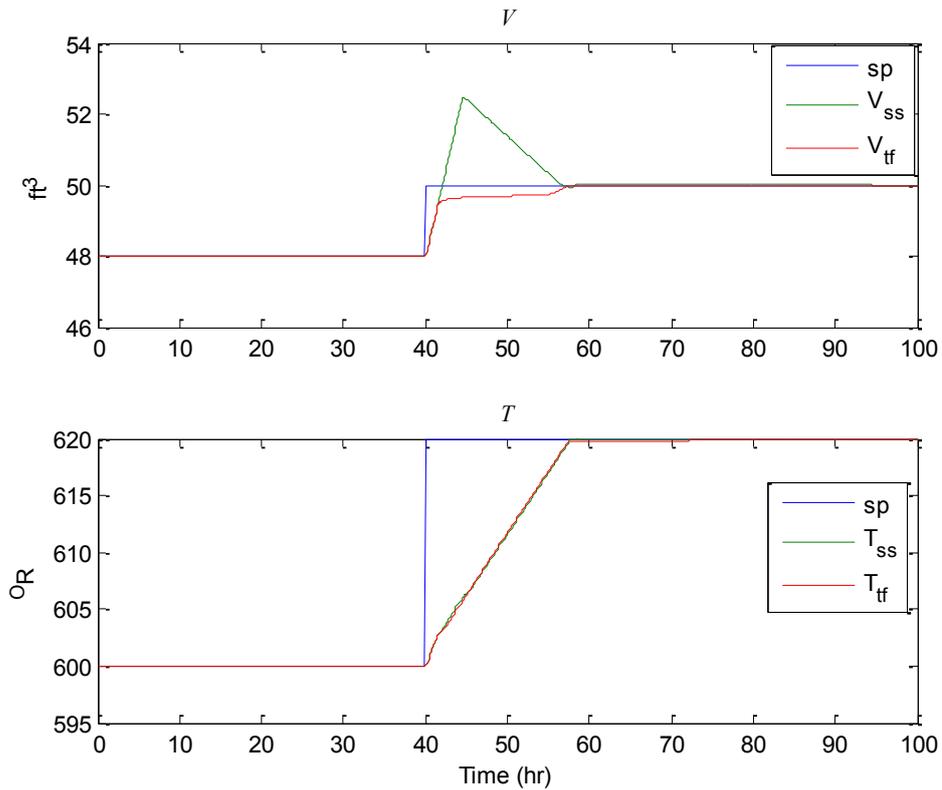


Figure 3.9: Controlled variable trends MPC as supervisory control on the CSTR ($sp =$ set point. Subscripts $ss =$ state space. Subscripts $tf =$ transfer function)

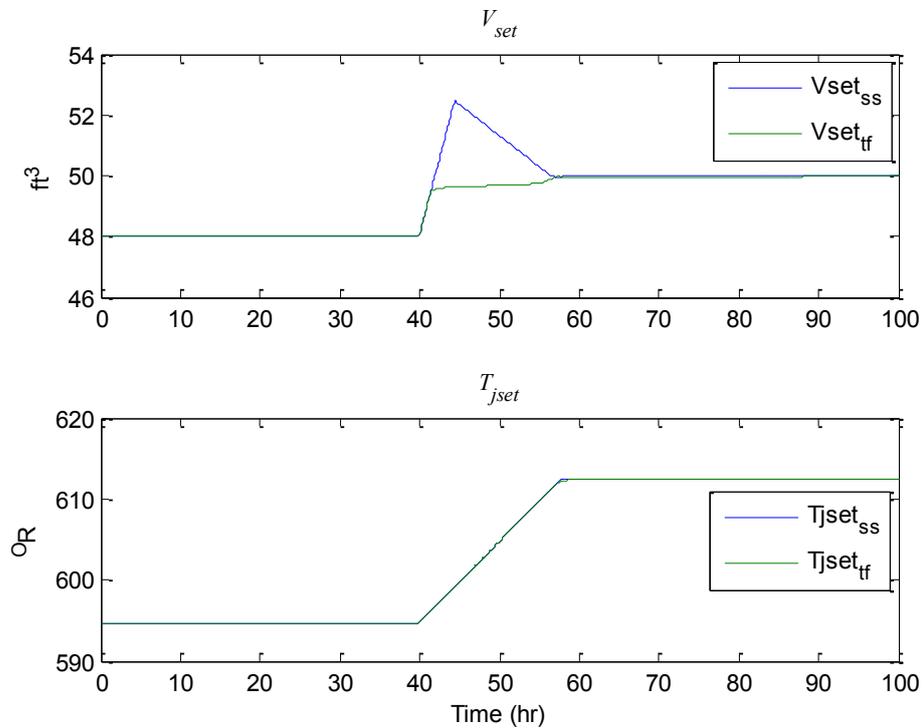


Figure 3.10: Manipulated variable (MPC output) trends MPC as supervisory control on the CSTR (Subscripts $ss =$ state space. Subscripts $tf =$ transfer function)

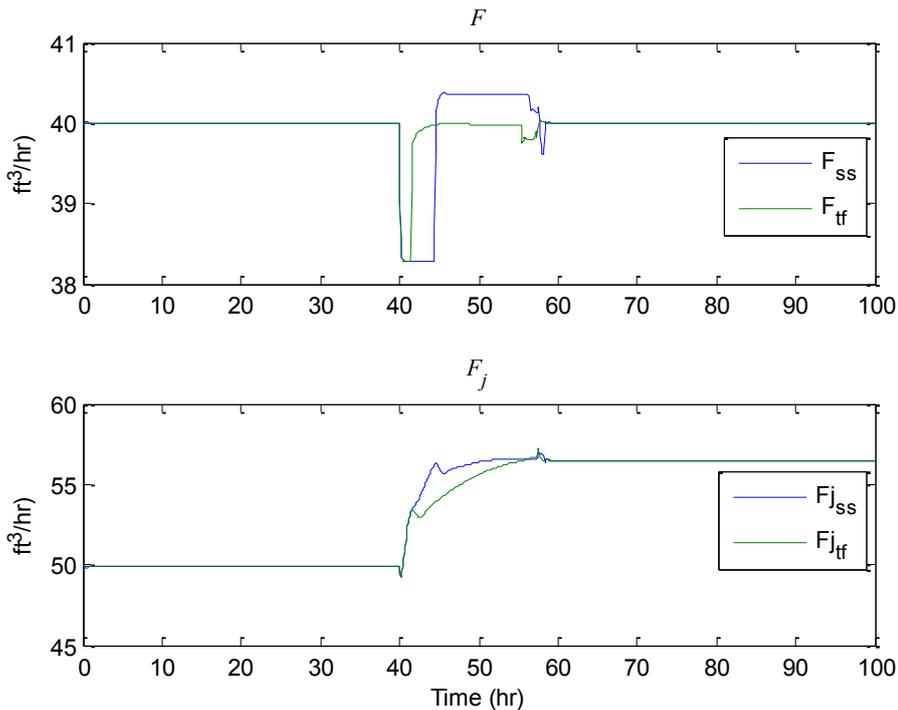


Figure 3.11: Manipulated variable (PI output) trends MPC as supervisory control on the CSTR Subscripts ss = state space. Subscripts tf = transfer function)

If the reactor temperature T is to be stepped up (as in figure 3.9), the cooling water flowrate (F_j) must also increase to compensate for increased cooling load demand (figure 3.10). Thus the new steady state cooling water flowrate is higher than that before the step. The reaction concentration C_A is tightly coupled to reactor temperature T (figures 3.7 and 3.8), and by extension to the coolant temperature T_j . Therefore any change in T has an immediate effect on C_A and T_j . It indicates that for the purpose of set point tracking, only one of the three (T , C_A , and T_j) may be chosen as a controlled variable at any one time. When the tank volume is stepped up (figure 3.9), the controller takes immediate action by reducing the outflow from the tank (figure 3.10). As soon as the tank reaches the new volume however, the flowrate from the tank returns to the initial resting value. This is typical of integrating systems.

3.3.3 Comparison of direct and supervisory MPC control of the CSTR

The low order transfer function model for the CSTR in open loop could not be obtained, because it is open loop unstable. This comparison could only be performed with linear state space models. Comparison plots of MPC as direct and MPC as supervisory control are shown in figures 3.12 and 3.13. Direct MPC exhibits a non-minimum-phase phenomena while supervisory MPC does not (figure 3.12). The settling time for MPC as supervisory control is shorter than for direct MPC. The trends of the controlled outputs for the MPC as supervisory control also show better transient dynamics in terms of overshoot when compared with direct MPC action (figure 3.12). The tight control action is also accompanied by very tight manipulation of inputs (F and F_j) to the CSTR process (figure 3.13). Here flow rates F and F_j are outputs from the PI controllers.

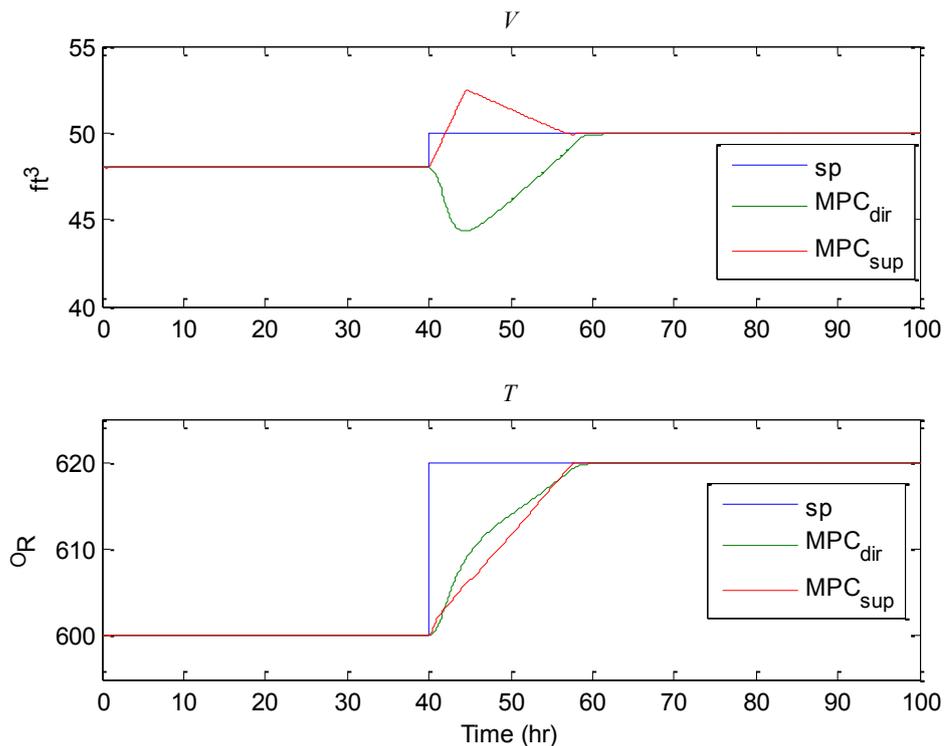


Figure 3.12: Controlled variable comparison: direct vs supervisory (Subscript dir = direct. Subscript sup = supervisory)

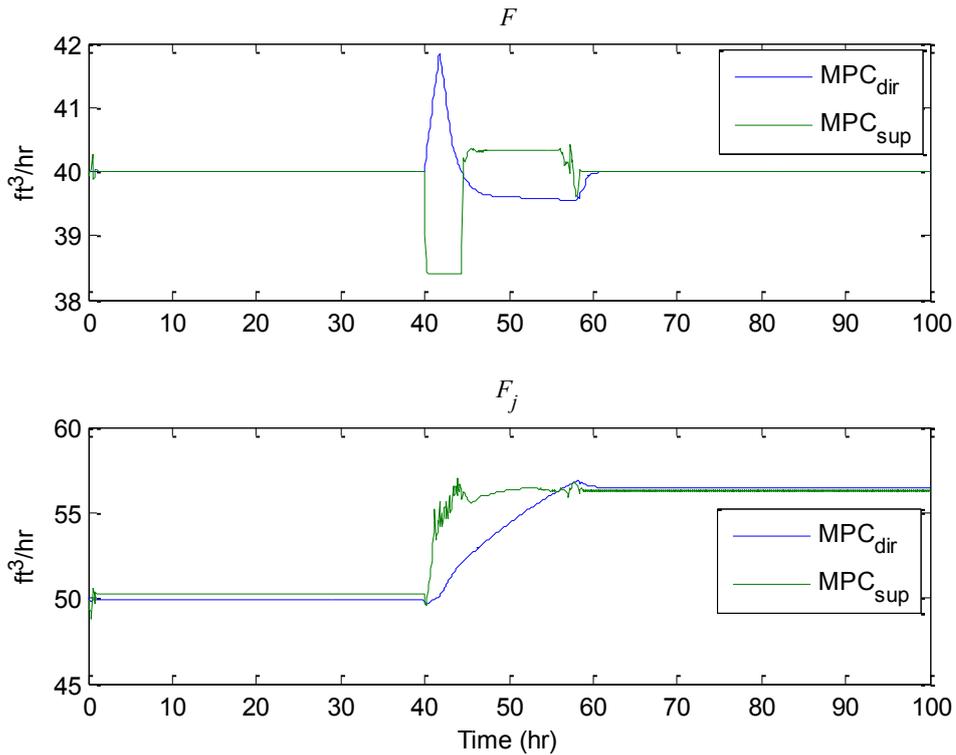


Figure 3.13: Manipulated variable comparison: direct vs supervisory (Subscript dir = direct. Subscript sup = supervisory)

3.4 The Evaporator Process

This non-linear model of a forced circulation evaporator system by Newell and Lee (1989) is partly based on the description given in Maciejowski (2002). An evaporator (figure 3.14) is used to remove a volatile liquid from a non-volatile solute, resulting in a concentrated solution. It consists of a heat exchanger (the evaporator), a separator and a condenser with a recirculating pump. A feed stream at a concentration X_1 , temperature T_1 , with a flow rate F_1 , is mixed with recirculating liquor pumped at a flow rate F_3 . The mixture enters the evaporator and is heated by steam flowing at a rate F_{100} with entry temperature T_{100} and pressure P_{100} . The heating of the mixture of feed and recirculating liquor inside the evaporator results in a mixture of higher concentration. The operating pressure inside the evaporator is P_2 . The resulting concentrated mixture of vapour and liquid enters a flash separator, in which the liquid level is L_2 . The flash separator is used to separate the mixture into vapour and liquid. The separated vapour

exits the separator and enters the condenser at flow rate F_4 and temperature T_3 , where it is cooled using cooling water flowing through the condenser at flowrate F_{200} and entry temperature of T_{200} and exit temperature of T_{201} . The condensed liquid exits the condenser at a flowrate of F_5 . The liquid from the separator becomes the recirculating liquor. The product stream, with concentration X_2 is drawn off at a flow rate F_2 and temperature T_2 from the recirculating liquor. In this description the controlled variables are specified as L_2 , X_2 and P_2 , while the manipulated variables are F_2 , P_{100} and F_{200} .

The governing equations for the process are given in equations 3.7 to 3.18. The input and output variables, their nominal operating values, and parameter values are given in tables 3.9 to 3.11 below.

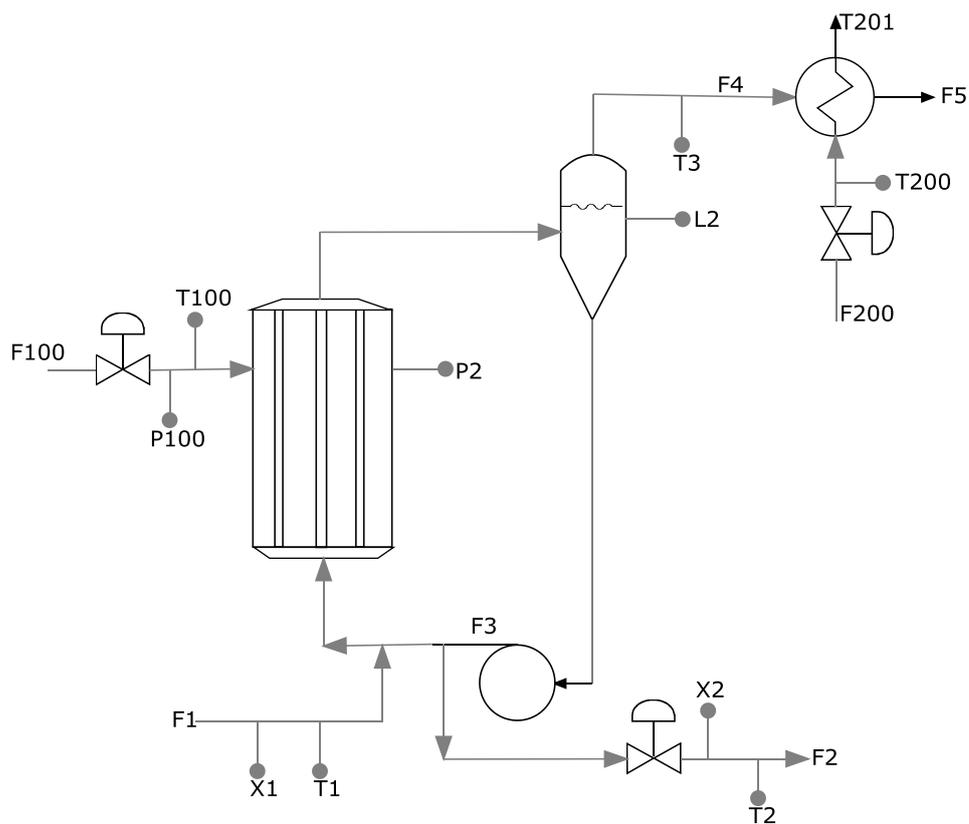


Figure 3.14: The Evaporator

$$\rho A \frac{dL_2}{dt} = F_1 - F_2 - F_3 \quad \dots(3.7)$$

$$M \frac{dX_2}{dt} = (F_1)(X_2) - (F_2)(X_2) \quad \dots(3.8)$$

$$C \frac{dP_2}{dt} = F_4 - F_5 \quad \dots(3.9)$$

$$T_2 = 0.5616(P_2) + 0.3126(X_2) + 48.43 \quad \dots(3.10)$$

$$T_3 = 0.507(P_2) + 55.0 \quad \dots(3.11)$$

$$F_4 = (Q_{100} - (F_1)C_p(T_2 - T_1))/\lambda \quad \dots(3.12)$$

$$Q_{200} = \frac{UA_2}{1 + UA_2/(2C_p F_{200})} \quad \dots(3.13)$$

$$T_{201} = \frac{Q_{200}}{F_{200}C_p} + T_{200} \quad \dots(3.14)$$

$$F_5 = Q_{200}/\lambda \quad \dots(3.15)$$

$$T_{100} = 0.1538(P_{100}) + 90 \quad \dots(3.16)$$

$$Q_{100} = 0.16(F_1 + F_3)(T_{100} - T_{200}) \quad \dots(3.17)$$

$$F_{100} = Q_{100}/\lambda_s \quad \dots(3.18)$$

Table 3.9: Input variables of the evaporator and the equilibrium values

Variable		Symbol	Nominal value	Unit
Product flow rate	Manipulated	F2	2.0	kg/min
Cooling water flow rate		F200	208	kg/min
Steam pressure		P100	194.7	kPa
Circulating flow rate	Measured disturbance	F3	50.0	kg/min
Feed flow rate	Unmeasured disturbance	F1	10.0	kg/min
Feed concentration		X1	5.0	%
Feed temperature		T1	40.0	°C
Cooling water inlet temperature		T200	25.0	°C

Table 3.10: Output variables of the evaporator and the equilibrium values

Variable		Symbol	Nominal value	Unit
Separator level	Controlled	L2	1	m
Product composition		X2	25	%
Operating pressure		P2	50.5	kPa
Separated vapour flowrate	Uncontrolled	F4	8.0	kg/min
Condensate flowrate		F5	8.0	kg/min
Steam flowrate		F100	9.3	kg/min
Product temperature		T2	84.6	°C
Separated vapour exit temperature		T3	80.6	°C
Steam temperature		T100	119.9	°C
Cooling water exit temperature		T201	25.0	°C
Heat duty		Q100	339.2	kW
Condenser duty		Q200	308.0	kW

Table 3.11 Evaporator parameters

Parameters	Symbol	value	unit
Liquid hold up in the tank	M	20	kg/m
Coefficient	UA_2	6.84	kW/K
Coefficient	C	4.0	Kg/kPa
Heat capacity of water	C_p	0.07	kW/kg.min
Latent heat of evaporated water	λ	38.5	kW/kg.min
Latent heat of step at saturation	λ_s	36.6	kW/kg.min

3.4.1: Open Loop Simulation of the Evaporator

Maciejowski (2002) suggests a scheme (figure 3.15) in which three manipulated variables F2, F200 and P100, together with the measured disturbance F3, are controlled by local servos actuated valves with time constant T. Figure 3.16 shows the responses to unit step increases in these three manipulated variables. The open loop steady state

gains between the manipulated variable and the controlled variables are shown in table 3.12.

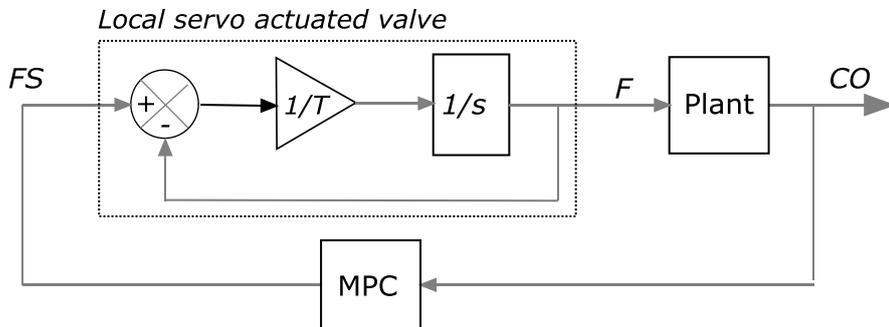


Figure 3.15 Plant input controlled by a local servo actuated valve

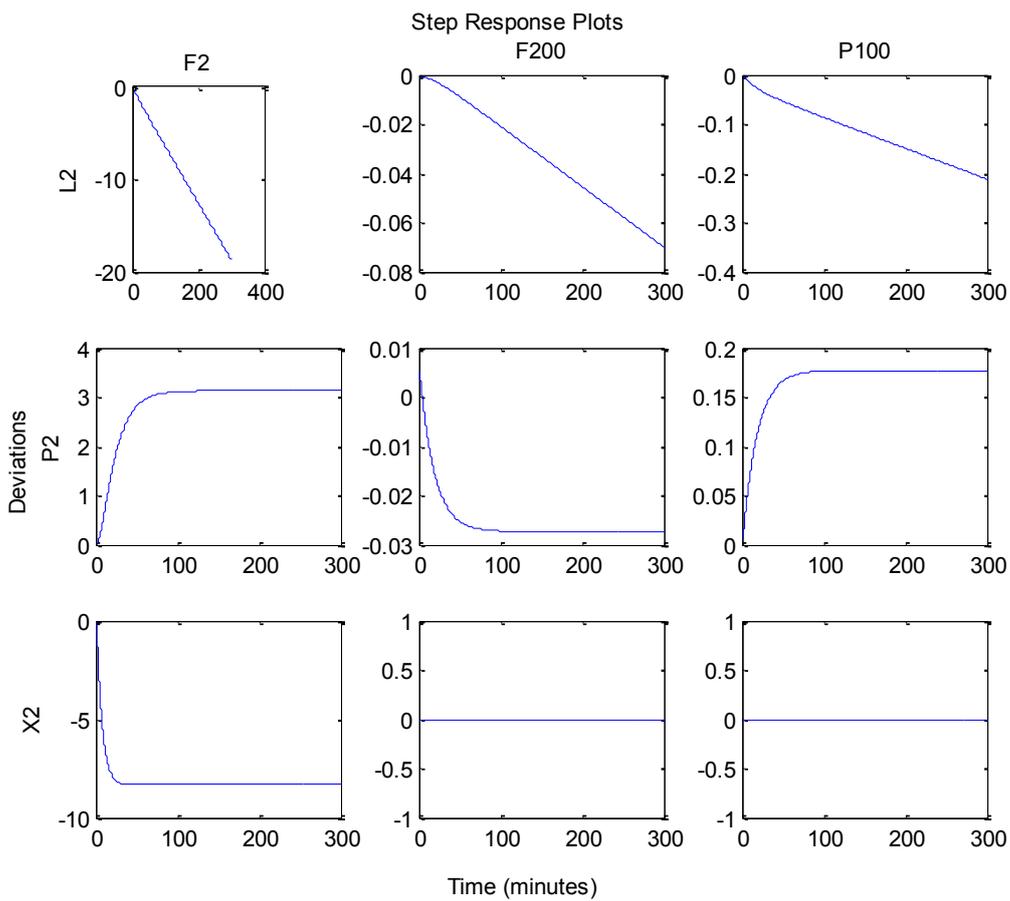


Figure 3.16: Step response of nonlinear plant to unit step changes in MVs in open loop

Table 3.12: Steady state MV-CV gains of the evaporator in open loop

	F2	F200	P100
L2	-ve ramp	-ve ramp	-ve ramp
P2	3.3	-0.027	0.177
X2	-8.33	0.00	0.00

From figure 3.16, it can be seen that loop F2-L2 is an integrating loop. This means that the separator level L2 ramps if any change is made to the product flow rate F2. A regulator with proportional action only would suffice in removing the integrator, but the regulator must be used with an integral action to ensure a zero steady state offset. The steady gains of loops F200-X2 and P100-X2 are zeros. They also do not exhibit any transient dynamics at all. These mean that changes in the manipulated variables F200 and P100 do not affect the controlled variables P2 and X2 respectively in any significant way. The loops F200-L2 and P100-L2 are also integrating loops, though the values of their slopes are very small. In a control sense, F200 and P100 have very little effect on L2. Input variable F2 has a significant effect on all the outputs.

Using the step response data obtained directly from the plant, the plots of which are shown in figure 3.15 above, the identified low order transfer function matrix for the open loop evaporator is obtained as shown in table 3.13 below. Also, the linear state space model of the open loop plant (obtained by linearizing about the operating point specified in table 3.9 and 3.10), is shown in equation 3.19 below. The plots of figure 3.17 compare the actual step response data from the plant with the step response plots of the two linear models (the state space model and the transfer function model).

Table 3.13: Approximate transfer function matrix for open loop evaporator

	F2	F200	P100	F3
L2	$\frac{-0.062}{s}$	$\frac{-0.00027}{s}$	$\frac{-0.00063}{s}$	$\frac{-0.0024}{s}$
P2	$\frac{3.13}{1 + 22.8s}$	$\frac{-0.027}{1 + 18.3s}$	$\frac{0.177}{1 + 18.2s}$	$\frac{0.66}{1 + 18.2s}$
X2	$\frac{-8.3}{1 + 7.2s}$	0	0	0

$$A = \begin{bmatrix} -0.08333 & 0 & 0 & 0 & 0 & 0 & -1.042 \\ -0.0182 & -0.05027 & 0 & -0.002037 & 0.00688 & 0.02628 & 0 \\ 0.00364 & 0.00654 & 0 & 0 & -0.001376 & -0.005255 & -0.04167 \\ 0 & 0 & 0 & -0.8333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.8333 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.8333 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.8333 \end{bmatrix};$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}; \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

...(3.19)

The open loop steady state gain values are shown in table 3.6.

Table 3.14: Steady state MV-CV gains of the linear model of the evaporator open loop

	F2set	F200	P100
L2	-ve ramp	-ve ramp	-ve ramp
P2	4.69	-0.033	0.172
X2	-12.5	0.000	0.000

For the transfer function matrix, each input-output model is of order one. Compared with the spate space model, the transient dynamics are similar (rise time, settling time, time constant), except for some difference in steady state gains as shown in tables 3.12 and 3.14.

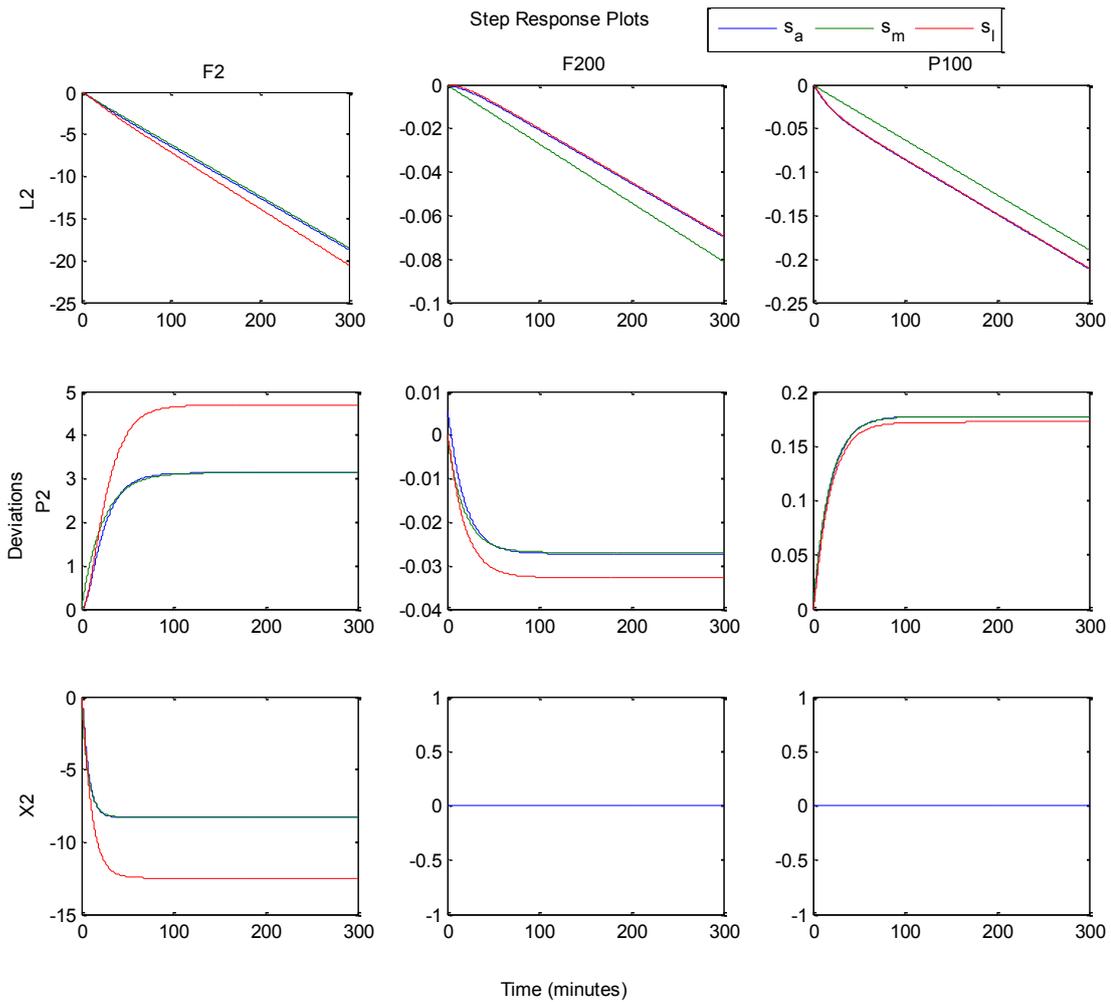


Figure 3.17: Step response plots of the models of the closed loop evaporator (actual (blue), reduced order (green) and full linear (red))

As explained by Maciejowski (2002), this particular plant requires adaptive MPC (periodic re-linearization of the internal model), if it has to be used at an operating point far removed from the nominal values given in tables 3. 9 and 3.10. Most real industrial applications do not operate this way however, because they usually have a fixed model which they expect to work for some acceptable range of operating conditions. Again MPC is implemented both directly on the open loop plant and as supervisory control when at least one of the loops is closed.

3.4.2 Direct MPC

Two MPCs, one using the linear state space model (MPC_{SS}), and the other using the reduced transfer function model (MPC_{TF}), were implemented in turn on the evaporator. The MPCs manipulate the set-points to the local servo actuated valves. MPC settings for the two implementations are given in table 3.15 below:

Table 3.15: MPC parameters for direct MPC control of evaporator

Parameters	Values
Prediction horizon	30
Control horizon	3
Sampling interval	1
MV weights [F2 F200 P100]	[0.1 0.1 0.1]
CV weights [L2 P2 X2]	[1000 100 100]
MV Constraints [F2 F200 P100]	Max = [4 400 400] Min = [0 0 0]

Trends obtained are shown in figures 3.18 and 3.19 below. Once again two stages are simulated: operation at an equilibrium point from zero to 500 minutes, after which time small steps are applied simultaneously to the set-points of the controlled variables at the 500 minute mark. The two MPCs exhibit good set-point tracking capabilities for the small step changes in the controlled variables (figure 3.18). Although the steps applied are positive, not all the manipulated variables increase positively from their equilibrium positions. The step on L2 causes F2 to change momentarily, before returning to its pre-step value. This is typical of an integrating loop. The MPCs decrease F200 after the steps were applied, indicating that there is an inverse relationship between it and at least one of the controlled variables. The non-minimum-phase phenomenon observed in the P2 responses also shows that its dynamics are affected by at least two of the manipulated variables. The manipulated variables are well within their constraint values and there is no input saturation (figure 3.19)

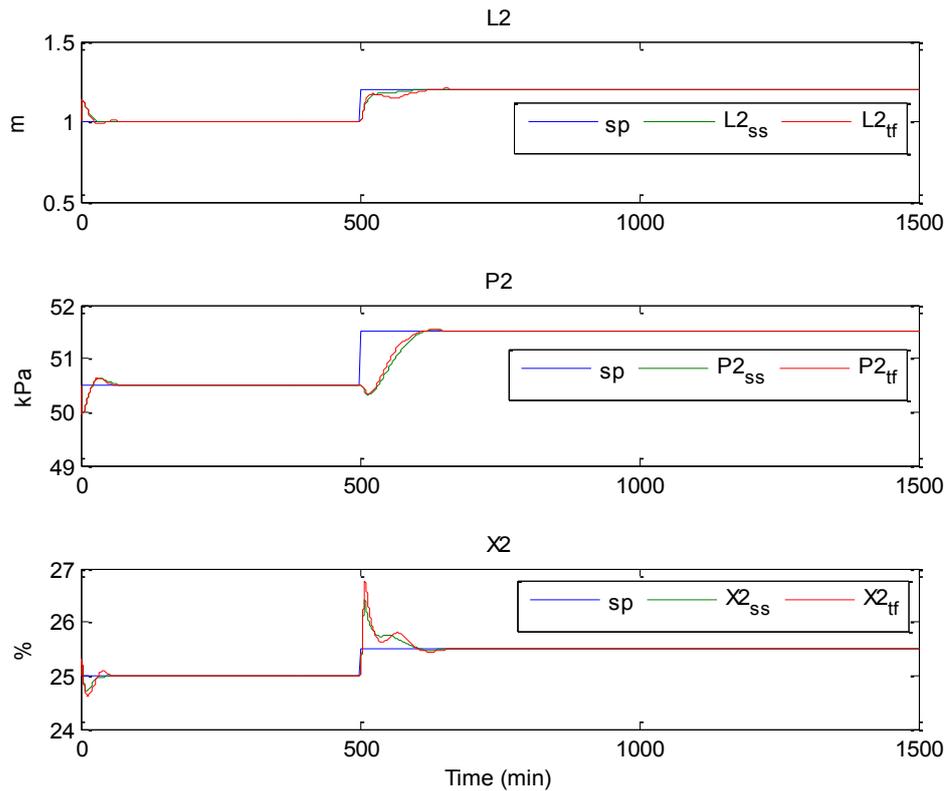


Figure 3.18: CV trends of direct MPC controlled nonlinear evaporator step with changes in set points (sp = set point, subscripts ss = state space, subscripts tf = transfer function)

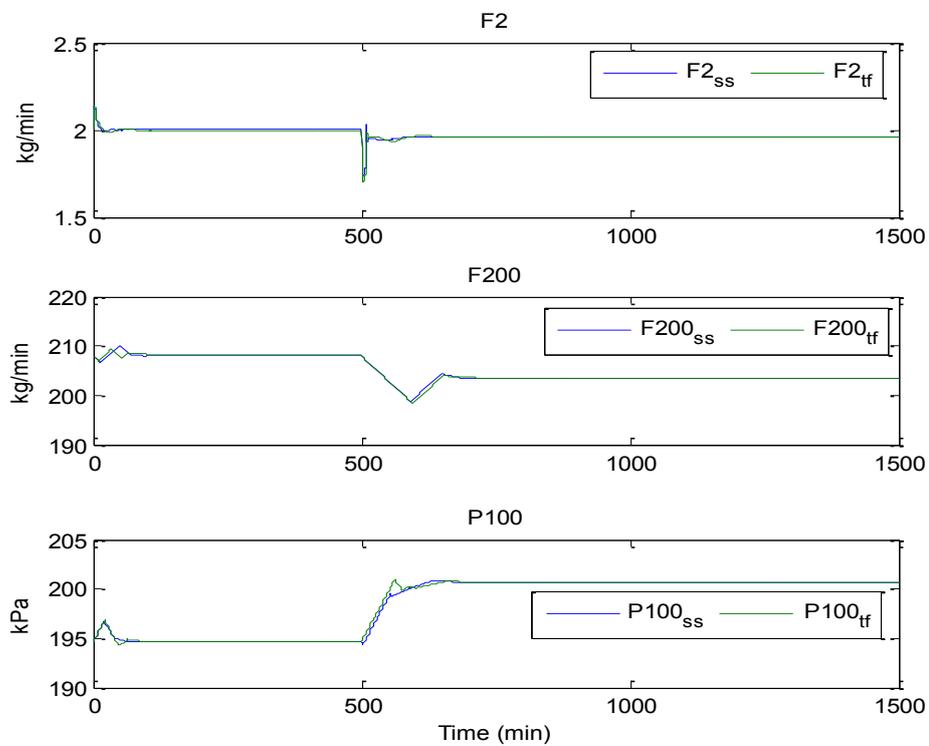


Figure 3.19 MV trends of direct MPC controlled nonlinear evaporator with step changes in set points (subscripts ss = state space, subscripts tf = transfer function)

3.4.3 MPC in a supervisory capacity

In any industrial application, it is practical to expect that the integrating loop F2-L2 of the evaporator would be closed first with the aid of a proportional (P) or proportional plus integral (PI) controller before implementing MPC as supervisory control. In closing the loop F2-L2, the local feedback loop around the valve was removed and a PI regulator is placed between the plant and the MPC. The MPC then manipulates the set-point to the PI controller. The other three local servo operating valves are maintained as before. The response plots to unit step increases in the three manipulated variables are shown in figure 3.20 when only regulatory control is present. The associated closed loop steady state gains for the manipulated variable versus controlled variables loops are shown in Table 3.16.

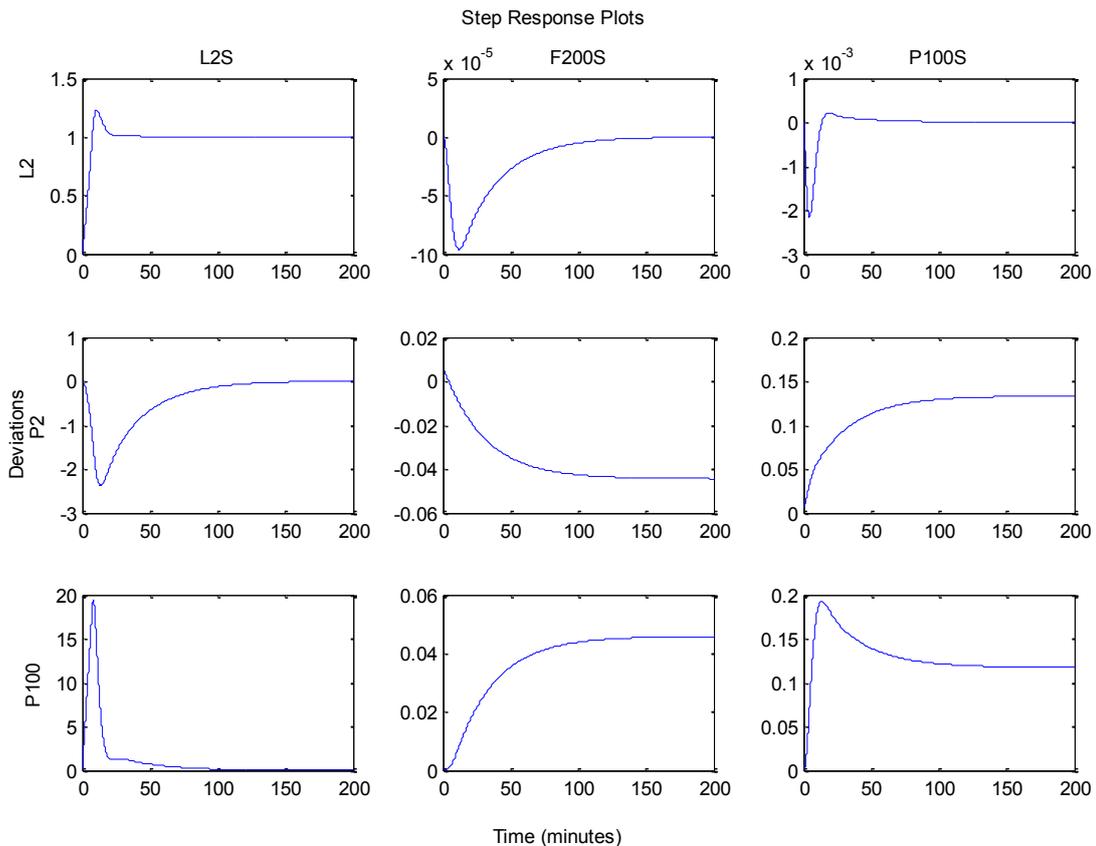


Figure 3.20: Step response under regulatory control

Table 3.16 Steady state gains of the evaporator MV-CV loops for closed loop

	L2set	F200	P100
L2	1.0000	-0.0000	0.0000
P2	0.0053	-0.0445	0.1332
X2	0.0000	0.0458	0.1172

All the loops have finite steady state gain values, which indicate that the plant no longer has any integrating loops. The equivalent low order transfer function matrix identified from the step response data is shown in table 3.17. The linear state space model obtained by linearizing about the operating point of tables 3.9 to 3.11 is given in equation 3.20. The steady state gain for the linear state space model is given in table 3.18. Figure 3.21 compares plots of the step responses of the two linear models (the state space and the transfer function matrix) with the direct step response from the plant.

$$\begin{aligned}
 A = & \begin{bmatrix} -0.1 & 0 & - & 0.625 & -1.25 & 0 & 0 & 0 \\ -0.02091 & -0.0558 & 0 & 0 & 0 & -0.001829 & 0.009588 & 0.03672 \\ 0.004182 & 0.007512 & -0.025 & -0.05 & 0 & -0.001918 & -0.007343 & \\ 0 & 0 & 0.05 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.8333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.8333 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & - & 0.8333 \end{bmatrix}; \\
 B = & \begin{bmatrix} 0.625 & 0 & 0 \\ 0 & 0 & 0 \\ 0.025 & 0 & 0 \\ -0.05 & 0 & 0 \\ 0 & 0.8333 & 0 \\ 0 & 0 & 0.8333 \\ 0 & 0 & 0 \end{bmatrix}; \\
 C = & \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};
 \end{aligned}$$

...(3.20)

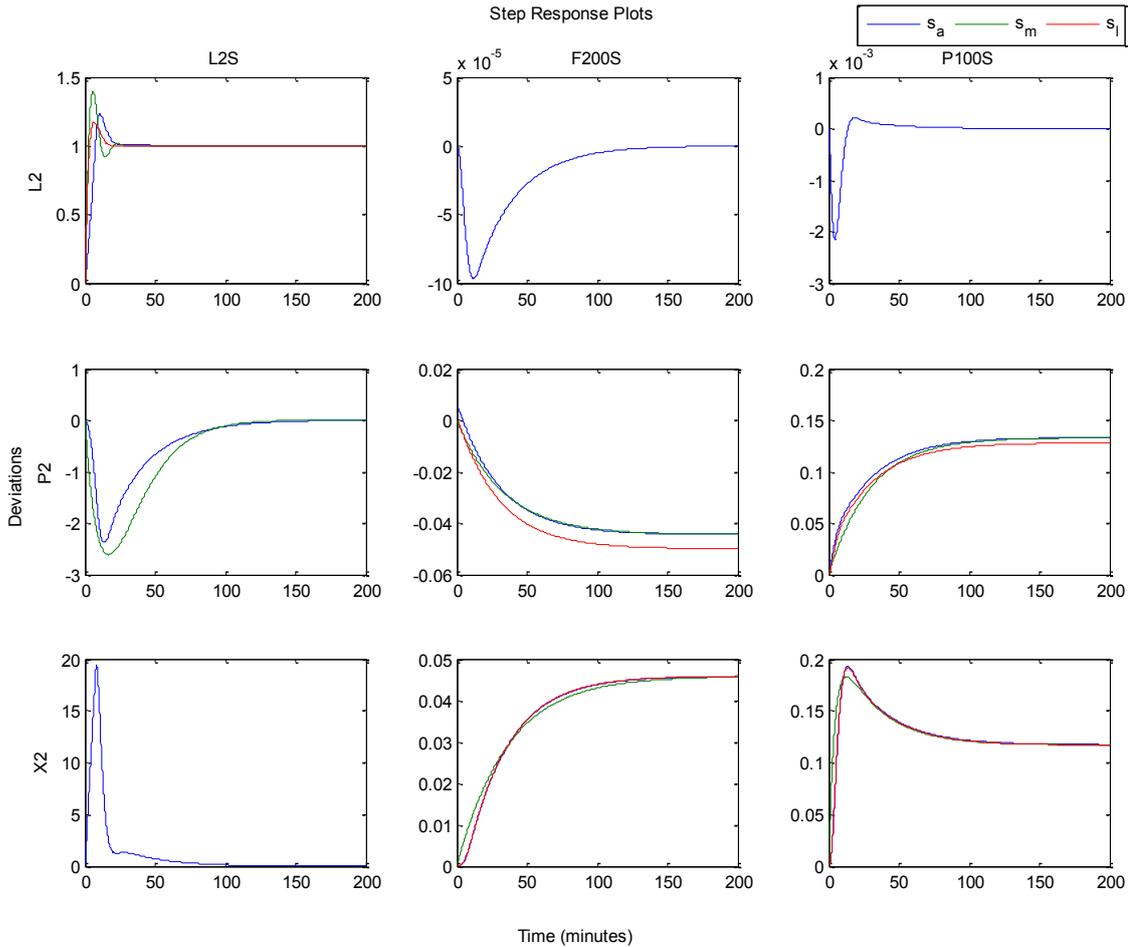


Figure 3.21: Step response plots for the models for the closed loop evaporator (s_a = actual. s_m = reduced model. s_l = linearised state space)

Table 3.17: Reduced transfer function matrix for evaporator under regulatory control

	L2set	F200	P100
L2	$\frac{1 + 2.9s}{1 + 1.15s + 6.25s^2}$	0	0
P2	0	$\frac{-0.073624}{1 + 33.5s}$	$\frac{0.1188}{1 + 29.7s}$
X2	0	$\frac{0.069}{1 + 35.96s}$	$\frac{0.1255(1 + 54s)}{1 + 33s + 121s^2}$

Table 3.18: Steady state gains for the state space model

	L2set	F200	P100
L2	1	0	0.1188
P2	-0.002137	-0.07357	0.164
X2	0.02397	0.069	0.1254

The transient dynamic responses of all the models are very similar, suggesting that the two derived linear models (the state space model and the transfer function matrix) are fair representations of the closed loop dynamics of the evaporator.

The two linear models (state space and transfer function matrix) are implemented in two MPCs that act as supervisory controllers. The output and input trends for the two supervisory MPC implementations are shown in figures 3.22, 3.23 and 3.24 below.

The trends of the controlled variables (figure 3.22), the manipulated variable from the MPC (figure 3.23) and the manipulated variables from the PI controllers and the local servo controllers (figure 3.24) show that using either the low order transfer function model or the state space model give good set point tracking for operation close to the nominal point.

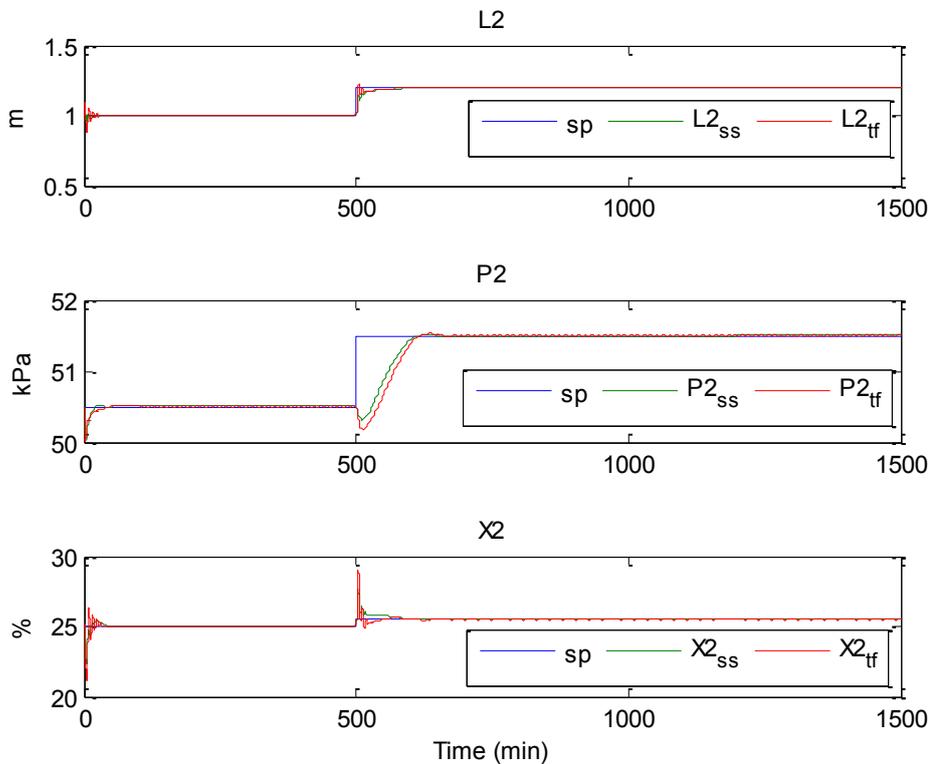


Figure 3.22: Trends of the controlled variables under supervisory MPC (s = set point, subscripts ss = state space, subscripts tf = transfer function)

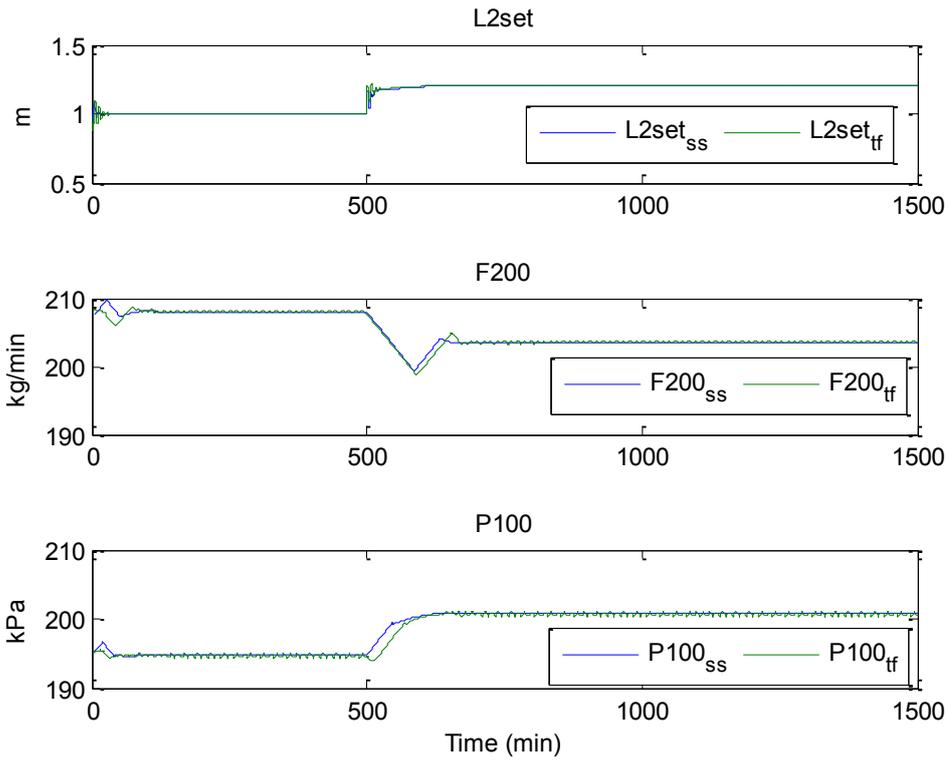


Figure 3.23: Trends of the manipulated variables (MPC outputs) under supervisory MPC (subscripts ss = state space, subscripts tf = transfer function)

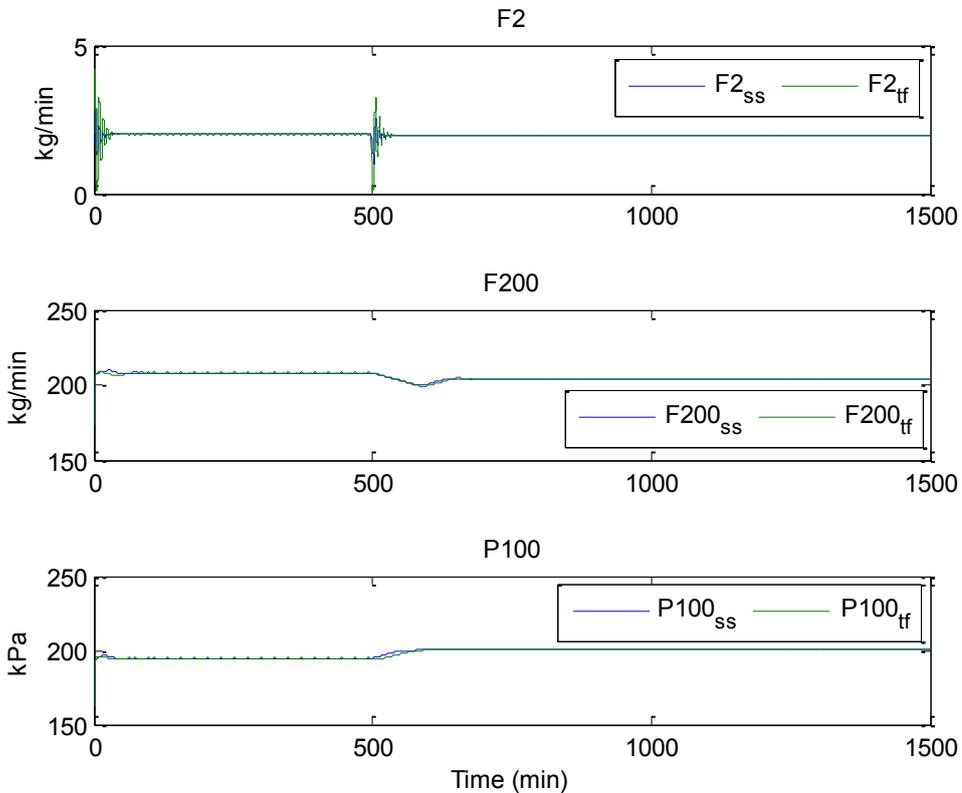


Figure 3.24 PI outputs under supervisory MPC control (subscripts ss = state space, subscripts tf = transfer function)

Figures 3.25 and 3.26 compare the output and input trends of direct and supervisory MPC using the low order transfer function implementation. In both situations, the implementation achieved the objective of driving the outputs to their set points (figure 3.25) though the performance is less oscillatory for direct MPC ($L2_o$, $P2_o$, and $X2_o$) compared with supervisory MPC ($L2_{cl}$, $P2_{cl}$, and $X2_{cl}$). The input trends (figure 3.26) even show more oscillations for supervisory MPC ($F2_{cl}$, $F200_{cl}$, and $P100_{cl}$) compared with direct MPC ($F2_o$, $F200_o$, and $P100_o$).

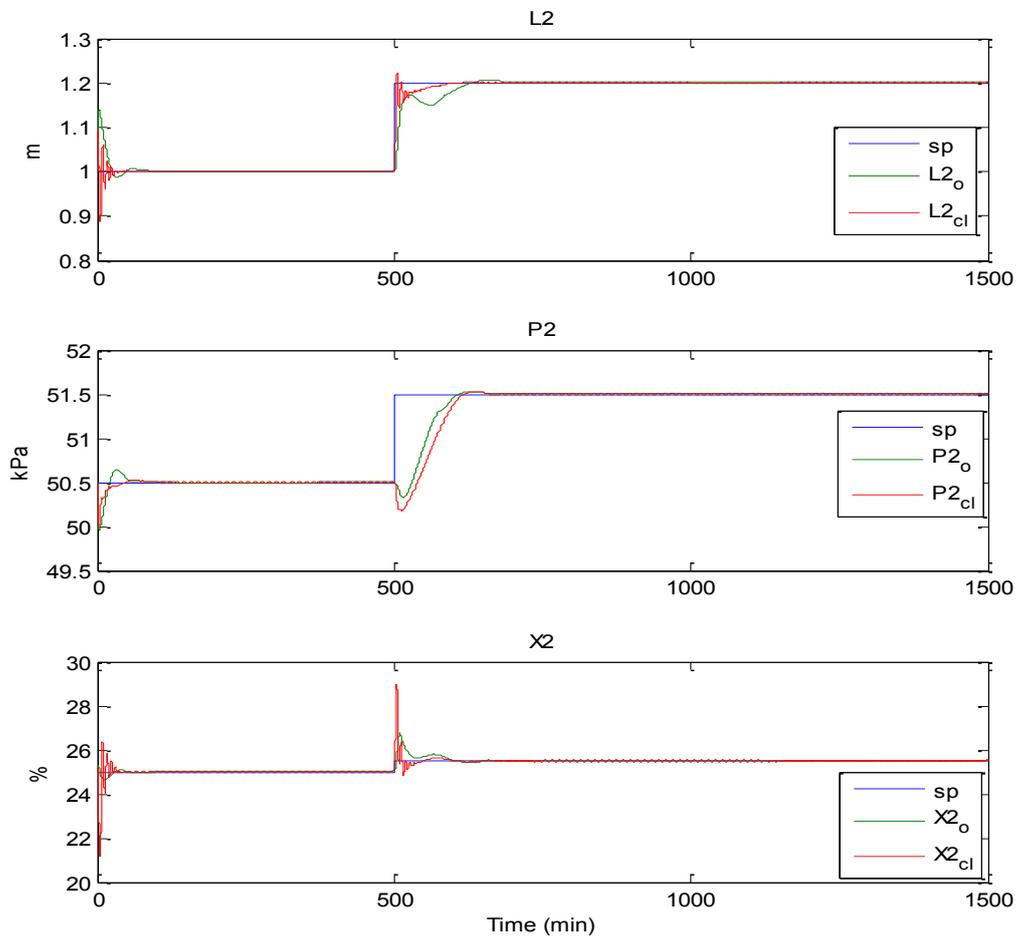


Figure 3.25: Controlled variables trends under direct control and supervisory MPC (subscripts o and cl for open loop and closed loop respectively. sp = set point)

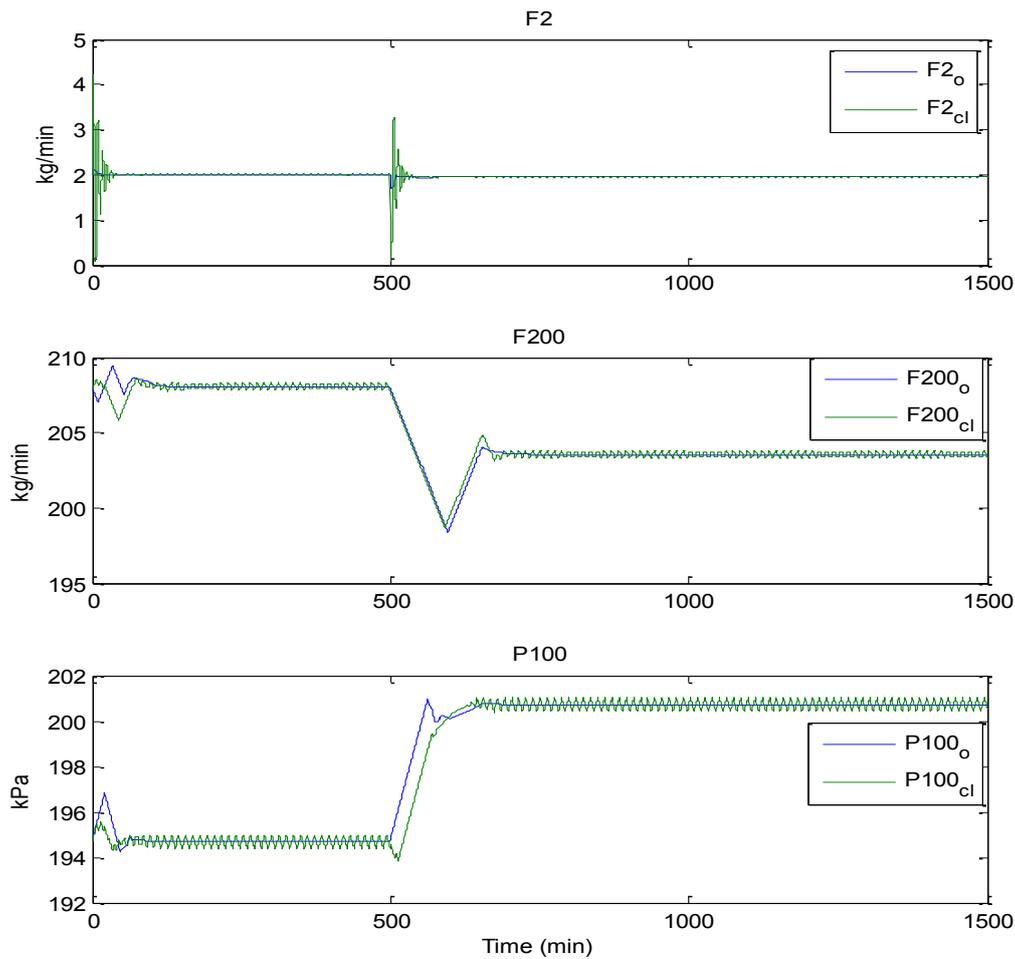


Figure 3.26 Manipulated variables trends under direct control and supervisory MPC (subscripts o and cl for open loop and closed loop respectively)

3.3 The FCCU model and properties

The nonlinear model of the fluid catalytic cracking unit (FCCU) presented by McFarlane et al. (1993) provides a very realistic industrial vehicle for the study of very complex, nonlinear and strongly interacting processes. The model includes important system, equipment and operating constraints typical in real plants.

In the FCCU model, heated fresh feed F3 (made up of a mixture of gas oil F0, wash oil F1 and diesel F2) is mixed with hot slurry recycle F4 (from the bottom of the main fractionator) See figure 3.27.

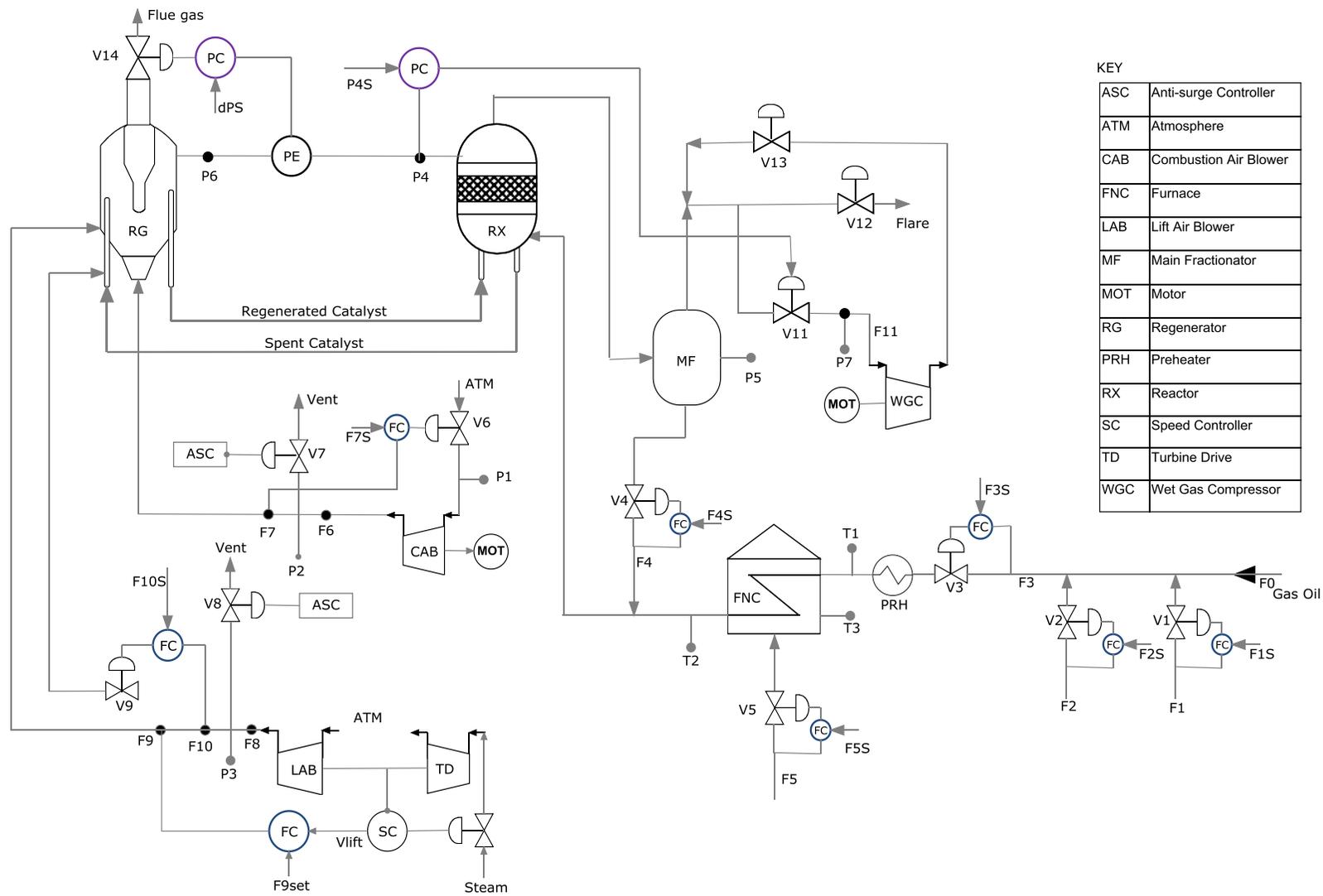


Figure 3.27: Flow diagram of the FCCU model (adapted from McFarlane et al. (1993))

The fresh feed is injected into the reactor riser, where it mixes with a hot regenerated catalyst and totally vaporizes. The hot catalyst provides the sensible heat, heat of vaporization and heat of reaction necessary for the endothermic cracking reactions. As a result of the cracking reactions, a carbonaceous material (coke) is deposited on the surface of the catalyst. Coke on spent catalyst is usually 5-10% hydrogen, depending on the coking characteristics of the feedstock. Since coke poisons the catalyst, continuous regeneration is required.

Separation of catalyst and gas occurs in the disengaging zone of the reactor. The entrained catalyst is removed in cyclones. Catalyst is returned to the stripping section of the reactor where steam is injected to remove entrained hydrocarbons. Reactor product gas is passed to the main fractionator for heat recovery and separation into various product streams. Wet gas from the overheads of the main fractionator (C, and lighter) is compressed for further separation in downstream fractionators. Spent catalyst is transported from the reactor to the regenerator through the spent catalyst U-bend.

Air is injected into the bottom of the regenerator lift pipe to assist the circulation of the catalyst. The catalyst in the regenerator is fluidized with air flow provided by the lift and combustion air blowers. Carbon and hydrogen on the catalyst react with oxygen to produce carbon monoxide, carbon dioxide and water. While most of the reactions occur in the fluidized bed some reaction does occur in the disengaging section above the bed, where some catalyst is still present. Gas travels up the regenerator into the cyclones where the entrained catalyst is removed and returned to the bed.

The regenerator is run at conditions of temperature and excess oxygen to ensure that virtually all carbon monoxide produced in the bed is converted to carbon dioxide before entering the cyclones (referred to as total CO burn). Since little catalyst is present in the

cyclones to absorb generated heat, a significant quantity of carbon monoxide in the cyclones, in the presence of oxygen, can result in sufficient heat generation to produce potentially destructive temperatures.

Regenerated catalyst flows over a weir into the regenerator standpipe. The head produced by catalyst in the standpipe provides the driving force for catalyst flow through the regenerated catalyst U-bend to the reactor riser.

3.3.1 FCCU Simulation

This non-linear model of the FCCU has twenty one governing differential equations and over forty associated algebraic equations. In McFarlane et al. (1993), a total of twelve regulatory controllers were suggested as necessary for lower level regulatory control. They include five regulatory flow controllers (for: the wash oil flowrate F1, the diesel flowrate F2, the fresh feed flowrate F3, the slurry recycle flowrate F4 and the fuel gas flow rate F5), two pressure controllers (for: the reactor pressure P4 and the differential pressure between the reactor and the regenerator ΔP), three air blower controllers (for: the combustion air flow to the regenerator F7, the lift air flow to the regenerator F9, and the spill air flow to the regenerator F10) and two air blower antisurge controllers (for the lift and combustion air blowers). This simulation incorporated all the flow, pressure and air blowers suggested. The antisurge controllers were not included. Table 3.19 below shows the PI regulatory controllers included in this simulation. The dynamics of the regulatory flow controllers were not provided in this model; first-order lags local were implemented with the same time constant of 1.2 seconds each to relate flow set-points to flow. All the other regulatory controllers were implemented as PI controllers. To prevent algebraic loop issues during simulation, lag filters with time constants of 1.2 seconds each were placed in series with the proportional gains of the pressure and air blower PI controllers.

Table 3.19: Parameters of the regulatory controllers

S/N	Loop	Regulatory Controllers	Controller MV	Proportional Gain (P)	Integral/Servo Time Constant (Ti) (Sec)
1	F1S-F1	Flow Controllers	F1		1.2
2	F2S-F2		F2		1.2
3	F3S-F3		F3		1.2
4	F4S-F4		F4		1.2
5	F5S-F5		F5		1.2
6	F7S-F7	Air Blower Controllers	V6	0.002	5
7	F9S-F9		Vlift	0.02	5
8	F10S-F10		V9	2.0	5
9	P4S-P4	Pressure Controllers	V11	-0.01	10
10	Δ PS- Δ P		V14	-0.001	5

The input (manipulated and disturbance) and output variables, as well as their nominal values, as used or obtained from the simulations, are given in tables 3.20, 3.21 and 3.22 below.

Table 3.20: Manipulated input variables

S/N	Variables	Symbol	Nominal Values	Units
1	Wash oil flow set point	F1S	13.80	Ib/s
2	Diesel flow set point	F2S	0.00	Ib/s
3	Fresh feed flow set point	F3S	126.00	Ib/s
4	Slurry recycle flow set point	F4S	5.25	Ib/s
5	fuel gas flow set point	F5S	34.00	Scf/s
6	Combustion air blower flowrate	F7S	61.0	Ib/s
7	Lift air flow to the regenerator set point	F9S	14.5	Ib/s
8	Spill air flowrate	F10S	0.0	Ib/s
9	Reactor/regenerator diff pressure set point	Δ PS	-3.38	psia
10	Reactor pressure set point	P4S	33.00	psia

Table 3.21: Disturbance input variables

S/N	Variables	Symbol	Nominal Value	Unit
1	Ambient air Temperature (unmeasured disturbance)	Tatm	75.0	^o F
2	Effective coking factor (unmeasured disturbance)	Ψ f	1.0	
3	Preheat outlet temperature (measured disturbance)	T1	460.9	^o F

Table 3.22: Output variables

S/N	Variables	Symbols	Nominal Values
1	Riser temperature	Tr	995.4
2	Furnace firebox temperature	T3	1608
3	Regenerator cyclone Temperature	Tcyc	1283
4	Cyclone/Regenerator temperature difference	Tdiff	10.27
5	Regenerator Temperature	Treg	1272
6	Concentration of oxygen in the regenerator stack gas	C _{O₂,sg}	1.479
7	Concentration of carbon monoxide in the stack gas	C _{CO,sg}	78.08
8	Regenerator standpipe level	Lsp	11.37
9	Lift air blower suction and surge flow difference	Flab	2134
10	Combustion air blower suction and surge flow difference	Fcab	4451
11	Wet gas compressor suction and surge flow difference	Fwgc	7308
12	Total air flowrate to the regenerator	Ft	75.5
13	Reactor Pressure	P4	33.0
14	Reactor/regenerator differential pressure	ΔP	-3.38
15	Combustion air blower suction pressure	P1	
16	Combustion air blower discharge pressure	P2	
17	Lift air blower discharge pressure	P3	
18	Main fractionator pressure	P5	
19	Regenerator pressure	P6	
20	Wet gas compressor suction pressure	P7	
21	Temperature of fresh feed entering the reactor riser	T2	
22	Combustion air blower throughput	F6	
23	Combustion air flow to the regenerator	F7	
24	Lift air blower throughput	F8	
25	Lift air flow to the regenerator	F9	
26	Spill air flow to the regenerator	F10	
27	Wet gas flow to the vapour recovery unit	F11	
28	Air flowrate into the regenerator	Fair	
29	Speed of the lift air blower	sa	

The FCCU process was simulated with regulatory controllers only, using the nominal input variables specified in tables 3.20 and 3.21, and with the initial values of the state variables specified in table 3.23 below. The trends of the process outputs and inputs (regulatory controller outputs) obtained from the simulation are shown in figures 3.28 and 3.29 respectively. The trends indicate that the process is stable under regulatory control. The simulation confirmed that the equilibrium values are as given in table 3.22 above. The equilibrium values for the valves opening are as given in table 3.24 below.

Table 3.23: Initial values of state variables

S/N		Output Variables	Initial Values
1	Treg	Regenerator Temperature	1271.96
2	T3	Furnace firebox temperature	1607.55
3	Tr	Riser temperature	995.13
4	Wreg	Regenerator catalyst inventory	273,742.7
5	Wc	Regenerator carbon inventory	1297.62
6	Crgc	Carbon concentration on regenerated catalyst	8.7296E-4
7	Wsp	Regenerator standpipe catalyst inventory	3566.80
8	P6	Regenerator pressure	29.64
9	n	Moles of gas in regenerator	245.92
10	Plift	Catalyst density in lift pipe	3.251
11	P2	Combustion air blower discharge pressure	35.19
12	P3	Lift air blower discharge pressure	40.50
13	Csc	Concentration of coke on spent catalyst	1283
14	Wr	Reactor catalyst inventory	10.27
15	P5	Main fractionator pressure	1272
16	P7	Wet gas compressor suction pressure	1.479
17	T2	Fresh feed temperature	78.08
18	P1	Combustion air blower suction pressure	14.63

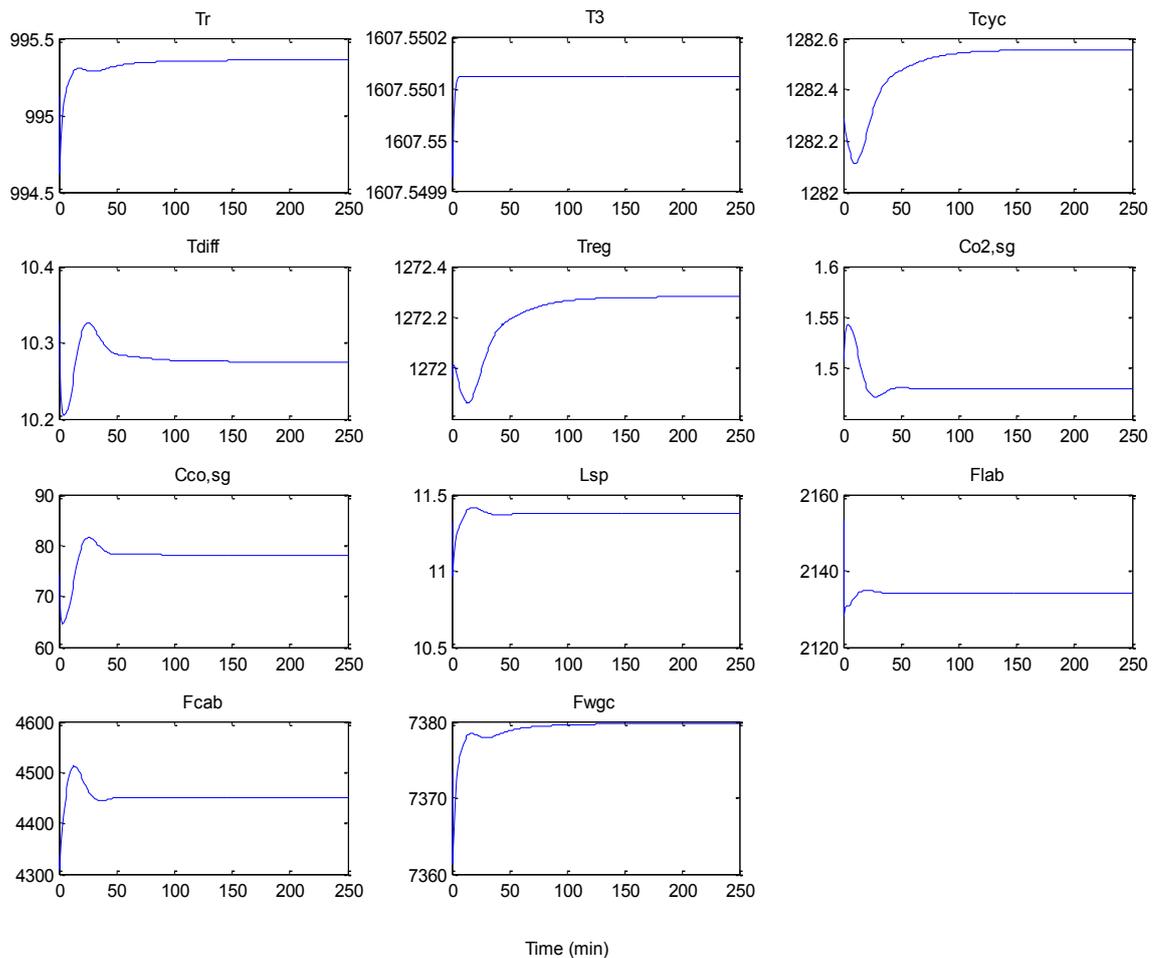


Figure 3.28: Output trends of the FCCU under regulatory control

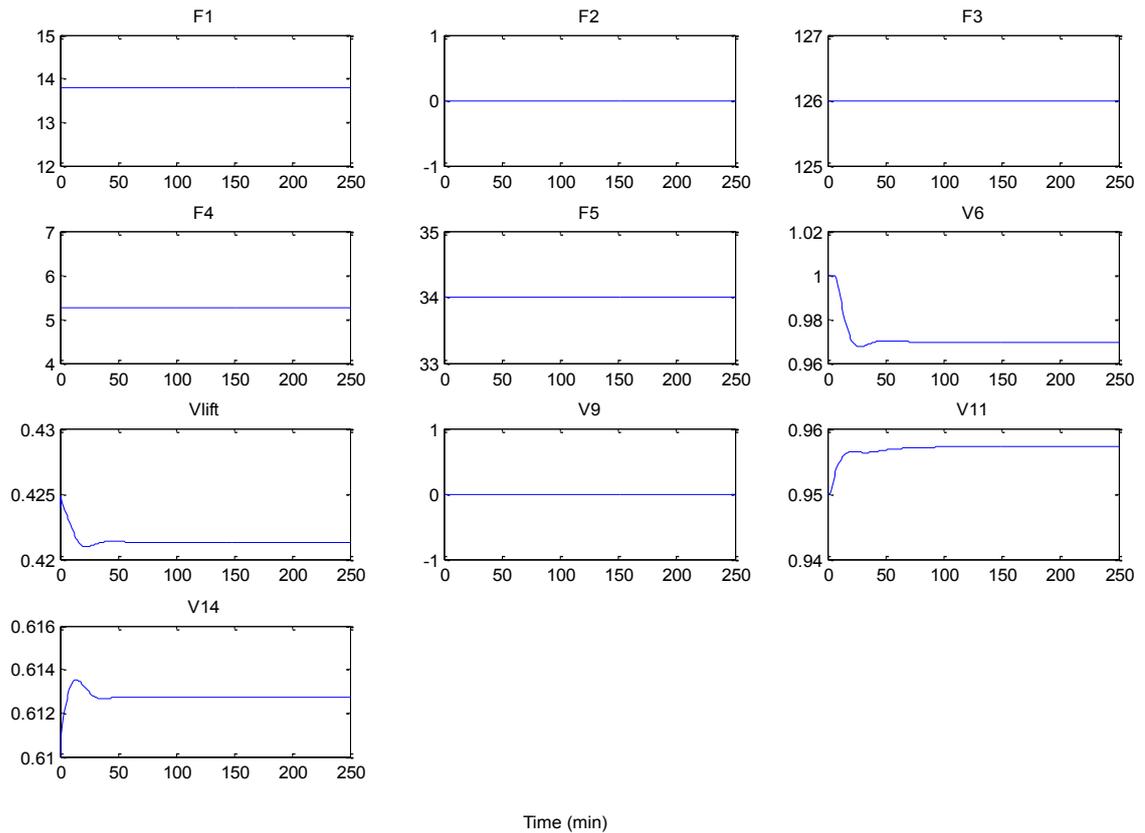


Figure 3.29: Input trends of the FCCU under regulatory control

Table 3.24: Equilibrium valve openings

S/N		Valve openings	Nominal values
1	V6	Combustion air blower suction valve position	0.9696
2	Vlift	Lift air blower steam valve position	0.4213
3	V9	Spill air valve position	0
4	V11	Wet gas compressor suction valve position	0.9573
5	V14	Stack gas valve position	0.6127

3.3.2 FCCU system identification

The equilibrium values of tables 3.22, 3.23 and 3.24 are used to simulate the process for the purpose of system identification and subsequent MPC implementation.

All the manipulated input variables of table 3.20 and the first fourteen output variables of table 3.22 (the reason is explained later in this section) were used in the identification. Similar to the procedure adopted for the two previous process examples, the state space model of the FCCU with these inputs and outputs was obtained by

linearization about the equilibrium values. Then by applying unit steps to the state space model, the equivalent low order transfer function models were obtained for all input output pairs by applying the simple direct procedure described in Appendix A. Steady state gain values for the FCCU are shown in table 3.25 below.

Table 3.25: Steady gain values for the FCCU (from the identified transfer function models)

	F1S	F2S	F3S	F4S	F5S	F7S	F9S	F10S	P4S	ΔPS
Tr	2.50	-1.00	-0.24	2.45	1.08	1.61	2.31	1.61	-5.68	-16.68
T3	0	0	-0.71	0	27.32	0	0	0	0	0
Treg	6.44	-2.58	1.78	8.38	0.93	1.29	2.32	1.29	-7.98	-24.10
Tdiff	-0.41	0.16	-0.02	-0.45	-0.13	-0.37	-0.35	-0.37	0.11	-0.27
Tcyc	6.85	-2.74	1.80	8.83	1.06	1.67	2.67	1.67	-8.09	-23.84
C _{O2,Sg}	-0.30	0.12	-0.13	-0.43	-0.01	0.01	-0.08	0.01	0.70	2.07
C _{CO,Sg}	32.75	-13.10	19.08	51.33	-2.86	-4.88	9.64	-4.88	-147.52	-375.70
Lsp	0.10	-0.04	-0.002	0.11	0.02	0.37	0.50	0.37	-1.05	-6.27
Flab	0.05	-0.02	-0.05	0.01	0.05	-0.26	682.05	806.11	-114.31	-108.71
Fcab	5.39	-2.15	1.40	6.94	0.83	637.49	19.43	19.37	-1283.9	-1284.8
Fwgc	63.48	-25.39	35.97	104.30	27.42	40.96	58.72	40.96	-144.16	-423.35

Table 3.26 below shows the identified transfer function matrix for the FCCU. Figure 3.30 shows the step response plots of the linear state space model (bold line) and the step response of the transfer function model (dotted line).

Both the step response plots and the transfer function matrix show that the settling time for the FCCU system varies from as low as 41 seconds (0.7 minute), between F5S and T3, to 17472 seconds (291 minutes), between P4S and Treg. The steady state gain values (shown in table 3.26) range from 0 to 1285 (absolute). It shows that the furnace temperature's dynamics is dependent almost solely on the furnace fuel flowrate F5; the total fresh feed flowrate F3 s contributes little.

3.3.3 MPC Control of the FCCU

The approach adopted in the control of the FCCU process is to make the MPC hold the riser temperature T_r at set-point while implementing all the outputs as zone control, and maximizing the total feedrates to the system, subject to the input and constraints given in tables 3.27 and 3.28 below. This strategy is necessary because of the very complex, mildly nonlinear and highly coupled nature of the process. In the zone control, the output weights on all the output variables except T_r are set to zero. The MPC implemented here has no way of output constraint ranking. As long as the other ten output variables are within specified boundaries, their exact values are of less importance.

The manipulated and disturbance inputs are the ones given in tables 3.21 and 3.22 respectively. The controlled output variables, with their nominal values, are the first eleven variables of table 3.20.

The FCCU control problem in some way supports the decision to use state space MPC instead of step response MPC, despite the transparency of the identified step response model. From the transfer function matrix (table 3.26) and the step response plot (figure 3.30), the dominant setting time for the system is about 17500 seconds (291 minutes). This means that given a sampling time (T_s) of 60 seconds (1 minute) for example, the step response model requires at least 291 step response coefficients (N) for the step response MPC algorithm.

Table 3.26: Identified Transfer Function Matrix for the FCCU

	F_{1s}	F_{2s}	F_{3s}	F_{4s}	F_{5s}	F_{7s}	F_{9s}	F_{10s}	P_{4s}	ΔP_s
T_r	$\frac{2.48}{41s+1}$	$\frac{-1}{41s+1}$	$\frac{-0.24(1+205s)}{84s^2+46s+1}$	$\frac{2.45}{43s+1}$	$\frac{1.08(1+27s)}{51s^2+50s+1}$	$\frac{1.61(1+33s)}{56s^2+60s+1}$	$\frac{2.31(1+23s)}{69s^2+50s+1}$	$\frac{1.61(1+27s)}{44s^2+53s+1}$	$\frac{-5.68}{60s+1}$	$\frac{-16.68(1+15s)}{34s^2+47s+1}$
T_3	0	0	$\frac{-0.71}{s+1}$	0	$\frac{27}{0.17s+1}$	0	0	0	0	0
T_{reg}	$\frac{6.44}{34s+1}$	$\frac{-2.58}{34s+1}$	$\frac{1.78}{43s+1}$	$\frac{8.38}{36s+1}$	$\frac{0.93}{55s+1}$	$\frac{1.29(1-21s)}{34s^2+47s+1}$	$\frac{2.32}{45s+1}$	$\frac{1.29(1-22s)}{34s^2+47s+1}$	$\frac{-7.98}{64s+1}$	$\frac{-24.1}{40s+1}$
T_{diff}	$\frac{-0.41(1-217s)}{110s^2+24s+1}$	$\frac{0.16(1-250s)}{112s^2+24s+1}$	$\frac{-0.02(1+53s)}{136s^2+28s+1}$	$\frac{-0.45(1+533s)}{121s^2+23s+1}$	$\frac{-0.13}{6s+1}$	$\frac{-0.37(1+s)}{s^2+s+1}$	$\frac{-0.35(1+s)}{0.84s^2+s+1}$	$\frac{-0.37(1+s)}{1.6s^2+s+1}$	$\frac{0.11}{15s+1}$	$\frac{-0.27(1+43s)}{82s^2+28s+1}$
T_{cyc}	$\frac{6.85}{34s+1}$	$\frac{-2.74}{34s+1}$	$\frac{1.80}{45s+1}$	$\frac{8.83}{37s+1}$	$\frac{1.06}{51s+1}$	$\frac{1.67}{48s+1}$	$\frac{2.67}{46s+1}$	$\frac{1.67}{48s+1}$	$\frac{-8.09}{73s+1}$	$\frac{-23.84}{42s+1}$
$C_{o,sg}$	$\frac{-0.3}{13s+1}$	$\frac{0.12(1+75s)}{384s^2+96s+1}$	$\frac{-0.13}{27s+1}$	$\frac{-0.43(1+58s)}{391s^2+84s+1}$	$\frac{-0.01(1-107s)}{274s^2+52s+1}$	$\frac{0.01(1+400s)}{44s^2+40s+1}$	$\frac{-0.08(1-41s)}{25s^2+40s+1}$	$\frac{0.01(1+410s)}{25s^2+40s+1}$	$\frac{0.70}{42s+1}$	$\frac{2.07(1+25s)}{136s^2+44s+1}$
$C_{CO,sg}$	$\frac{32.8(1+12s)}{54s^2+9s+1}$	$\frac{-13(1+13s)}{54s^2+10s+1}$	$\frac{19.08}{9s+1}$	$\frac{51.33(1+5s)}{37s^2+7s+1}$	$\frac{-2.86(1+130s)}{225s^2+47s+1}$	$\frac{-4.88(1+157s)}{34s^2+35s+1}$	$\frac{9.64(1-54s)}{40s^2+38s+1}$	$\frac{-4.88(1+150s)}{17s^2+33s+1}$	$\frac{-147.52}{32s+1}$	$\frac{-375.7}{0.2s+1}$
L_{sp}	$\frac{0.1(1-14s)}{1273s^2+72s+1}$	$\frac{-0.04(1-12s)}{1469s^2+74s+1}$	$\frac{-0.002(1+2505s)}{44s^2+67s+1}$	$\frac{0.11(1-7s)}{114s^2+64s+1}$	$\frac{0.02(1-145s)}{25s^2+40s+1}$	$\frac{0.37(1+46s)}{54s^2+59s+1}$	$\frac{0.5(1+48s)}{64s^2+64s+1}$	$\frac{0.37(1+52s)}{44s^2+67s+1}$	$\frac{-1.05}{44s+1}$	$\frac{-6.27}{0.3s+1}$
F_{lab}	$\frac{0.05(1-16667s)}{1003s^2+63s+1}$	$\frac{-0.02(1-27500s)}{1004s^2+64s+1}$	$\frac{-0.05(1+6672s)}{100s^2+60s+1}$	$\frac{0.01(1+30333s)}{114s^2+60s+1}$	$\frac{0.05(1-5583s)}{56s^2+45s+1}$	$\frac{-0.26(1+1633s)}{27s^2+52s+1}$	$\frac{682}{s+1}$	$\frac{806.11}{s+1}$	$\frac{-114.31}{39s+1}$	$\frac{-108.7(1+s)}{0.09s^2+0.5s+1}$
F_{cab}	$\frac{5.39(1-1517s)}{995s^2+65s+1}$	$\frac{-2.15(1-1533s)}{994s^2+64s+1}$	$\frac{1.40(1+2001s)}{117s^2+65s+1}$	$\frac{6.94(1-1750s)}{910s^2+66s+1}$	$\frac{0.83(1-3617s)}{56s^2+45s+1}$	$\frac{637.49}{2s+1}$	$\frac{19.43(1-300s)}{44s^2+53s+1}$	$\frac{19.37(1+167s)}{44s^2+53s+1}$	$\frac{-1283.9}{39s+1}$	$\frac{-1284.8(1+2s)}{0.11s^2+24s+1}$
F_{wgc}	$\frac{63.48}{40s+1}$	$\frac{-25.39}{40s+1}$	$\frac{35.97(1+20s)}{178s^2+51s+1}$	$\frac{104.3}{42s+1}$	$\frac{27.42(1+37s)}{156s^2+63s+1}$	$\frac{40.96(1+21s)}{64s^2+48s+1}$	$\frac{58.72(1+25s)}{44s^2+53s+1}$	$\frac{40.96(1+27s)}{44s^2+53s+1}$	$\frac{-144.16}{60s+1}$	$\frac{-423.35(1+15s)}{34s^2+47s+1}$

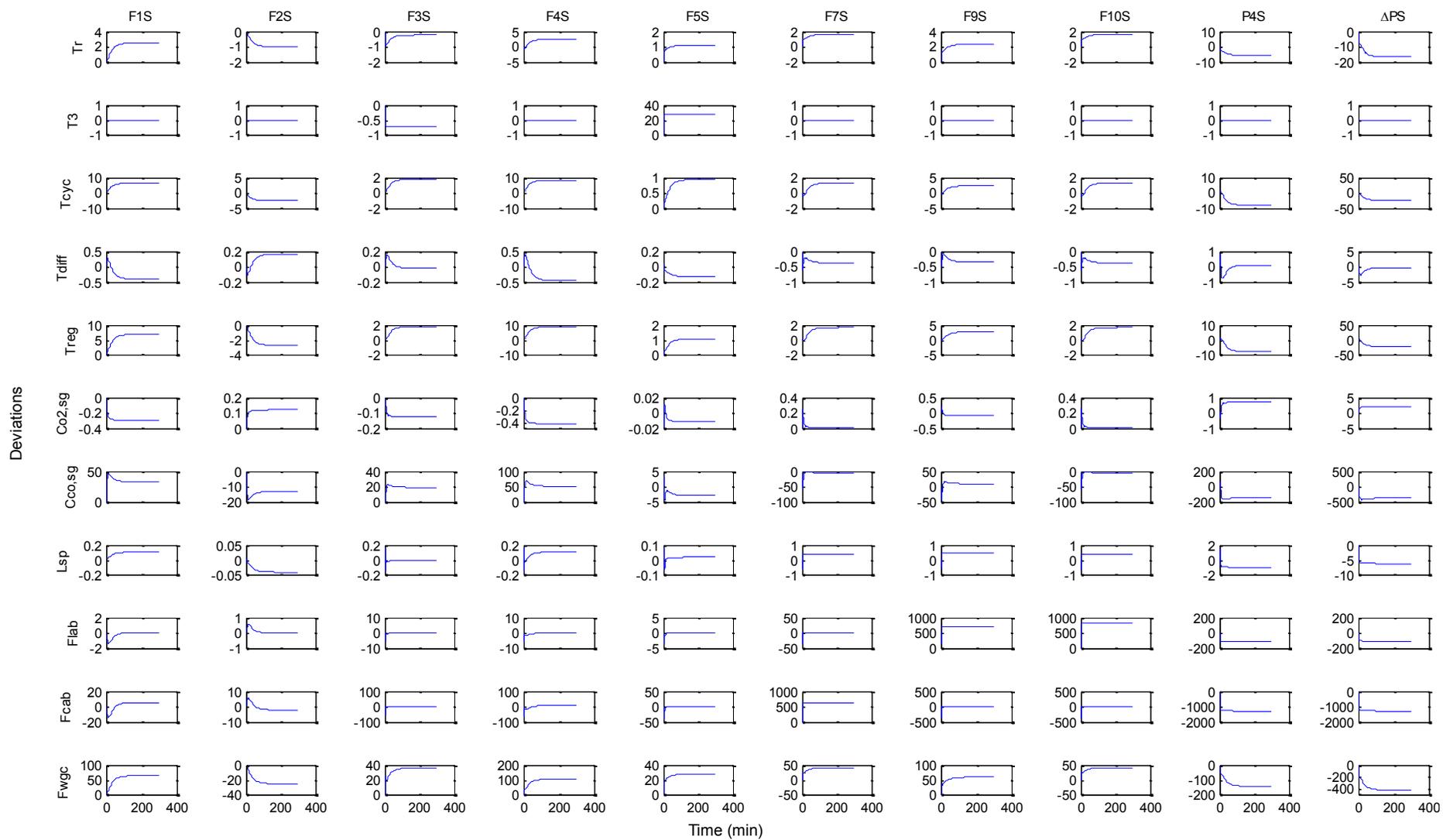


Figure 3.30: Step response plots for the FCCU

Table 3.27: Constraints on outputs (soft, operating constraints)

S/N		Output Variables	Constraints
1	F_{lab}	Lift air blower surge limit	$F_{sucn,lift} - F_{surge,lift} > 0$
2	F_{cab}	Combustion air blower surge limit	$F_{sucn,comb} - F_{surge,comb} > 0$
3	F_{wgc}	Wet gas compressor surge limit	$F_{sucn,wg} - F_{surge,wg} > 0$
4	$C_{O2,sg}$	Concentration of oxygen in the regenerator stack gas	$C_{O2,sg} > 1.5\%$
5	$C_{CO,sg}$	Concentration of carbon monoxide in the stack gas	$C_{CO,sg} \leq 350 \text{ ppm}$
6	T_{reg}	Regenerator Temperature	$T_{reg} \geq 1265^{\circ}F$
7	T_{diff}	Cyclone/Regenerator temperature difference	$T_{diff} = T_{cyc} - T_{reg} \leq 20^{\circ}F$
8	T_{cyc}	Regenerator cyclone Temperature	$T_{cyc} \leq 1310^{\circ}F$
9	dP	Reactor/regenerator differential pressure	$-5.0 \leq \Delta P \leq 2.0 \text{ psi}$
10	L_{sp}	Regenerator standpipe level	$0 \leq L_{sp} \leq 20 \text{ ft}$
11	T_r	Riser temperature	$T_r \leq 995^{\circ}F$
12	T_3	Furnace firebox temperature	$T_3 \leq 1700^{\circ}F$

Table 3.28: Constraints on inputs (hard, equipment constraints)

S/N		Input Variables	Constraints
1	F1	Wash oil flow rate	$0 \leq F1S \leq 17$
2	F2	diesel flow rate	$0 \leq F2S \leq 16$
3	F3	fresh feed flow rate	$0 \leq F3S \leq 144$
4	F4	slurry recycle flow rate	$0 \leq F4S \leq 10$
5	F5	fuel gas flow rate	$0 \leq F5S \leq 40$
6	F7	combustion air flow to the regenerator	$0 \leq F7S \leq 61.5$
7	F9	lift air flow to the regenerator	$0 \leq F9S \leq 41$
8	F10	spill air flow to the regenerator	$0 \leq F10S \leq 40$
9	P4	Reactor maximum pressure	$P4 < 49.2$
10	P5	Main Fractionator maximum pressure	$P5 < 39.7$
11	P6	Regenerator maximum Pressure	$P6 < 39.7$

For this multivariable system with 11 outputs and 10 inputs, the memory requirement for storing the step response MPC algorithm was huge, and the huge matrices caused the computations to progress very slowly. Also the number of step response coefficients is well outside the range $30 \leq N \leq 120$ recommended by Seborg et al. (2010). Of course the memory requirement could be reduced by using a sampling time as high as 540 seconds (9 minutes), resulting in 30 step response coefficients, which greatly

reduces the memory space requirement, thereby speeding up computations. The downside is that a high value of T_s is not advisable, given that there are input-output channels with very fast dynamics. Again when the transfer function matrix is converted directly to state space, the size of the state becomes huge (263). It is observed that this is not even a good choice especially if the transfer function matrix includes input delays.

Without losing much of the transparency offered by the identified transfer function matrix, the transfer function matrix is converted to an appropriate low order state space by using the algorithm described in chapter two. First the unit step is applied to the transfer function matrix model to obtain step response coefficients at intervals of 60 seconds (1 minute). Then the step response coefficients are converted to a corresponding pulse response matrix. The matrix of pulse response is then assembled into a block Hankel matrix. Then the Hankel matrix is used to obtain the appropriate matrices A , B and C of a state space model of order r ($r = 3500$) which most approximates the transfer function matrix. The singular values ($\sigma_1, \sigma_2, \dots, \sigma_r$) of the Hankel matrix (of rank r) are obtained through singular value decomposition. Figure 3.31 shows the plot of the first 30 singular values of the Hankel matrix. On the figure, the number of singular values at the point where the plot becomes asymptotic to zero corresponds to the least order to which the original state space model of order r may be reduced to. The plot indicates that a state space model having 12 states may be derived. Following the procedure outline in the algorithm of Section 2.14, the matrices A , B , and C are truncated appropriately. In this case the state space equation is truncated to an order of 20.

The plots of the step responses of the approximated 20 order state space equation and the identified step response are compared in figure 3.32 below. It is noted that for most cases the approximate state space equation response follows the intricate patterns of the step response of the original transfer function matrix.

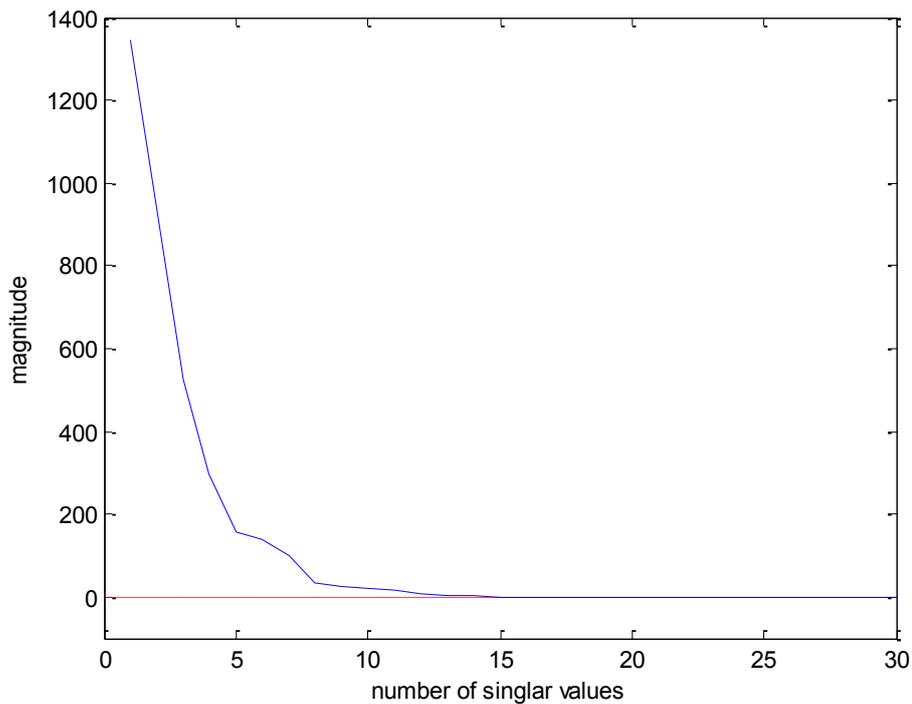


Figure 3.31: Plot of the first 30 singular values of the Hankel matrix for the FCCU step response model

The estimated state space model is used in the MPC which is implemented on the FCCU simulation. The plots of the trends of the controlled variables and the manipulated variables from the simulations are shown in figures 3.33 to 3.35. Again simulation of the supervisory MPC implementation is in two parts: the first part lasts for 120 minutes and aims to keep Tr at a constant value of 995⁰F. At the 120 minute mark, the set-point of Tr is stepped down to 994⁰F. Throughout all the other controlled variables were allowed to operate in zones, meaning that their output weights were all equal to zero.

The plots of figures 3.33 show that after the unit step was applied to the set-point of T_r , it took about 10 minutes for the transient dynamics of T_r to die out and for it to track its new set-point. Figure 3.34 shows that all the manipulated variables respond to the step change in the set point of T_r , but almost all of them, except ΔP return to their pre-step values after the transients have died down. Pressure differential ΔP settled to a higher value over its pre-step one. This is as predicted in the step response matrix of table 3.26 and figure 3.30 and 3.32. All the manipulated variables were well within their constraints. The entire zone controlled output variables were also within their constraint values.

3.4 Summary

This chapter describes the simulation of the three nonlinear plants (CSTR, evaporator and the FCCU), their control using PI regulators, and their identifications using two approaches: direct linearization about their operating points to obtain their state space models, and the derivation of low order transfer function approximations of the state space models. Comparison of the step response plots of the two models with direct step response trends of the plants showed that they are adequate. The models were used as internal model to design the MPC implemented on the plants. The trends obtained from the simulations show that all the plant exhibit good performance when operated close to their specified operating points.

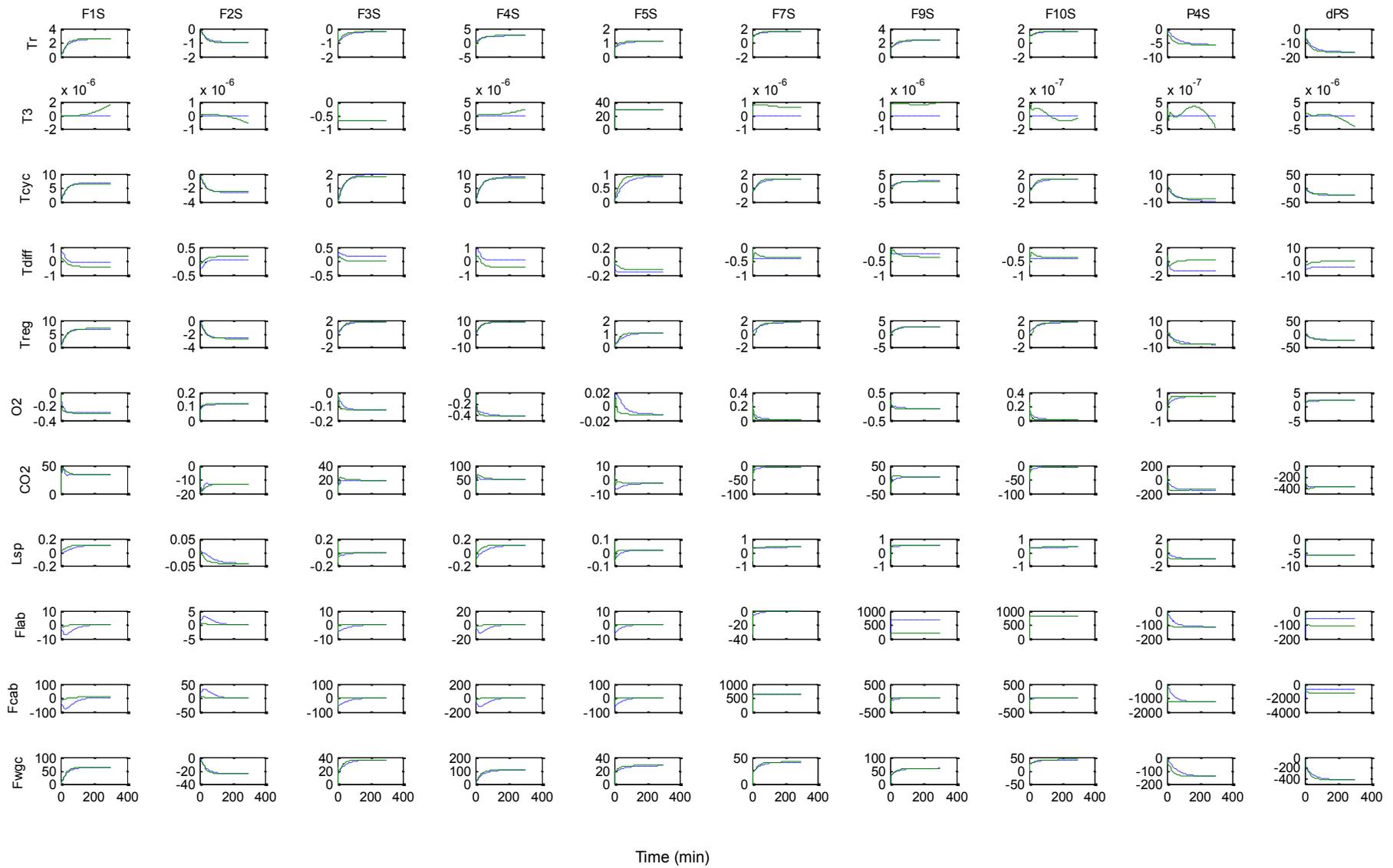


Figure 3.32: Comparison of reduced model response (green) with the original response (blue)

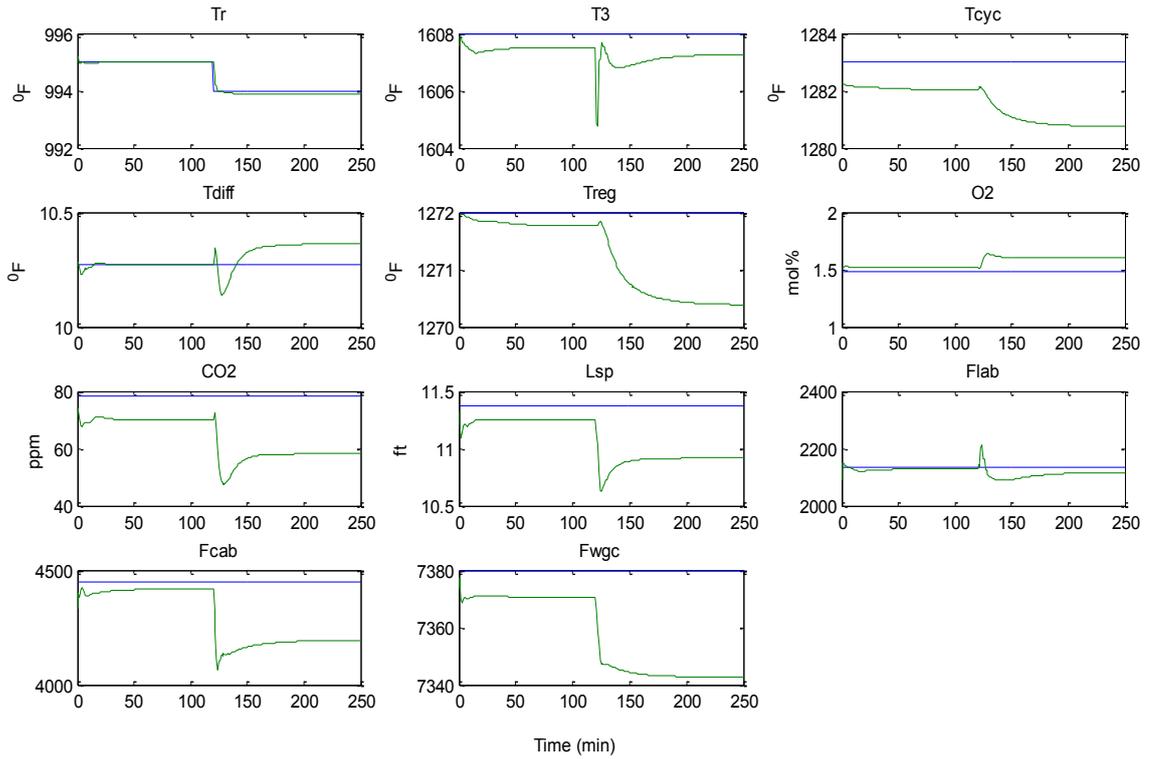


Figure 3.33: Trends of the controlled variables (green lines) and their nominal values (blue lines) under MPC

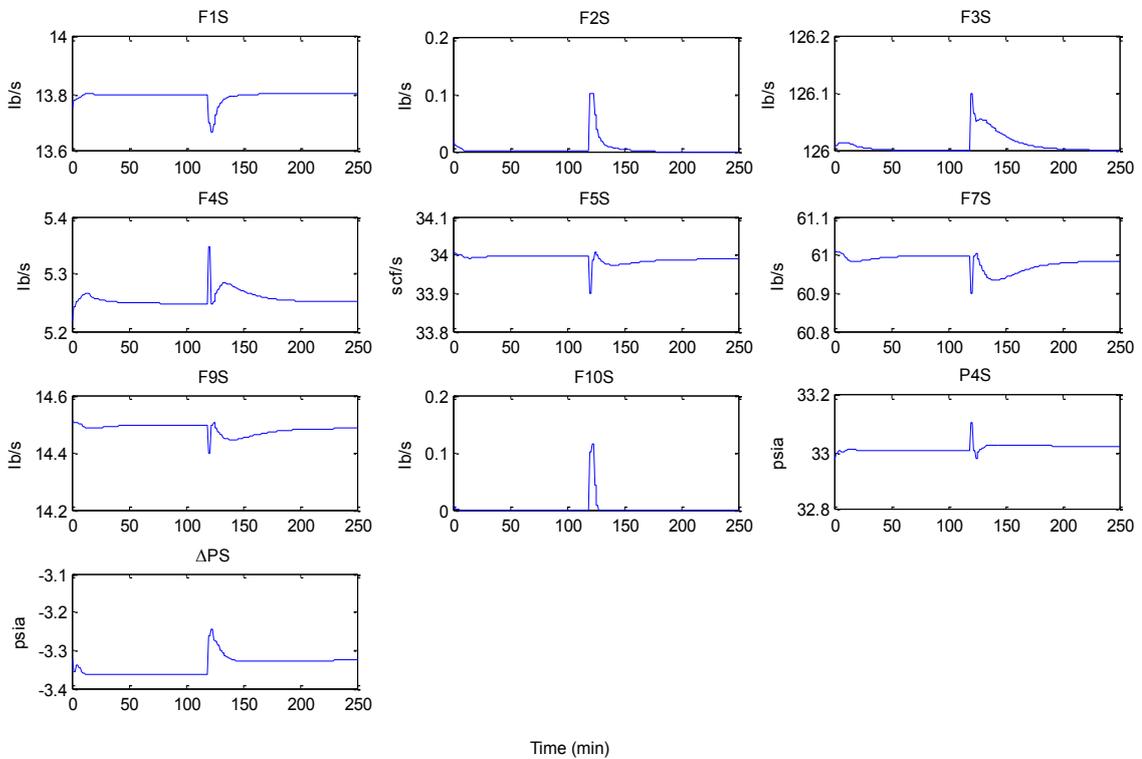


Figure 3.34: Trends of the manipulated variables (MPC outputs) under MPC

Chapter Four

Common MPC Faults and their Isolation

4.1 What-if Simulations of Common MPC Faults: An Overview

In the previous chapter, it was shown how MPCs might be implemented on certain nonlinear processes, to give good performance under normal operating conditions. More importantly it was shown that for very complex multivariable plants represented by high order nonlinear models, equivalent linear low order transfer function input-output models can be used as internal models for MPC implementation. These case studies are now carried forward into this chapter to provide vehicles on which to study how various causes of MPC performance degradation might be isolated. The approach is to induce symptoms into these simulations, and then, in each case, to ask the question “how might the operator detect and reason about what is observed?” The aim is to construct a set of rules and procedures, which might then form the basis of a general tool that the operator could use if these symptoms were ever to be observed on the operator’s real plant. It is assumed that the plant operator would have access to a simple linear model of the plant on which the MPC is implemented. By applying MPC similar to the one implemented on the ‘real plant’ (nonlinear model) to the linear model, the operator might be able to identify the source of a malfunction and possibly recover from the degradation.

Using the linear model, the operator could make changes to the MPC parameters, model configurations and operating conditions, based on informed guidance, and see what steps or actions might be taken to return the MPC to good performance. Or the operator might replicate the trend observed on the real plant to gain insight into what variables or

parameters are responsible for the degradation. The recovery procedure suggested could be as simple as simply retuning the MPC parameters, or as complex as outright redesign or reconfiguration of the MPC and/or of the plant. The tool helps to improve the confidence of the operator in his interaction with the MPC by improving his understanding of the processes involved.

State space MPC is used throughout in the what-if-simulations. MPC internal models are based on the transfer function matrices derived by applying steps directly to the plants (CSTR and Evaporator) and those obtained by applying steps on the linearized state space model of the plant (the FCCU). The transfer function models are used mainly for the transparency that they offer. They are actually converted to state space format before they are used in the state space MPC. Where the resulting state space dimension is large, an appropriate low order state space is estimated through the algorithm described in section 2.14.

Common situations that might give rise to poor controller performance, and in particular to MPC degradation, as reported by experiences from industrial applications (Darby and Nikolaou, 2012, Huang et al., 2000, Jelali, 2006, Jiang et al., 2011, Lee et al., 2004, Schäfer and Cinar, 2004), are summarised as follows:

- controller design (inadequate configuration, wrongly specified measured disturbances, improper controlled variable and manipulated variable selection);
- changes in operational conditions leading to changes in system dynamics (operation far from equilibrium points, excessive process drift due to plant aging, large and unmanageable changes in measured and unmeasured disturbances);
- tuning parameters specifications (output weights, prediction and control horizons, input and output constraints);

- equipment and hardware problems (regulatory controllers, valve stiction, deadband etc.

4.2 The Relative Weight Array

When analysing various situations, it was found that a simple, but new and novel measure, called the Relative Weight Array (RWA) might help the operator when thinking about MPC output weights. In selecting the output weights for set-point tracking MPC, it is usually stated that the relative sizes of these weights depend on “numerical values” of the variables involved. The steady gain array can be used to calculate a relative weight array, which can be used to obtain insight into the measure of these numerical values.

For a multivariable system with m outputs and q inputs, and with the steady gain matrix K (dimension $m \times q$), the RWA is a matrix of the same dimension as K . The elements of the RWA are obtained from:

$$RWA(i, j) = \frac{\max(K_a(:, j))}{K_a(i, j)}, i = 1, 2, \dots, m \quad j = 1, 2, \dots, q \quad \dots(4.1)$$

where

K_a is the absolute value of matrix K

$K_a(:, j)$ is all rows of column j of matrix K_a

Consider a system with steady state gain given below:

$$K = \begin{bmatrix} 1.5 & 0.5 \\ 2.0 & 1.7 \\ 3.4 & 2.9 \end{bmatrix}$$

the relative weight array is given as:

$$RWA = \begin{bmatrix} 2.27 & 5.80 \\ 1.70 & 1.71 \\ 1.00 & 1.00 \end{bmatrix}$$

The values in a column (an input) of the RWA give an indication of the relative numerical values or impact of that input on the associated outputs. The lowest value in a column (unity) points the output that is most affected by the associated input. The output with the highest RWA value for an input is the least affected by that input. When two or more outputs are to be set-point tracked, an indication of their baseline relative weights can be obtained by looking up the RWA values of the outputs against the associated manipulated inputs. For example if the first and the third outputs of the system with RWA given above are to be set-point tracked, the first output's weight can be from 2.27 to 5.80, while the second output's weight can be 1.00. These values are meant to be as starting guides only. The eventual output weights may be very different, as they might depend on other factors, including the order of importance of the output to be tracked. The RWA is used together with the relative gain array (RGA) to suggest acceptable baselines for output weights.

The case studies chosen for the what-if simulations are carefully selected to address most of the common situations, as listed in Section 4.1 above that might give rise to MPC degradation. They include cases involving MPC parameter tuning, MPC design, actuator and PID degradation, variable selection, MPC constraints and model plant mismatch. The method adopted in this research was to simulate these degradations based on MPC simulation in chapter 3, to describe the resulting degradation and to give a typical operator's perspective about the degradation observed. This was followed by an expert's reasoning about the degradation and proffering solution to recover from the degradation.

The following sections examine how the operators might seek to address situations that lead to MPC degradation and in doing so we draw out pertinent features that might form the basis for a maintenance tool design.

4.3 Case 1: Example Relating to MPC Parameter Tuning

This case study is based on the CSTR example, in which MPC is implemented as a supervisory controller (Section 3.3.2) with two PI regulatory controllers. The PI settings for the regulatory controllers are as given in table 3.5, and the MPC settings are the same as in table 3.8. The MPC internal model is that given in table 3.7. The case relates to a situation where the output weight of a control variable (C_A) is set much lower than the minimum that would result in good performance.

4.3.1 The Scene

In Section 3.3.2, the controlled variables were specified as V (volume of reactants in the reactor) and T (the reactor temperature). A situation is now envisioned where the operators feel that they need to change the configuration so as to closely monitor the product concentration C_A . In other words they use both V and C_A as controlled variables instead of V and T so as to better adhere to strict requirements on the product quality. They may not see any issues with this because C_A is measured already and is accounted for by the MPC model internal (similar to that given in figure 3.6, equation 3.8 and table 3.7). They also need to change the product concentration from 0.245 lb. mol A/ft³ to 0.2 lb. mol A/ft³.

Having made the change on the plant, the operators run the plant for 20 hrs before stepping down the set-point for the product concentration from 0.245 to 0.2 lb. mol A/ft³. Figure 4.1 shows trends obtained when the set-point was changed at the 20 hour mark. The figure shows that the variable C_A does not track its new set-point, while

variable V exhibits very small offset from the set-point. The operators have seen these trends, but they are not sure why they occur. Next the operators' possible perceptions of the degradation is discussed.

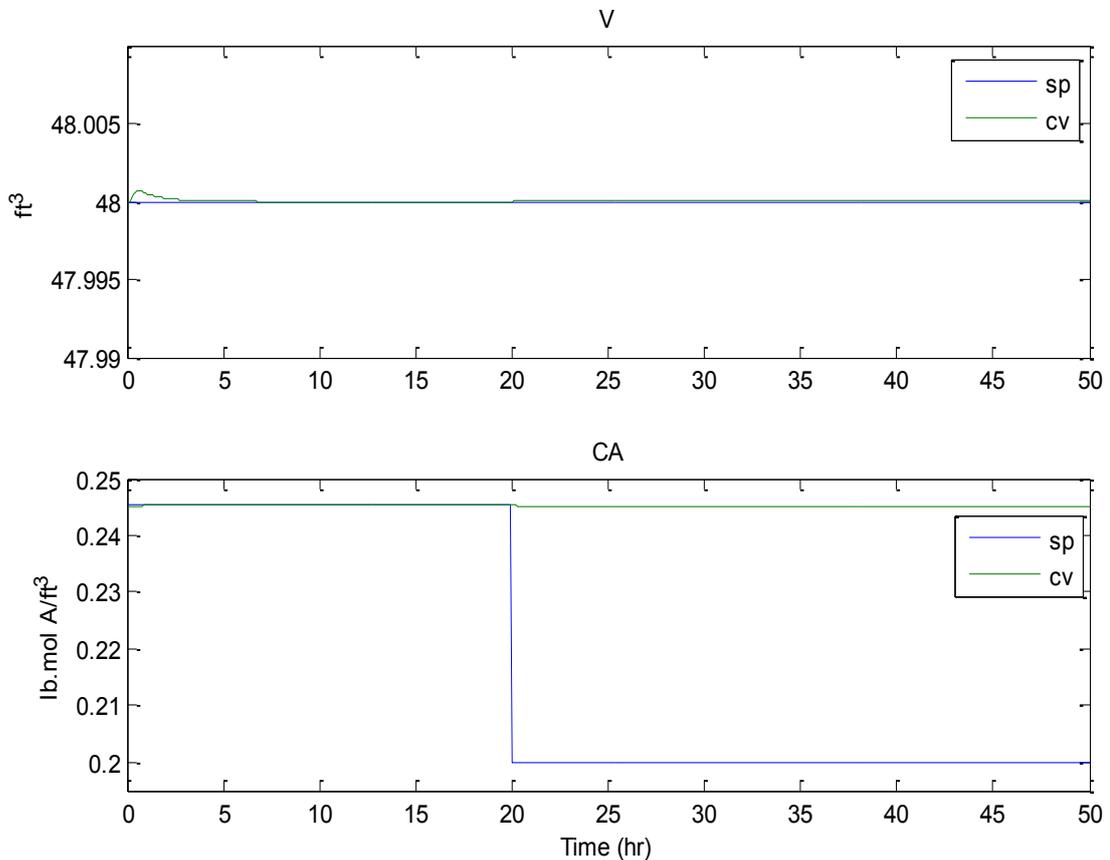


Figure 4.1: Controlled variables of the CSTR showing degradation

4.3.2 The operators' perspective

Having noted the CV trends, the operators are likely to study the trends of the manipulated variables: the MPC outputs, which are the set-points to the two PI controllers (figure 4.2), and the PI controllers' outputs which are the inputs to the plant (figure 4.3). The operators might observe that the manipulated variables exhibit steady values, without any oscillations or other unsteady trends. This observation should rule out causes associated with equipment degradation (e.g. valve sticking, deadband). They might also observe that the manipulated variables indicate no saturation or constraints activation, ruling out causes associated with input constraints.

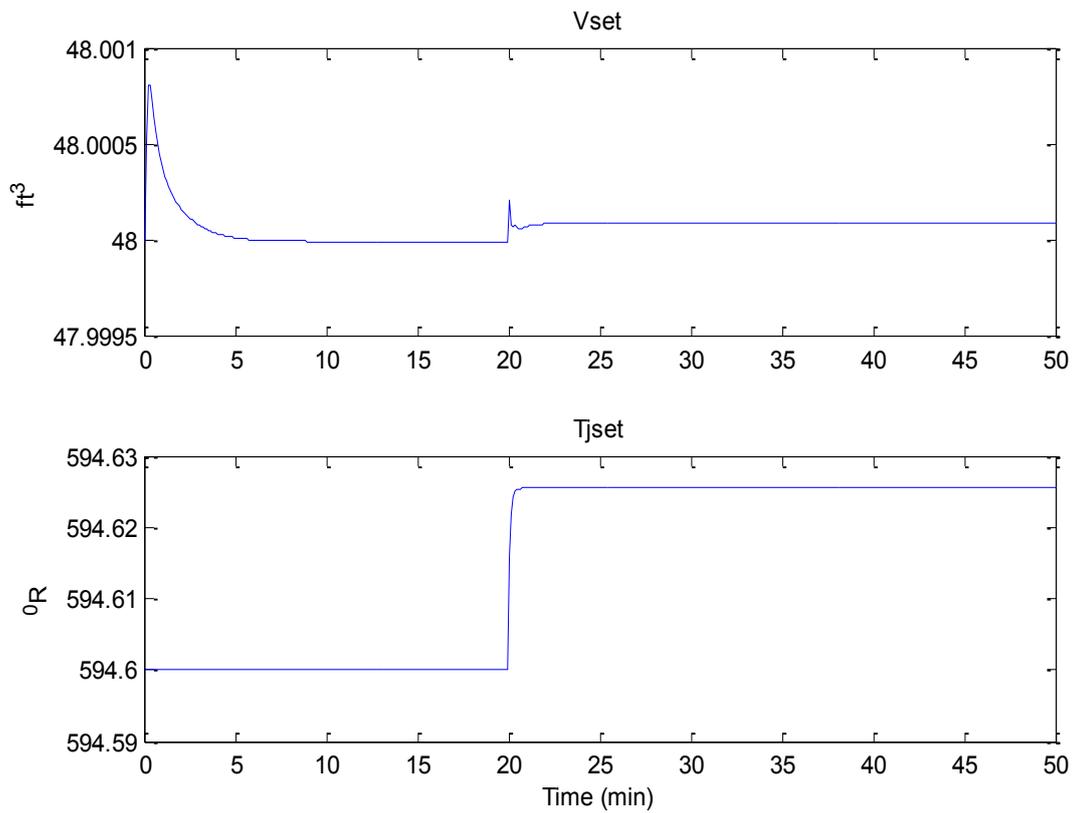


Figure 4.2: CSTR manipulated variables: MPC outputs

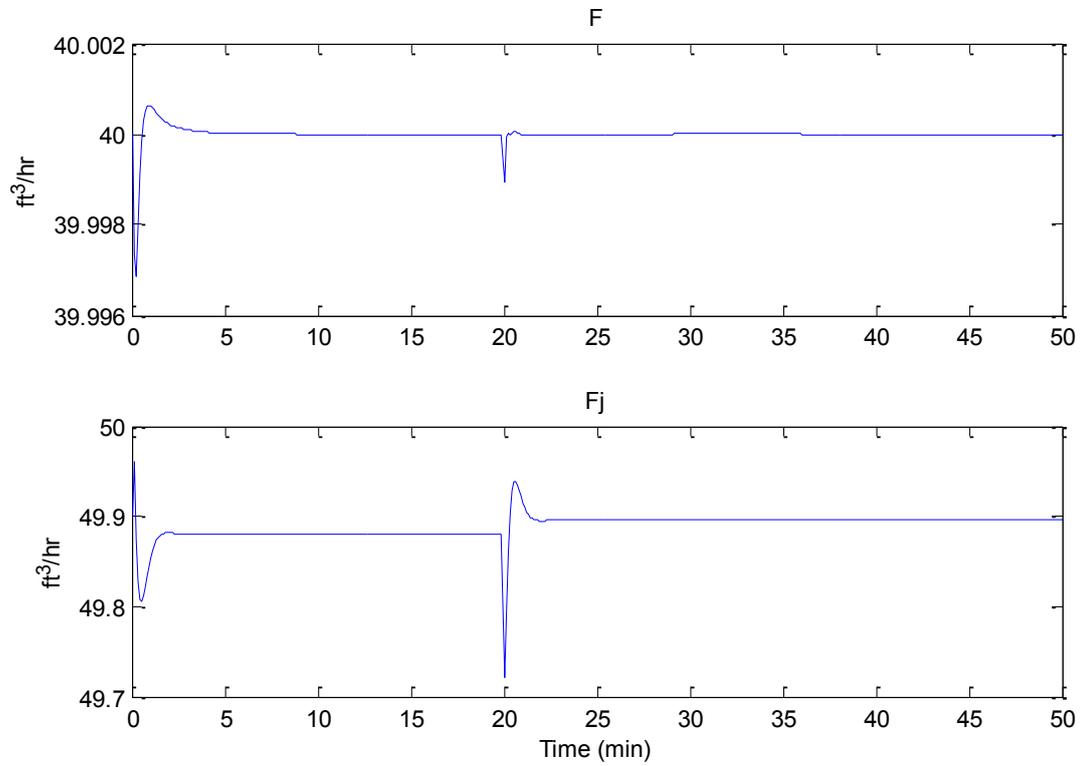


Figure 4.3: CSTR manipulated variables: PI outputs

The operators might now implement similar MPC on the linear model of the plant to create a virtual plant. In their simulations the trends of the controlled variables should be the same as those shown in figure 4.4 below. They might observe that the trends from the plant are very similar to those from their simulations. Observation like this should rule out causes due to process drift (contamination of feed streams, equipment aging or even a change in operating conditions). Linear MPC on linear models are not affected by changed operating condition except when constraints on the manipulated variables become active.

The operators might now narrow the possible cause(s) of the MPC degradation to either MPC design issues (overall configuration) or to poor parameter (input and output weights, variables selections etc.) specification.

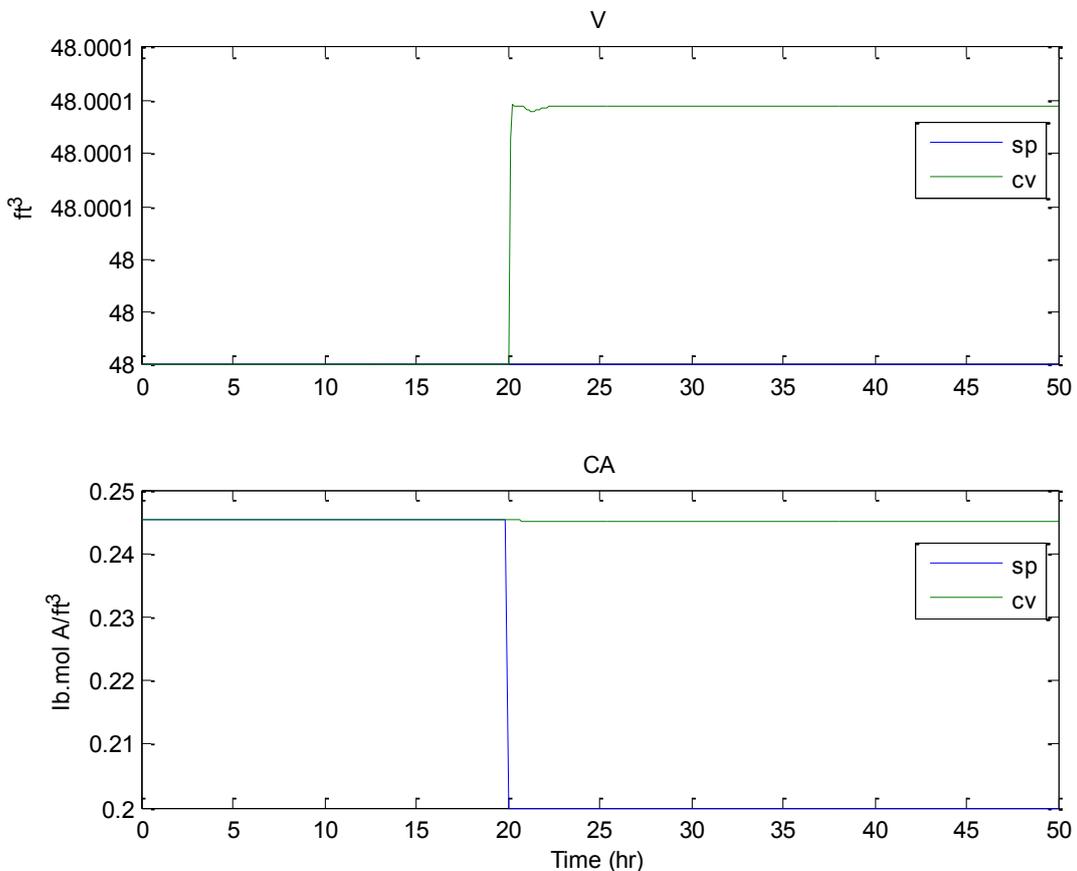


Figure 4.4 Controlled variables of the virtual plant showing degradation

4.3.3 Reasoning about the scene

The operators might now go back to see what worked successfully in the past. Thus they might look at the step response plots of the linear model shown in figure 3.6. They might observe similarities in the dynamics of the responses of three variables (the product concentration C_A , the product temperature T and the cooling water temperature T_j) when either a unit step is applied to the first manipulated variable V_{set} or to the second manipulated variable, T_{jset} . The fourth variable (the reactor volume V) has different dynamics. The trends of the variables CA , T , and T_j suggest that they are highly coupled, and that a change in one of the variables has an effect on the other two variables. Any one of these three variables should suffice as a control variable (but not two). The operators might now examine the relative gain array (RGA) of the system, obtained from the steady state gain of table 3.6, which is given below.

Table 4.1 Relative gain array of the CSTR

	V_{set}	T_{jset}
V	0.9972	0.0
C_A	4.8E-6	-1.5E-5
T	0.0028	0.5454
T_j	-3.9E-5	0.4546

In the RGA, the number in a cell gives an indication of the degree of effect the manipulated variable has over the corresponding controlled variable. A positive unity value indicates the greatest effect between the two. The RGA suggests that if V_{set} and T_{jset} are used as manipulated variables, V and T , or V and T_j may be used as controlled variables, because their RGA values are high. This simple approach does not completely rule out other pairings, such as V with C_A , however. They might now examine the relative weight array (RWA), which is a measure of the relative sizes of the steady state gains, shown in table 4.2.

Table 4.2 Relative weight array of the CSTR

	V_{set}	T_{jset}
V	1.0	inf
C_A	476.2	192.8
T	12.6	1.0
T_j	1000	1.1

RWA values give two indications: firstly a higher value suggests that the manipulated variable for that cell has a lower effect on the corresponding controlled variable. Secondly, the value gives an indication of what the output weight should be if the manipulated variable and the controlled variable are used in MPC design. Table 4.2 data suggests that if V is paired with T for example, the output weight can be as low as unity. On the other hand, if V and C_A are used as controlled variables, the weight of C_A should at least be as high as 193, possibly up to 476 or more.

The operators might now examine the effect of reversing the C_A output weight to 200, say, on the virtual plant (i.e MPC on linear model), all other MPC settings remaining as before. The operators should now observe satisfactory performance. If they were now to apply the same MPC setting (revised output weight) to the real plant, the outputs and inputs trends should be as shown in the following figures 4.5, 4.6 and 4.7. In figure 4.5, the two controlled variables now track their set-points. The manipulated variable from the MPC (the PI controller set points) shown in figure 4.6, and the plant inputs (PI controllers outputs) shown in figure 4.7 also exhibit satisfactory performance. The manipulated variables drive the outputs to their set-points without saturation.

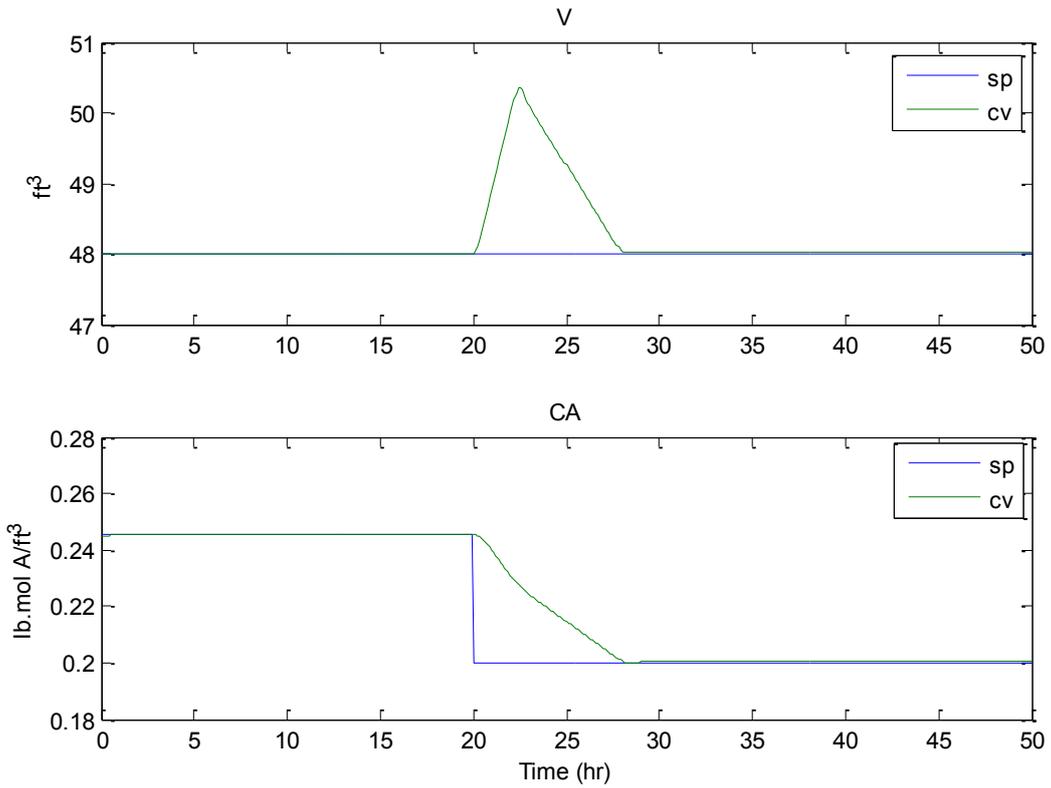


Figure 4.5 Controlled variables of the CSTR after recovery

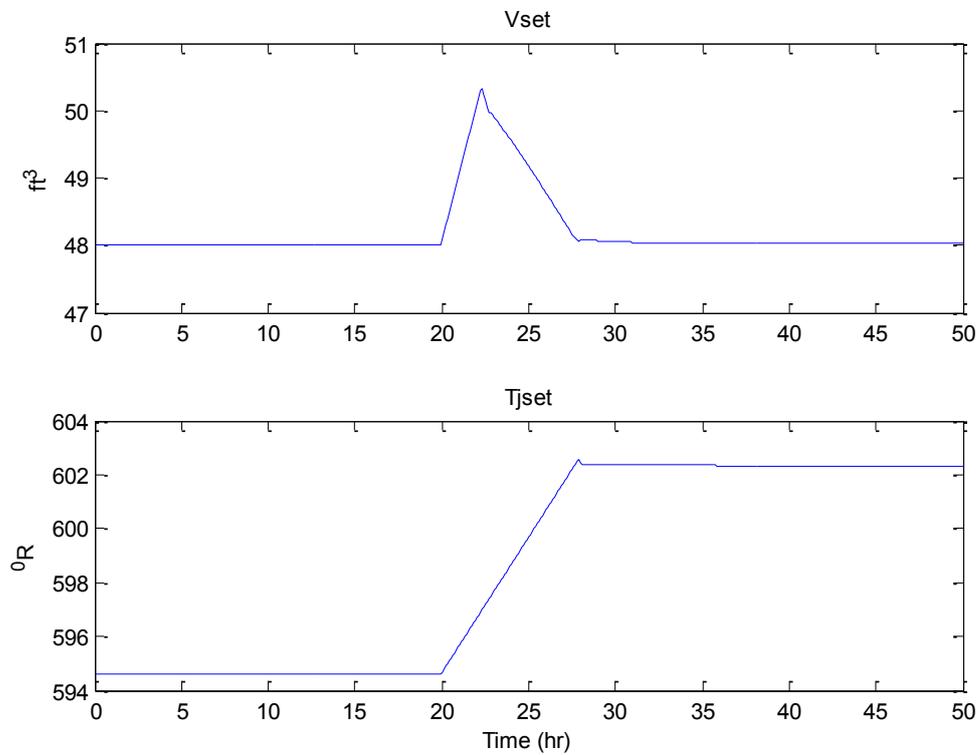


Figure 4.6: Manipulated variables of the CSTR (MPC outputs) after recovery

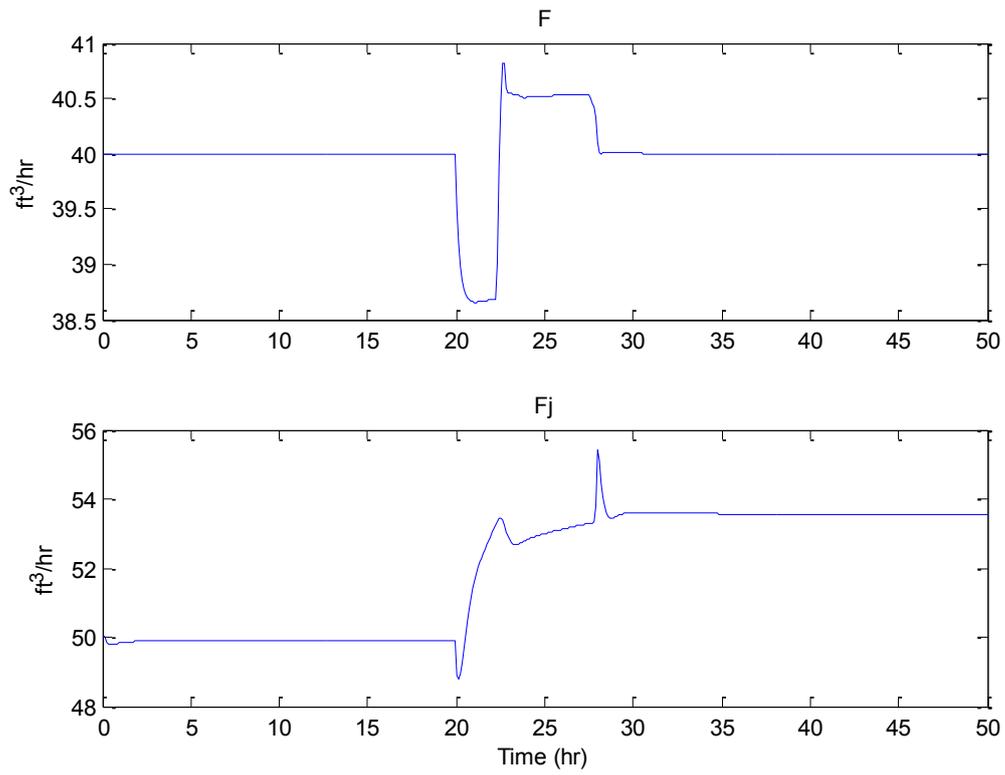


Figure 4.7 Manipulated variables of the CSTR (PI outputs) after recovery

4.3.4. Case 1 Conclusion

The following table summarises the path described above to reach the wrong output weight conclusion.

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	NO	
	Partial set-points tracking?	NO	
	Zero set-points tracking?	YES	Tuning parameters Unmeasured disturbance Process drift/model-plant mismatch Variable selection
	Unbounded outputs?	NO	
Observed MV Trends	Oscillations?	NO	
	MV saturates?	NO	
	Aggressive MV behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CV selections?	YES	Good variable selection
Examine CV and MV selection using RGA	Reasonable for given MV?	YES	Good variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	NO	Wrong output weights suspected.
Similar MPC on virtual Plant	MV and CV Trends similar to those of real plant?	YES	No process drift suspected
	Tuning improved performance?	YES	Wrong output weights confirmed
Result: Retuning of MPC outputs weight restored MPC performance			

4.4 Case 2: Example relating to MPC design

This case study is based on the evaporator example, in which MPC is implemented directly on the process (Section 3.4.2). The MPC settings are the same as in table 3.15.

The case relates to a situation where the plant is operated very far from its nominal operating point, leading to a change in dynamics, and therefore, degradation. The goal is to see how a different MPC design might prevent the degradation.

4.4.1 The Scene

In Section 3.4.1, direct linear MPC was used for set-point control of the three output variables L2, P2 and X2. Under this operation within the vicinity of the equilibrium point, MPC performance was satisfactory. All controlled variables tracked their set-points very reasonably. A situation is now envisioned where, due to operational requirements, the operators have to run the plant at a point far from this equilibrium point. In this imagined scenario, the operators ramp up the pressure P2 from 50.5 kPa to 70 kPa in 20 minutes, starting at the 20 minute mark. Simultaneously from the same point and within the same time frame, they ramp down the product concentration X2 from 25% to 15%. They keep the MPC tuning parameters as specified in table 3.14.

Figures 4.8 and figure 4.9 below show the plots of the output and input trends they observe. Of the three controlled variables, only P2 tracks its new set point, after the ramp (figure 4.8). The separator level L2 tracks its set point about halfway through the ramp, but deviates drastically from the set point thereafter, with a large steady value offset in the new operation condition. The product concentration X2 does not track the set-point during the ramp, and ends up with a small steady value offset in the new operating condition. In figure 4.9, all the manipulated variables (F2, F200 and P100) attain steady values different from the initial equilibrium values. The operators have seen these trends. What can they do?

4.4.2 The operator perspective

The operators might observe that the manipulated inputs (figure 4.9) do not exhibit any aggressive behaviour, which rules out poor input weight specification. There are also no input saturations or oscillations, ruling out causes associated with constraints and equipment degradation. Also there is no recorded change in the trend of the measured disturbance.

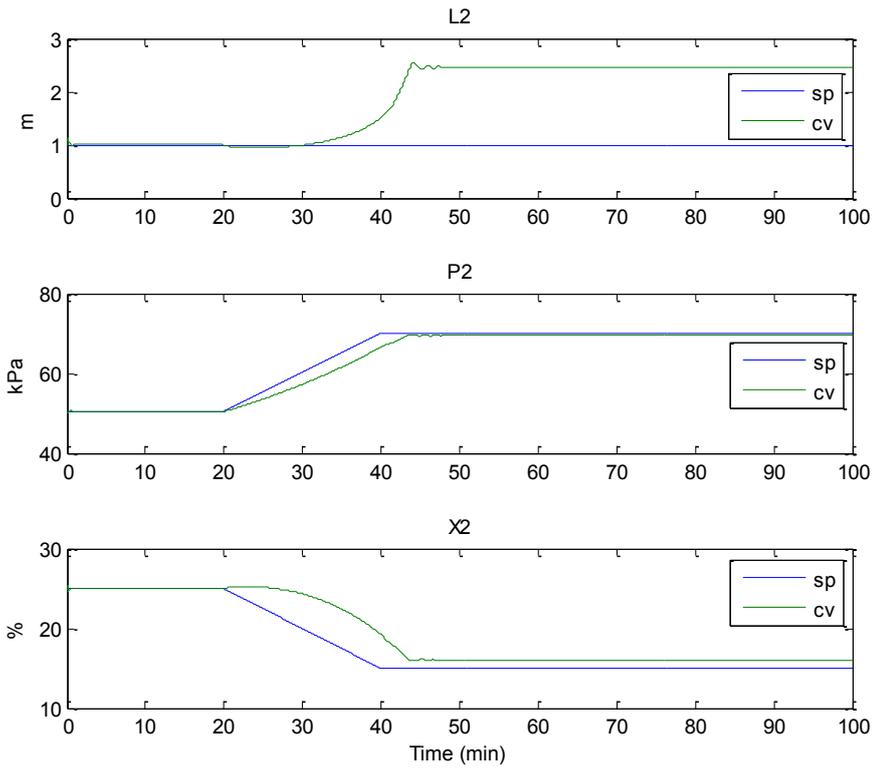


Figure 4.8: Controlled variables of the evaporator showing degradation (sp = set pint. cv = controlled variable)

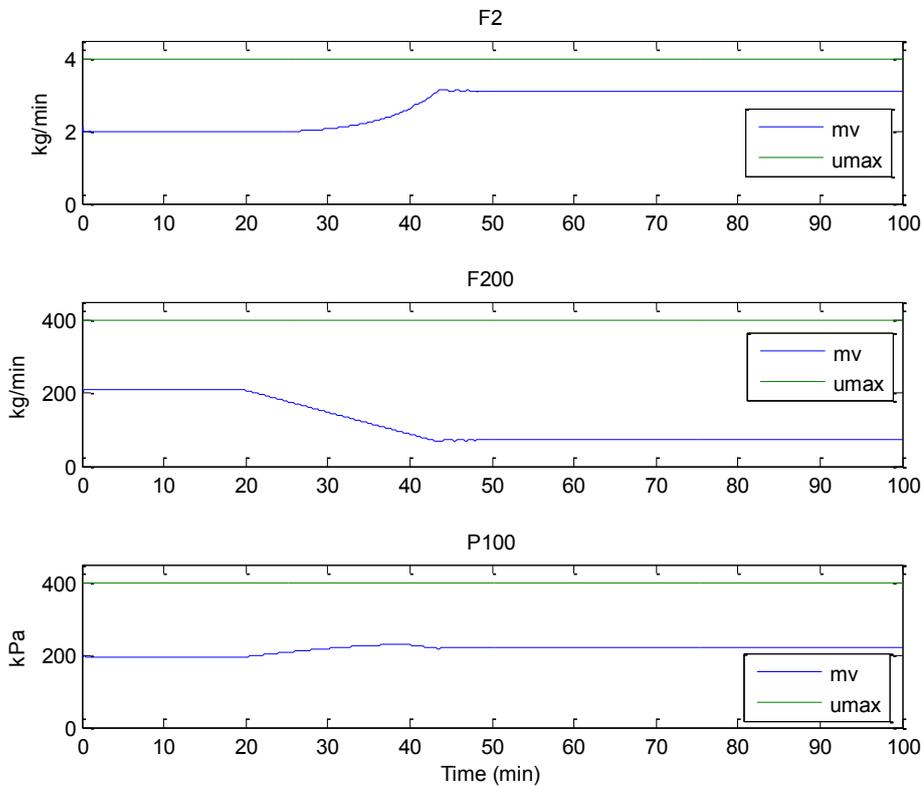


Figure 4.9: Manipulated variables of the evaporator (mv = manipulated variable. umax = upper constrain value of manipulated input)

The operator might make reference to the RWA of the linear model (table 4.3) and observe that the MPC output weight vector ([100 10 10]) is reasonable, ruling out poor output weight specification.

Table 4.3: RWA for open loop evaporator

	F2	F200	P100
L2	134.0	111.0	281
P2	2.7	1.0	1
X2	1.0	inf	inf

4.4.3 Reasoning about the scene

The operators might now implement similar MPC on the linear model of the plant to obtain a virtual plant. They would find that they could not change the MPC parameters to make their linear model exhibit good performance. For instance they might obtain trends like those shown in figures 4.10 and 4.11.

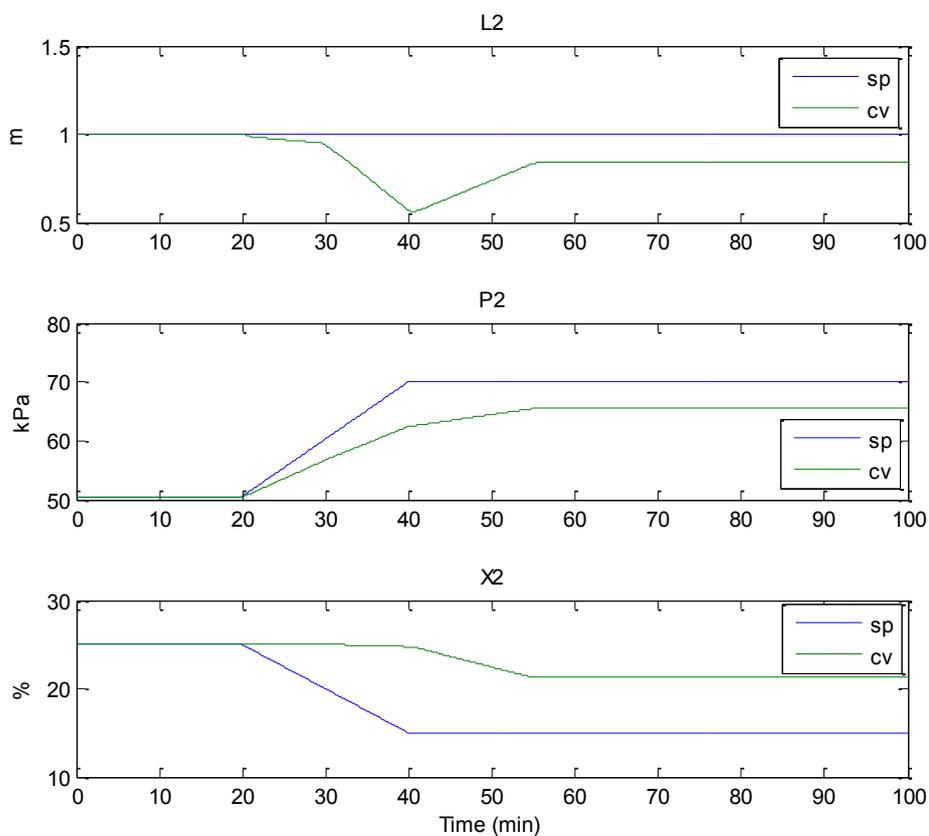


Figure 4.10 Controlled variables of the virtual plant showing degradation (sp = set pint. cv = controlled variable)

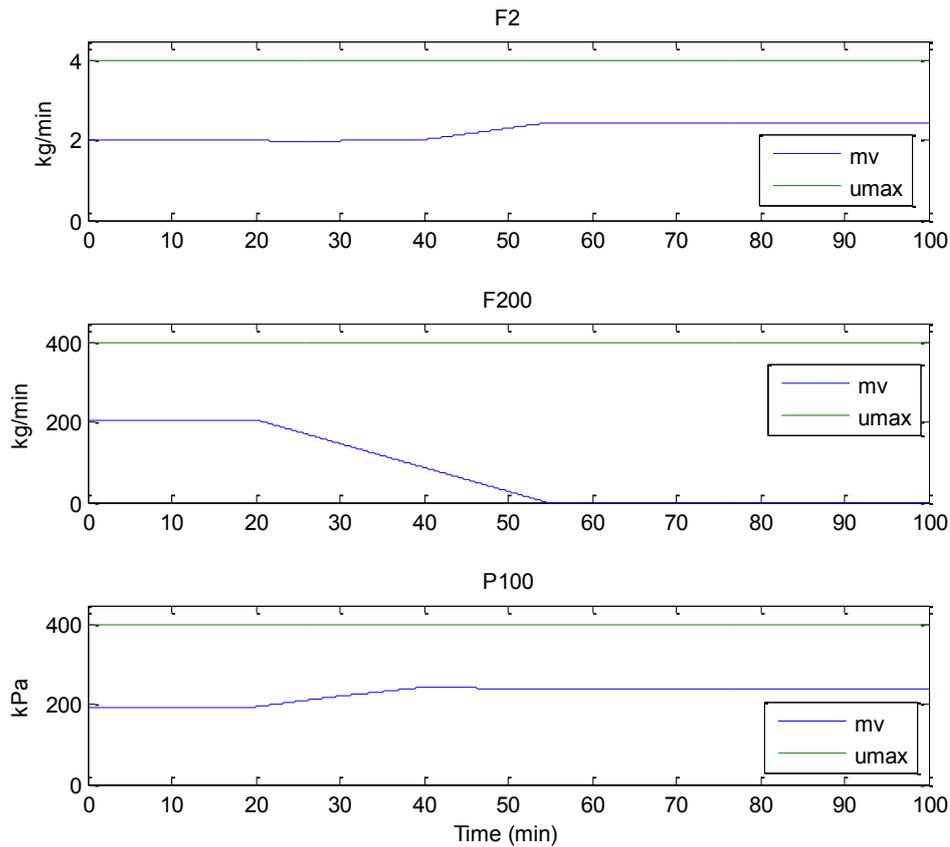


Figure 4.11 Manipulated variables of the virtual plant (mv = manipulated variable. umax = upper constrain value of manipulated input)

However by looking at the trends, the operators should be able to see that manipulated variable F200 continues to decrease even after the ramp has been completed, saturating at about the 55 minute mark, when it has reached its lowest physical value (0). The saturation of F200 seems to be preventing the controlled variables from reaching their intended set-point values. This suggests that F200 has a very significant effect on the controlled variables especially at the new operating point. The trends also suggest that the operating point for P2 and X2 should not exceed 65.5 kPa and 21.25% respectively, even for the virtual plant. The nonlinear plant may even require a lower operating point.

Once again the operators might wish to refer to what worked in the past, by studying the step response plots of figure 3.15 closely. The operators might observe that step

changes in all the manipulated variables lead to a ramp-like response in L2. With this situation, it might be difficult to control L2 directly. On the other hand, all the manipulated variables moved to suitable equilibrium values without constraints. This might not be possible with the change of operating point, because of physical limitations placed on all the manipulated variables, especially F200. Changes required by P2 and X2 in F2, F200 and P100 for offset free tracking all also affect L2 simultaneously. Also the integrating loop F2-L2 is the fastest loop, compared to the other integrating loops F200-L2 and P100-L2, because it has the highest slope (-0.062) compared to the others (-0.004 and -0.0063 respectively). The operator might also note that the relative gain array shown in table 4.4 indicates that influence F2→X2 is stronger than influence F2→L2. That means that under direct MPC, L2 would always struggle for control because X2 would always be of greater importance to F2. A change in output weight may help change this situation slightly. The operator might conclude that based on these observations, the current MPC design or configuration is not ideal for implementation at this new operating point. They now believe they have reached the point where expert intervention is required.

Table 4.4 RGA for open loop evaporator

	F2	F200	P100
L2	-0.0000	0.7166	0.2834
P2	-0.0000	0.2834	0.7166
X2	1.0000	0.0000	0.0000

4.4.4 Possible outcomes: experts perspective

One possible outcome for this particular plant is to implement adaptive MPC. This is a continuous re-linearization of the linear internal model at specific points. The approach is one of the methods suggested by Maciejowski (2002). This approach is unlikely to be implemented in general on real plants.

An alternative approach is to implement the MPC as a supervisory controller. The strong integrating loop F2-L2 is closed with the aid of a proportional plus integral (PI) controller. Then MPC is implemented as supervisory control. The procedure for doing this has been described in section 3.4.3. Closing the loop F2-L2 would ensure that the MPC is concerned with manipulating the set-point to the PI controller that in turn manipulates the flow rate F2 to the desired value. In essence closing the loop F2-L2 ensures that the plant could be operated over a range of operating points. The step response plots for the closed loop plant of figure 3.19 as well as the relative gain array (RGA) for the closed plant shown in table 4.5 show that the set-point to the PI controller is now virtually solely responsible for controlling L2. F200 and P100 now have very little impact on L2. Both F200 and P100 are now virtually devoted primarily to controlling P2 and X2. The conflicting control situation described earlier is no longer present.

Table 4.5 RGA for closed loop evaporator

	L2set	F200	P100
L2	1.000	0.00	0.00
P2	-0.0000	0.4608	0.5392
X2	0.0000	0.5392	0.4608

Plant simulation results for this supervisory controller are shown in figures 4.12, 4.13 and 4.14 below. Plots umax refer to upper constraint trends of the inputs. All the controlled variables attain their settling time about 5 minutes after the ramps on the set-points of P2 and X2 are completed.

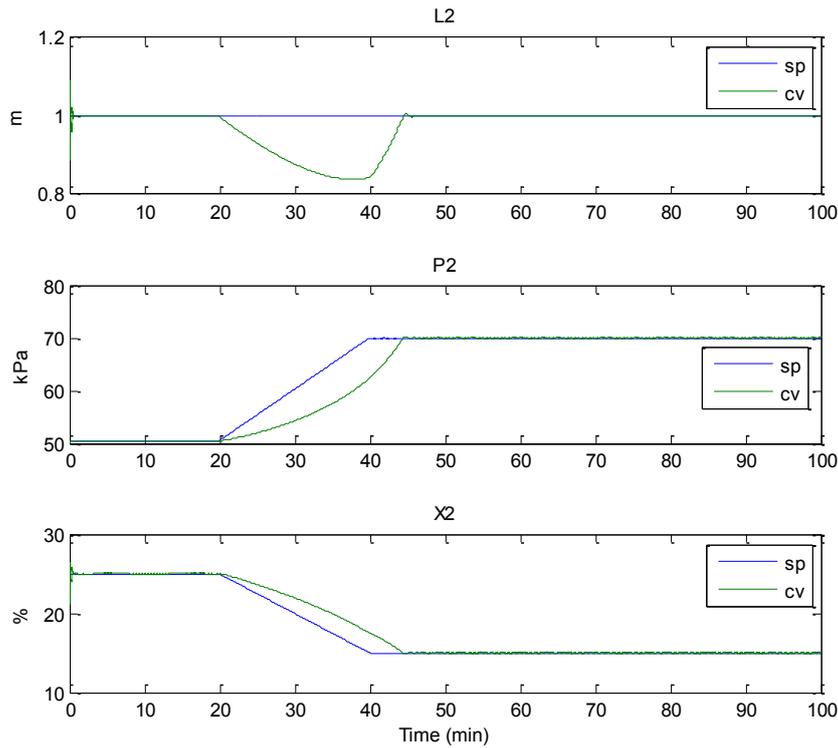


Figure 4.12: Closed loop evaporator showing improved MPC performance (sp = set pint. cv = controlled variable)

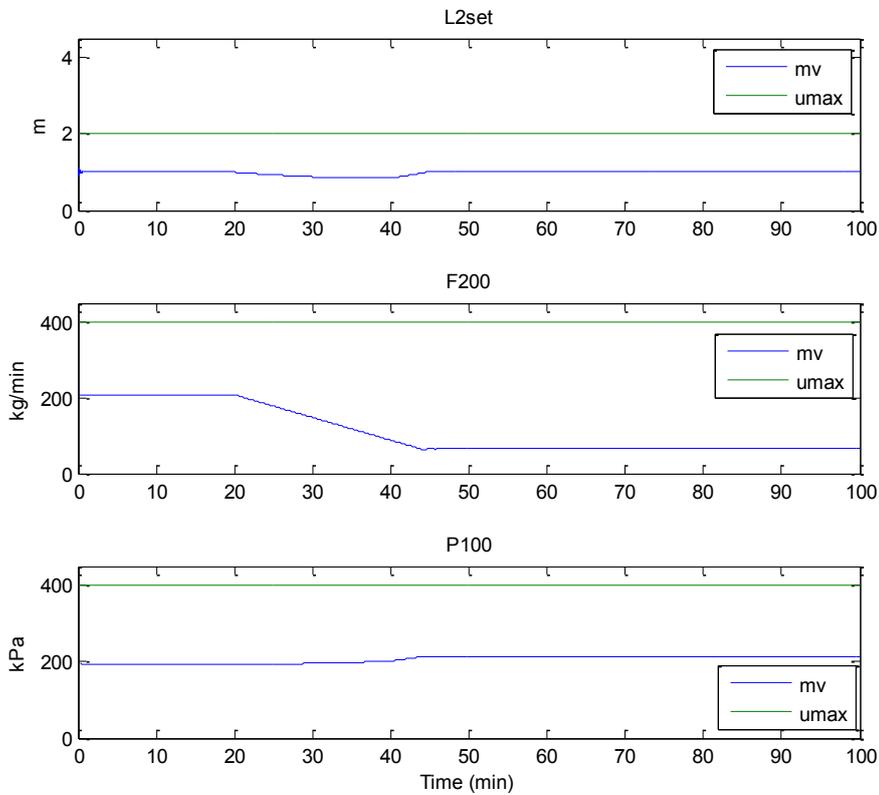


Figure 4.13: Controlled variables of evaporator showing improved MPC performance in closed loop (mv = manipulated variable. umax = upper constrain value of manipulated input)

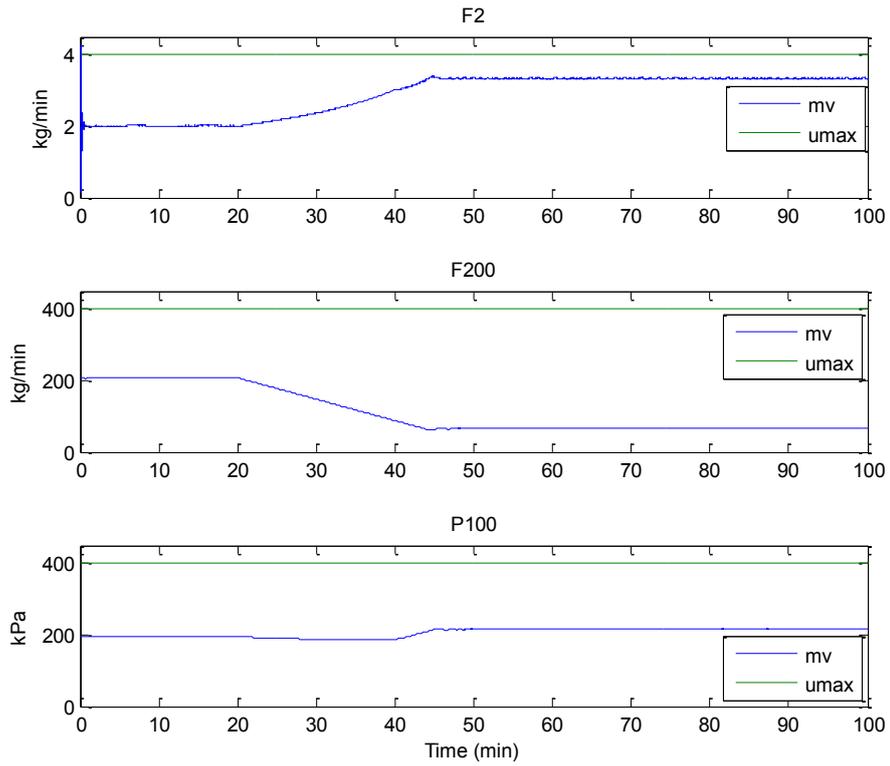


Figure 4.14 Manipulated variables of evaporator showing improved MPC performance in closed loop (mv = manipulated variable. umax = upper constrain value of manipulated input)

4.4.5 Case 2 Conclusion

The following table summarises the steps taken to reach the regulatory control conclusion.

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	NO	
	Partial set-points tracking?	YES	<ul style="list-style-type: none"> • Unmeasured disturbance • Process drift/model-plant mismatch
	Zero set-points tracking?	NO	
	Unbounded outputs?	NO	
Observed MV Trends	Oscillations?	NO	
	MV saturates?	NO	
	Aggressive MV behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CV selections?	YES	<ul style="list-style-type: none"> • Good variable selection • Presence of unclosed integrating loops noted
Examine CV and MV selection using RGA	Reasonable for given MV?	YES	Good variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	YES	Good output weights
Similar MPC on virtual Plant	MV and CV trends similar to those of real plant?	YES	<ul style="list-style-type: none"> • The new operating point not suitable for the model • Too large process drift • Design issue suspected • A configuration that permits wider range of operating points is considered
	MPC Tunings improved performance?	NO	
	Reversion to pre-ramp set-point improved performance?	YES	
Result: Re-design of the MPC allowed for operations from the old operating point to the new one			

4.5 Case 3: Example relating to sensor/actuator degradation

This case study is based on the direct implementation of MPC on the evaporator as described in section 3.4.2. The MPC settings are the same as in table 3.15. The case relates to a situation where there is equipment degradation in the form of an increased deadband in one of the actuators. This leads to degradation of the MPC.

4.5.1 The Scene

In Section 3.4.2, the evaporator process is operated at the equilibrium point by the MPC. A situation is now envisaged where the operators begin to observe performance degradation of the MPC from the 20 minute mark. The controlled variables begin to oscillate about their set-points as shown in the trends of figure 4.15 below. The oscillations are sustained, periodic and have definite patterns.

4.5.2 The operators Perspective

The operators might immediately examine the trends of the MPC outputs, which are the manipulated inputs to the plant. The trends, as shown in figure 4.16, exhibit patterns of oscillations in the three manipulated variables, though the oscillations appear more pronounced in F2. Oscillations are usually associated with equipment degradation or bad control tuning. Because the constraints on the manipulated variables have not at any point become active, the operators might reason that it is better to investigate a possible equipment degradation first before delving into MPC.

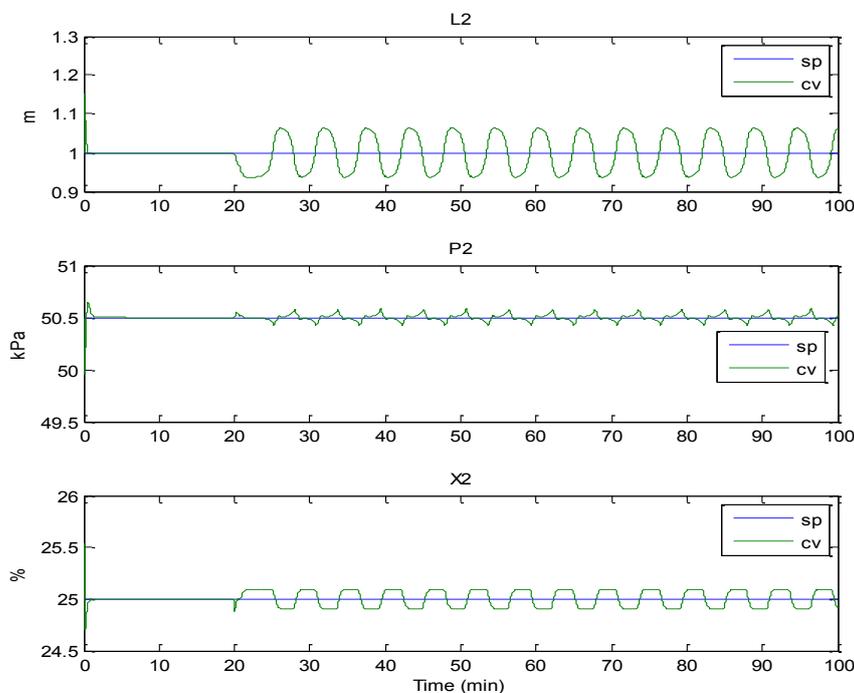


Figure 4.15: Controlled variables of the evaporator showing oscillations (sp = set pint. cv = controlled variable)

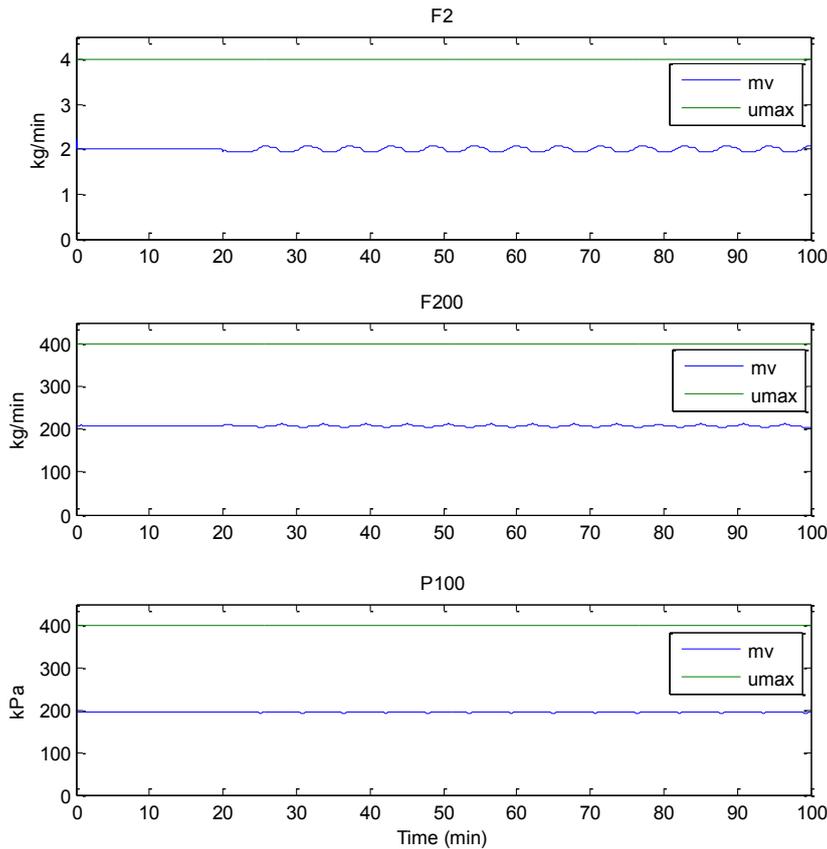


Figure 4.16 Manipulated variables of the evaporator showing oscillations (mv = manipulated variable. umax = upper constrain value of manipulated input)

In their paper Tatjewski (2010) state that plant oscillations are mostly as a result of nonlinearities in plants such as valve stiction, deadband, deadzone etc. They can also be due to poor process and control design, poor controller tuning, and oscillatory disturbances (Shoukat Choudhury et al., 2005).

Most plant control systems would record the inputs to each of the three valves (op) and the outputs from the valves (mv). The operators might produce limit cycle diagnostic plots from these records as shown in the figures below. In figures 4.17, 4.18 and 4.19, the upper parts are the plots of both op and mv against time, while the lower parts show plots of mv against op.

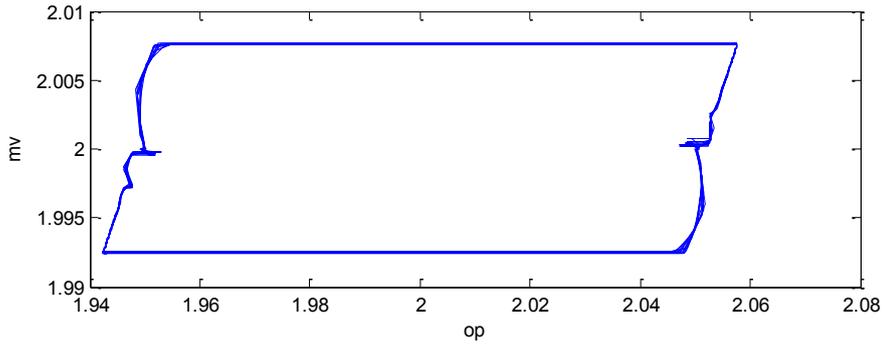
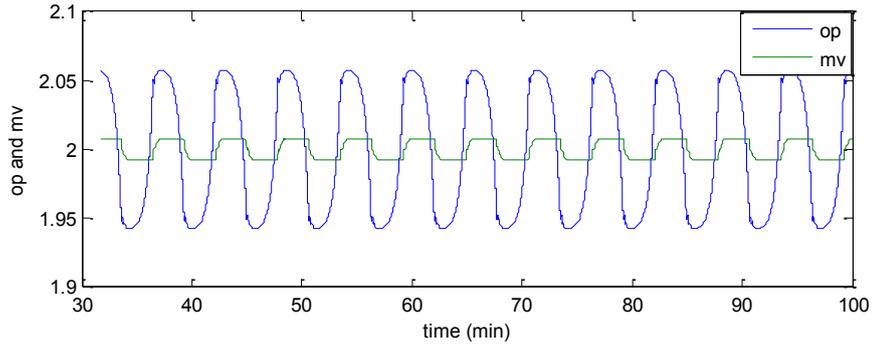


Figure 4.17: Limit cycle plots for valve on F2

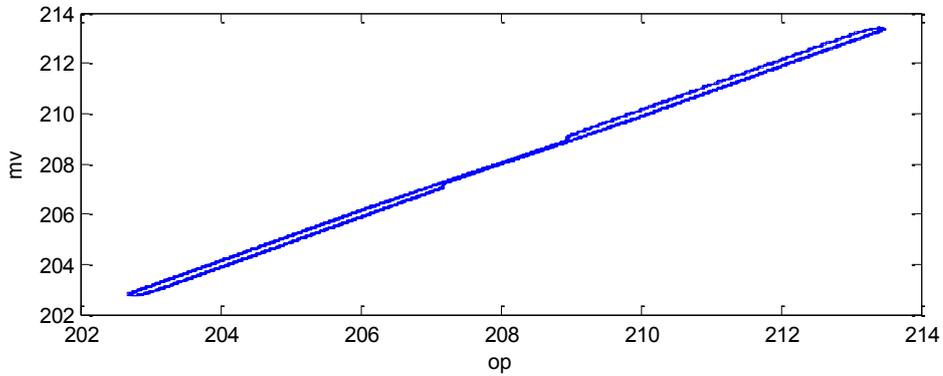
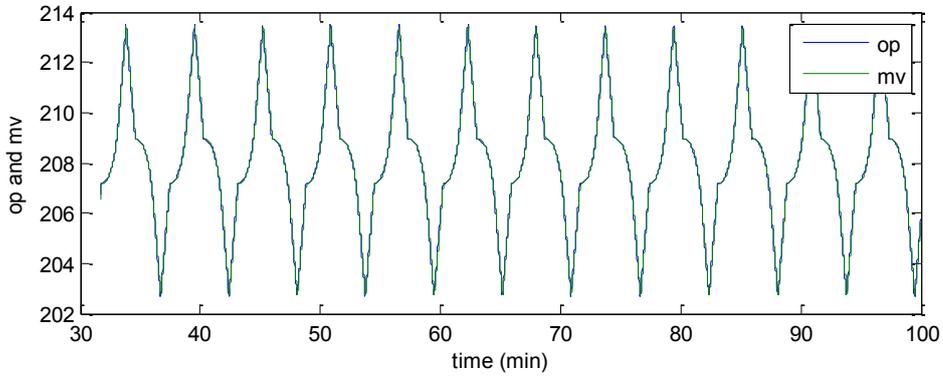


Figure 4.18: Limit cycle plots for valve on F200

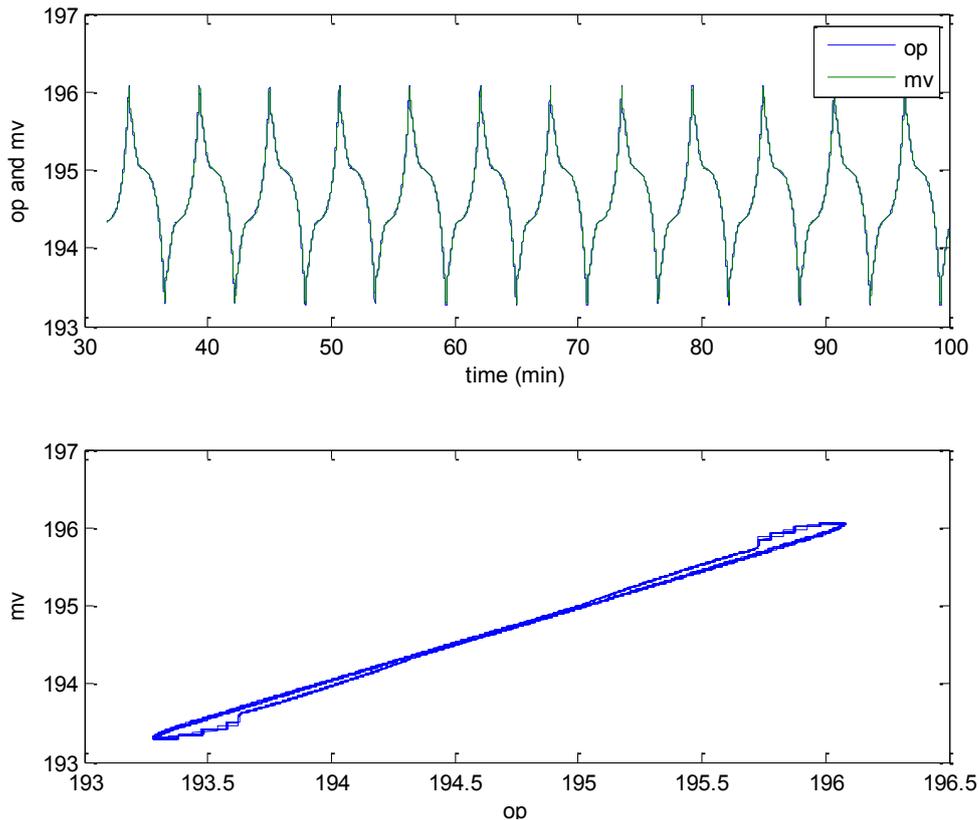


Figure 4.19: Limit cycle plots for valve on P100

4.5.3 Reasoning about the scene

The developed gear tooth-like wave of the **mv** in figure 4.17 suggests that the valve manipulating F2 is most probably affected by a nonlinearity caused by pure deadband (Patwardhan et al., 2002, Tatjewski, 2010). The parallelogram-like shape of the plot of **mv** against **op** shown in the same figure confirms that it is indeed a case of pure deadband. The fact that **op** plots coincide with **mv** plots in figure 4.18 and 4.19 suggest these valves are performing in a linear fashion, i.e. the valves on F200 and P100 are almost linear. The pattern observed in the **pv** plots of figure 4.15 is as a result of propagation of the oscillations of one variable into another.

Having isolated the source of the MPC degradation, the affected valve can be repaired or replaced and the performance of the MPC can be restored.

4.5.4 Case 3 Conclusion

The path taken to arrive at equipment degradation is summarised in the table below.

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	YES	Equipment failure PID degradation
	Partial set-points tracking?		
	Zero set-points tracking?		
	Unbounded outputs?	NO	
Observed MV Trends	Oscillations?	YES	Equipment failure PID degradation
	MV saturates?	NO	
	Aggressive MV behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CV selections?	YES	Good variable selection
Examine CV and MV selections using RGA	Reasonable for given MV?	YES	Good variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	YES	Good output weights
Similar MPC on virtual Plant (virtual plant does not include effect of PID or equipment)	MV and CV trends similar to those of real plant?	NO	<ul style="list-style-type: none"> • Good MPC performance • Investigate equipment failure
	MPC Tunings improved performance?		
Run the plant in open loop (No MPC)	Performance improved	NO	Valve failure suspected
Plot and study the valve limit cycles	Fit any known pattern of failure?	YES	
Result: Limit cycle plot reveal deadband problem in one of the valves			

4.6 Case 4: Example relating to MPC constraints

This case study is based on the evaporator example where MPC is implemented directly (described in Section 3.4.2) The MPC settings are the same as in table 3.15. The case relates to a situation where a hard constraint on one of the manipulated variables (F200) is set too tightly, leading to degradation of the MPC.

4.6.1 The Scene

In Section 3.4.2, the evaporator is operated at the equilibrium point specified in section 3.4. The equilibrium value of the product flowrate, F200, is 208 kg/min. A different scenario is now imagined, where as a result of the original plant specification, it was decided that the valve on this flow line should not be allowed to fluctuate more than 50% above or below the equilibrium opening. That is the hard constraints on F200 was set as 208 ± 104 kg/min. All other MPC settings were as in section 3.4.2. The plant operated sensibly for some time. Then one day they observe changes like those shown in figure 4.20. The MPC begins to degrade at the 50 minute mark. Two of the controlled variables (P2 and X2) deviate sharply from their set-points, while P2 exhibits a large overshoot and X2 shows large undershoot, with both ending up with large steady value offsets.

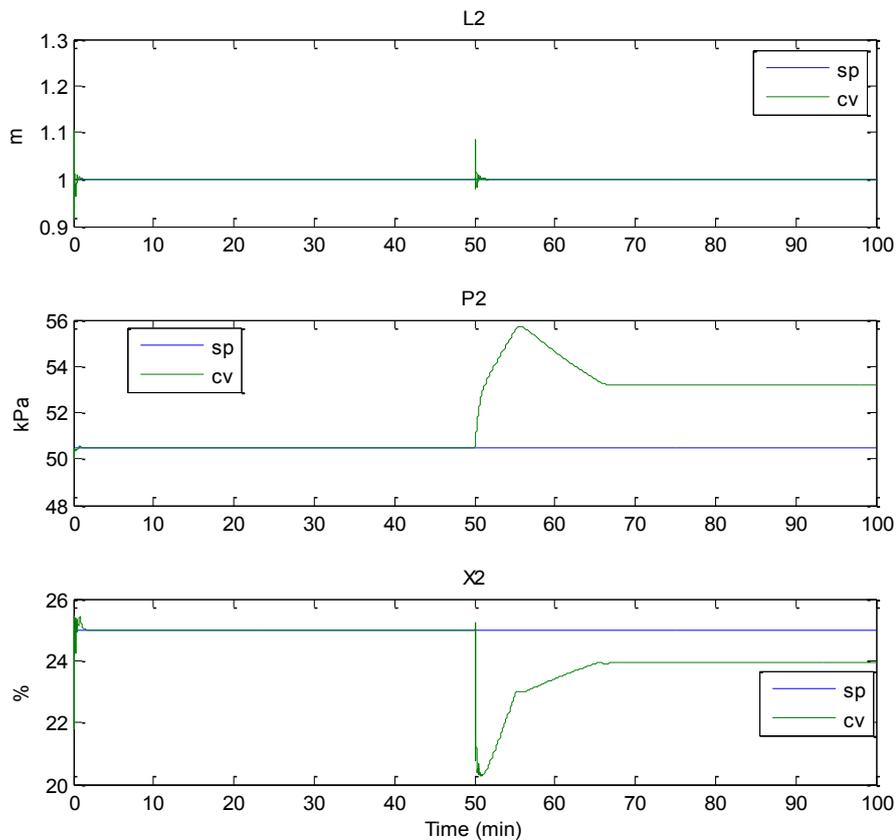


Figure 4.20: Trends of evaporator controlled variables showing degradation (sp = setpoint. cv = controlled variable)

4.6.2 The operators Perspective

The operators might be surprised at the sudden degradation of the MPC, because the parameters of the MPC had not been changed prior to the sudden appearance of degradation. For them, issues relating to tuning parameters could be eliminated initially at least. They might observe that the trends of the controlled variables do not show any oscillations so they might also eliminate issues relating to equipment degradation. The operators might now study the trends of the manipulated shown in figure 4.21 below.

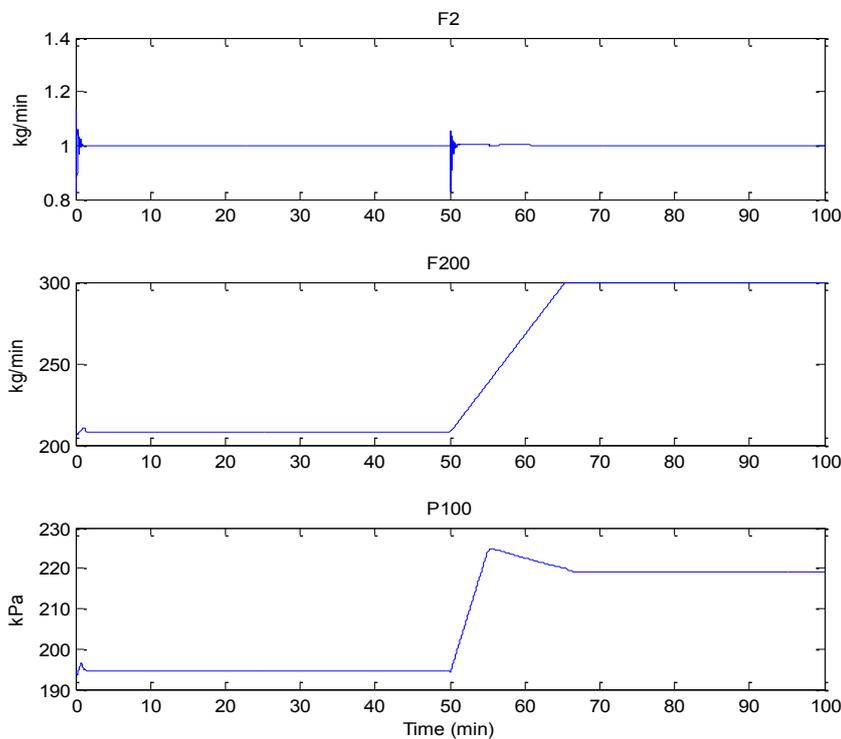


Figure 4.21: Trends of evaporator manipulated variables

4.6.3 Reasoning about the scene

From figure 4.21, the operators might observe that at the 50 minute mark when the controlled variables P2 and X2 begin to deviate from their set-points, both F200 and P100 move to bring them back to their set-points. At about the 65 minute point, F200 saturates and P100 attains a steady value of about 218 kPa. In essence the constraint on F200 has become active at about the 65 minute mark. They might suspect that the source of the MPC degradation is due to the constraint on F200. Ideally, equipment can

support F200 flow up to 400 kg/min and it is assumed that the operators are aware of this. Based on this observation, the operator might deduce that the constraint on F200 is too tight, and decide to relax it. The upper bound of the constraint on F200 was made 400 kg/min. The trends of the MPC implementation on the plant after the constraint is relaxed is shown in figure 4.22. Now all the controlled variables tracked their set-points.

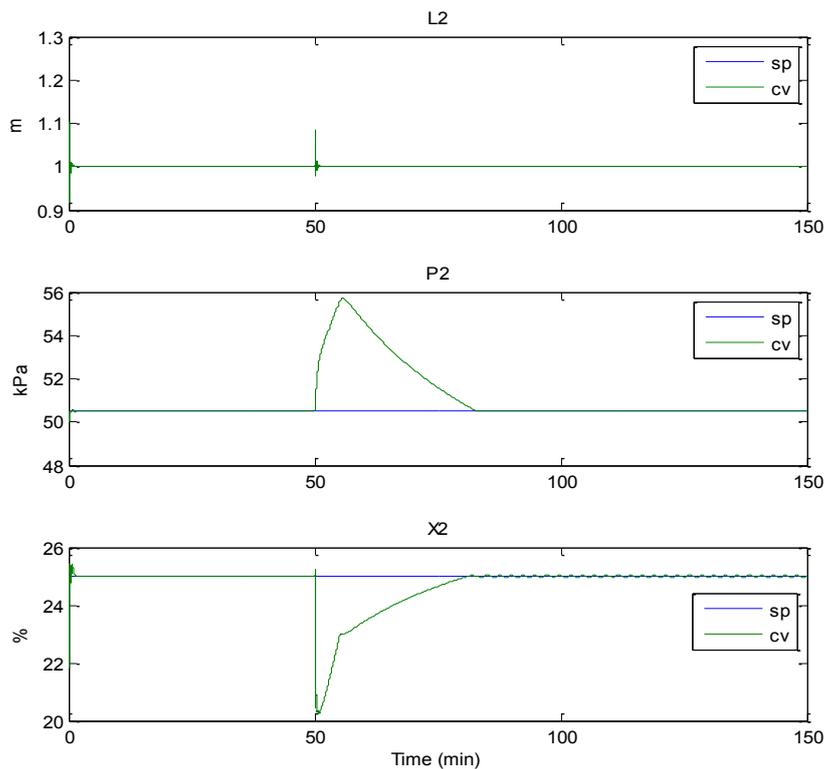


Figure 4.22: Trends of evaporator manipulated variables after constraint is relaxed

To the operators, the relaxation of the constraint on F200 has solved one problem: that is the elimination of offsets on P2 and X2. But it has not revealed the source of the sudden breakdown of the MPC and the need to adjust the constraint on F200. It is reasonable to assume that all MPC tuning parameters cannot all suddenly require retuning. The operators may arrive at two possibilities for the sudden degradation of the MPC:

- process drift due to equipment or feed fouling, and

- sudden appearance of unmeasured disturbance

The linear model available to the operators is assumed to include models for unmeasured disturbances F1, T1, T200 and X1. The operators might choose to experiment with these measured disturbances which are not included in the controller by examining their effects on the virtual plant. They might observe that trends similar to those that occurred on the real plant can be replicated when the unmeasured disturbance F1 is increased from 10.0 kg/min to 11.0 kg/min, with the constraints on F200 kept at 208 ± 104 kg/min. They might now trace the unmeasured disturbance back to a point in the plant where the effects caused by such a disturbance might be observed.

4.6.4 Case 4 Conclusion

The path taken is summarised for a case where a constraint is set inappropriately:

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	NO	
	Partial set-points tracking?	YES	Unmeasured disturbance
	Zero set-points tracking?		
	Unbounded outputs?	NO	
Observed MV Trends	Oscillations?	NO	
	MV saturates?	YES	Constraints on MV
	Aggressive MV behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CV selections?	YES	Good variable selection
Examine CV and MV selections using RGA	Reasonable for given MV?	YES	Good variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	YES	Good output weights
Similar MPC on virtual Plant	MV and CV trends similar to those of real plant?	YES	Too strict constraint confirmed
	Constraint relaxation improved performance?	YES	
Result: relaxation of constraint on an MV improved performance			

4.7 Case 5: Example Relating to Variables Selection

This case study is based on the FCCU example, where MPC is implemented as a supervisory, largely zone control problem as described in section 3.3.3. The hard constraints for the MPC are given in table 3.28. The soft constraints are given in table 3.27, except that there is now no constraint on $C_{O_2,sg}$, which is now implanted as set point control. The other MPC parameters are given in table 4.6 below. The case relates to a situation where two highly coupled variables (T_r and $C_{O_2,sg}$) are used for set point control at the same time, leading to MPC degradation

Table 4.6: MPC parameters for Case 5

Parameter	Symbol	Value
Sampling Time	T_s	120
Prediction horizon	P	30
Control horizon	M	3
Output weights	ow	[10,0,0,0,0,10,0,0,0,0,0];
Input weights	iw	[0.1,0.1, 0.1,0.1,0.1, 0.1,0.1,0.1, 0.1,0.1]

4.7.1 The Scene

A situation is envisaged where for operational reasons, the operators decide that the FCCU should be controlled with both the riser temperature T_r and the oxygen concentration in the regenerator stack gas $C_{O_2,sg}$ at their set-points. They test these requirements by applying unit step down to T_r at the 60 minute mark, and a unit step up to $C_{O_2,sg}$ at the 120 minute mark. The output and input trends for this envisaged scenario are shown in the figures 4.23 and 4.24 below. The trends show that both T_r and $C_{O_2,sg}$ track their set-points.

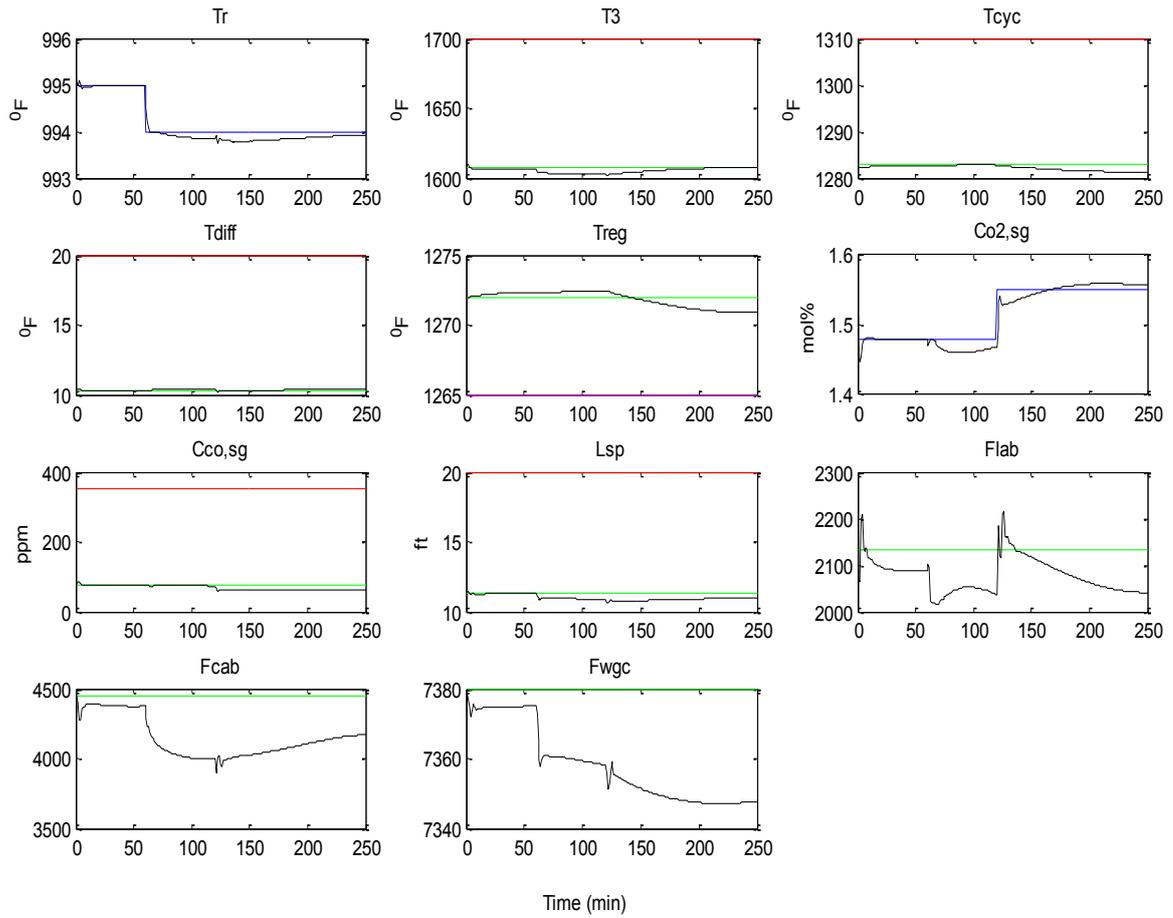


Figure 4.23: Controlled output trends with set-points on T_r and $C_{O_2,sg}$ no degradation (black: trend; blue: set-point; green: nominal; red: upper constraint; magenta: lower constraint)

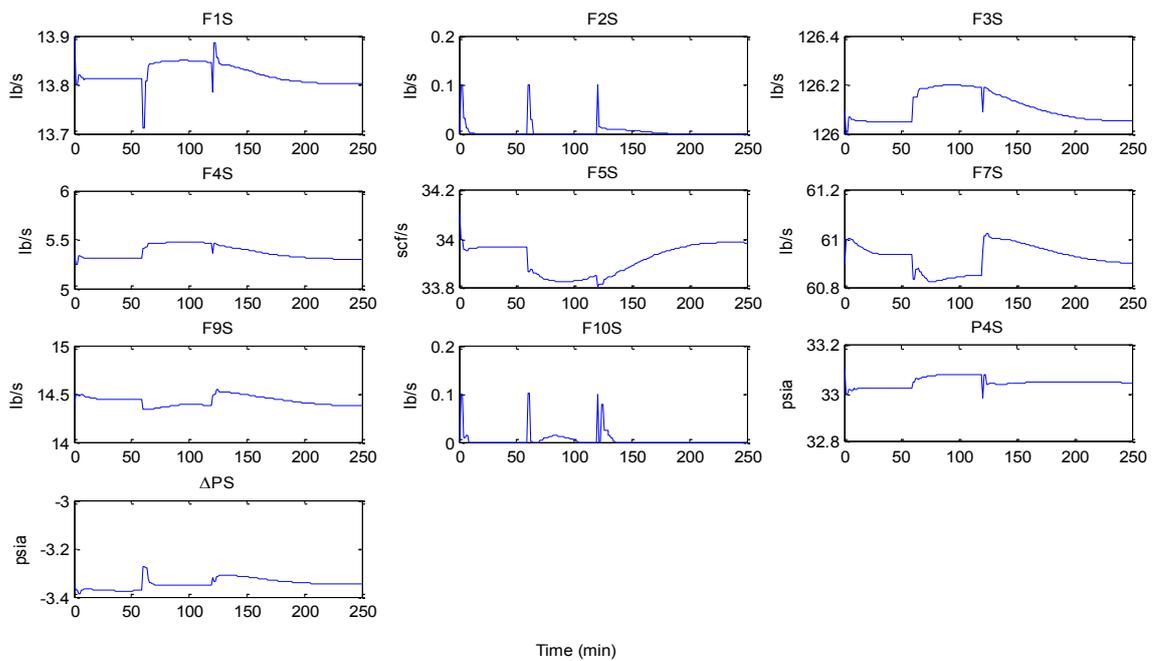


Figure 4.24: Manipulated inputs trends with set-points on T_r and $C_{O_2,sg}$ no degradation

But later they begin to observe the degradation, as shown from the 60 minute mark in figures 4.25 and 4.26. Both T_r and $C_{O_2,sg}$ begin to deviate from their set-points, and $C_{O_2,sg}$ in particular never track its set-point.

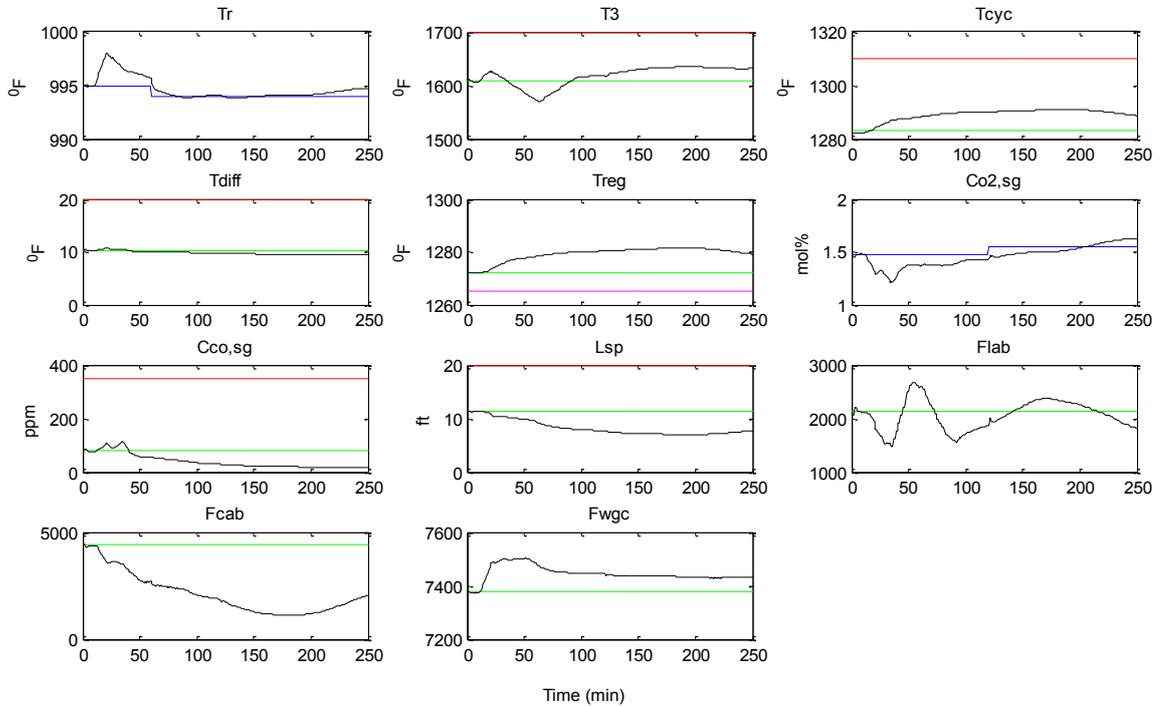


Figure 4.25: Controlled outputs trends with set-points on T_r and $C_{O_2,sg}$, with degradation (black: trend; blue: set-point; green: nominal; red: upper constraint; magenta: lower constraint)

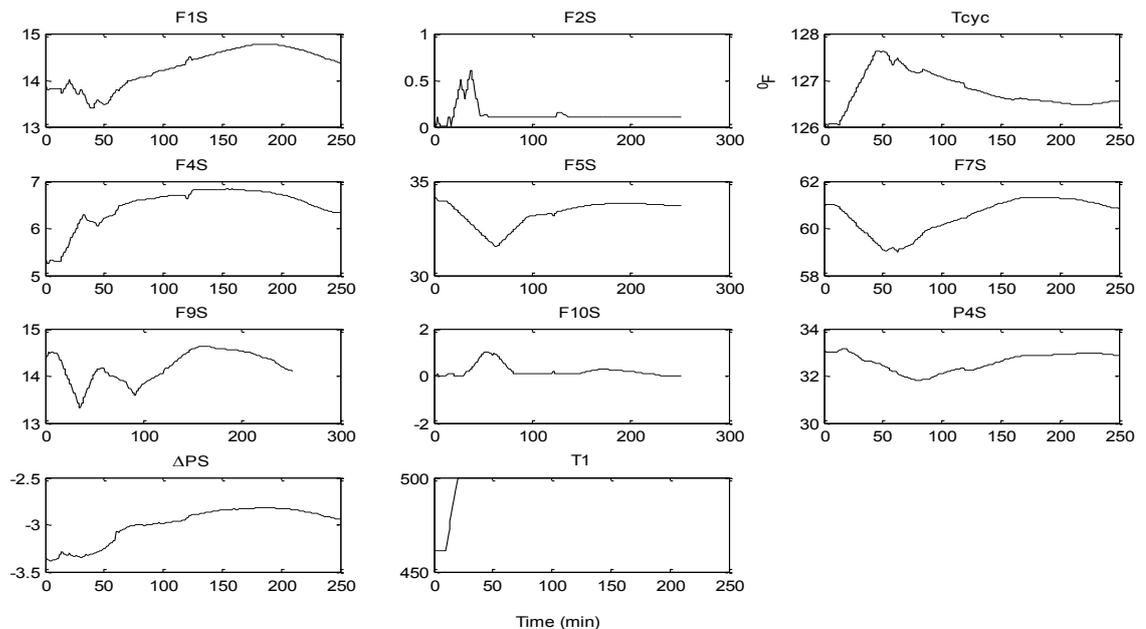


Figure 4.26: Inputs trends with set-points on T_r and $C_{O_2,sg}$ with degradation (T1 is a measure disturbance)

4.7.2 The Operator's Perspective

The operators supposedly go through the checklist of possible causes and confirm that virtually all MPC settings (MPC tuning parameters, constraints, PID settings) remain as before. They then note a change in T1, which is a measured disturbance. There was a ramp increase of 39.1⁰F in T1 over a ten minute period from the 10 minute mark. They might turn to the virtual plant, to see that this increase partly explains the MPC degradation. The trends of the virtual plant outputs are shown in figure 4.27 below.

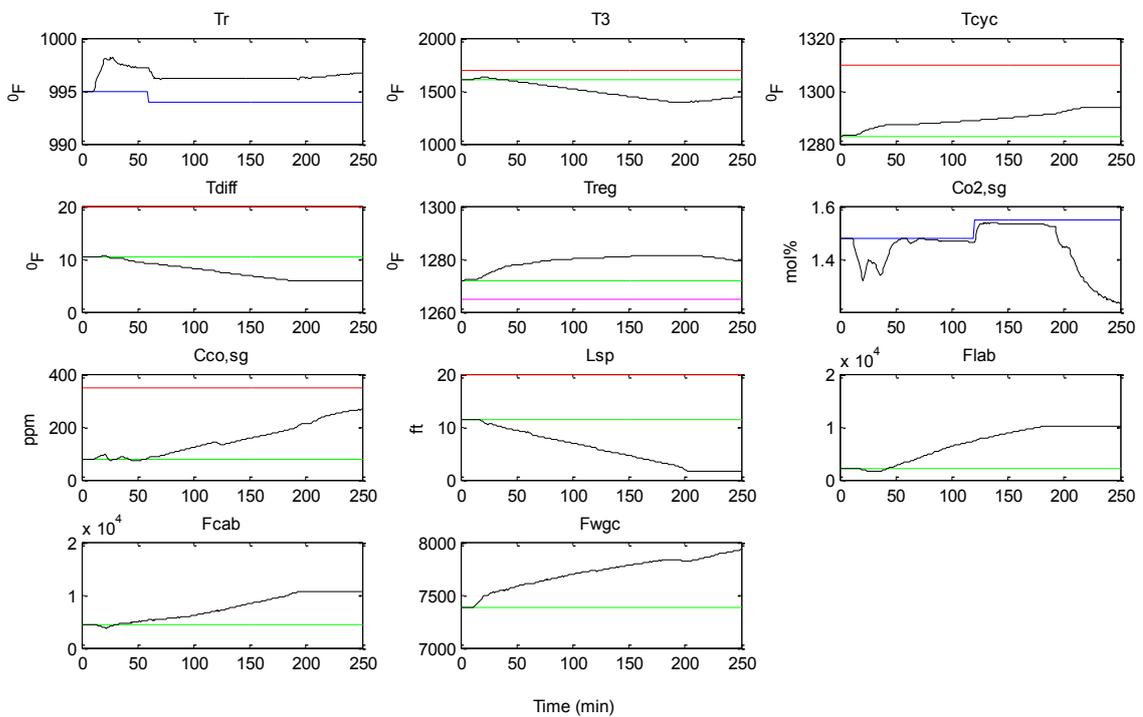


Figure 4.27: Controlled outputs trends from virtual plant with set-points on Tr and CO_{2,sg}, and with ramp increase in T1 (black: trend; blue: set-point; green: nominal; red: upper constraint; magenta: lower constraint)

4.7.3 Reasoning about the scene

The operators might now be fairly convinced that the ramp increase in T1 is partly or solely responsible for the degradation. Disturbances are clearly part of virtually all processes and they just cannot be wished away. The operators might ask: are there underlying factors that made the MPC so sensitive to changes in a disturbance of this nature? Again they might refer back to previous successful operation by studying the

transfer matrix model of the FCCU as shown in table 3.26 or the step response plots of the model shown in figure 3.30. From the table and the figure, the response of Tr due to step changes in each of the manipulated inputs is opposite to that of $C_{O_2,sg}$. In particular both Tr and $C_{O_2,sg}$ are both most strongly affected by the same input: ΔP . This clearly shows that both Tr and $C_{O_2,sg}$ should not have been paired for set-point control as both would always be in conflict. One option is to drop set-point control of $C_{O_2,sg}$. The virtual plant outputs then obtained are shown in figure 4.28.

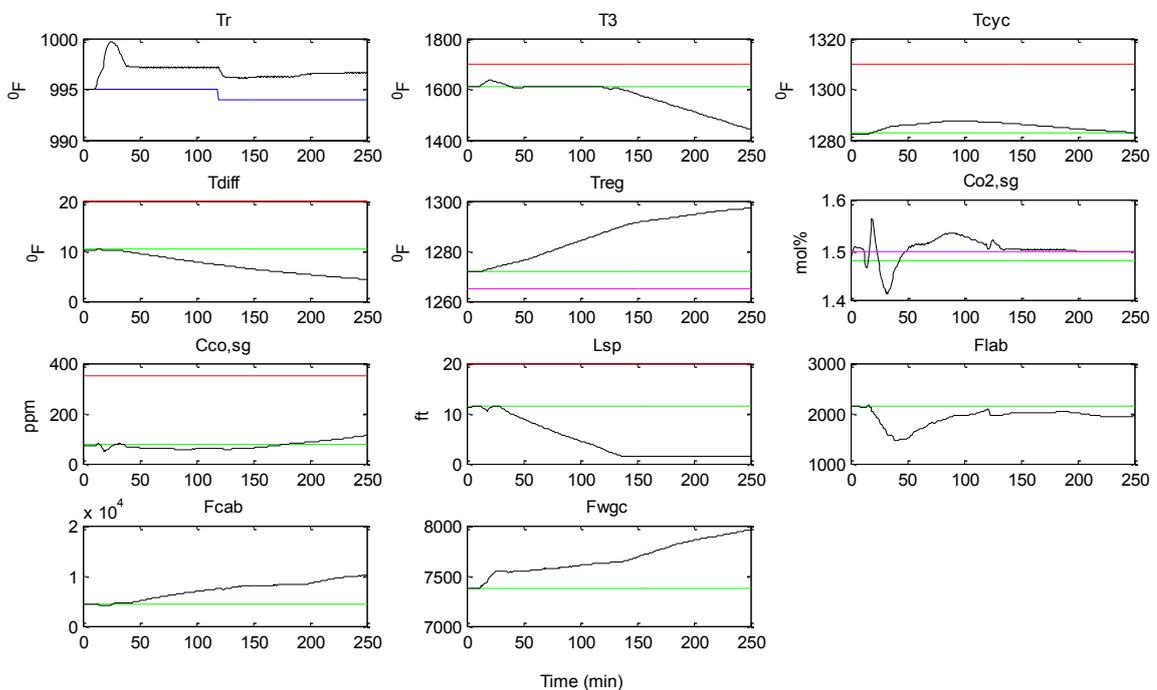


Figure 4.28: Controlled outputs from the virtual plant when set-point on Tr only, and with ramp increase in T1 (black: trend; blue: set-point; green: nominal; red: upper constraint; magenta: lower constraint)

The elimination of set-point control of $C_{O_2,sg}$ ensures stability in the trends of Tr and $C_{O_2,sg}$, but does not achieve good set-point tracking of Tr in the presence of the disturbance T1. In addition, many of the other outputs, though still within their constraints, are exhibiting long settling times. These trends indicate that the process is very sensitive to changes in T1. Further examination of the transfer function matrix of table 3.26 and step response plots of figure 3.30 show that Tr is tightly coupled to all

other outputs (changes in the manipulated inputs that affect Tr also affect other outputs, either completely opposite to Tr or in the same direction), with the exception of T3. T3 is almost solely controlled by F5 and with very fast dynamics.

So an alternative control strategy may be to consider having set-point tracking of T3 together with Tr, or to use a more advanced MPC control where the output constraints can be ranked.

4.7.4. Case 5 Conclusion

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	NO	
	Partial set-points tracking?	YES	Unmeasured disturbance Measured disturbance
	Zero set-points tracking?		
	Unbounded outputs?	YES	Variable selection
Observed MV Trends	Oscillations?	NO	
	MVs saturates?	NO	
	Aggressive MVs behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CVs selections?	NO	Bad variable selection
Examine CV and MV selections using RGA	Reasonable CVs for given MVs?	NO	Bad variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	YES	Good output weights
Similar MPC on virtual Plant	MVs and CVs trends similar to those of real plant?	YES	Bad design or variable selection
	MPC tuning improved performance?	NO	
Result: Virtual plant dynamics reveal that two of the controlled variables should not have been used together. Dropping one of them improved performance.			

4.8 Case 6: Example relating to absence of output constraints in zone control

This case study is based on the FCCU example, where MPC is implemented as a supervisory, largely zone control problem as described in section 3.3.3. The hard constraints for the MPC are given in table 3.28. The soft constraints are given in table 3.27, except that there is now no constraint on T3, which is now implanted as set point control. The other MPC parameters are given in table 4.13 above, except that the output weight of $C_{O_2,sg}$ is zero and that of T3 is 10. The case relates to a situation where though zone control is intended, the output constraints are not specified.

4.8.1 The Scene

A situation is envisaged where the control of the FCCU using MPC is implemented largely as a zone control problem, but with set-point control of Tr and T3. No constraints are specified for the outputs, and the operating conditions are assumed to be perfect, meaning that there are no disturbances. The thinking behind this scenario is that as long as Tr and T3 are kept at their set-points, all the other outputs that are noted to be highly coupled with Tr would either increase or decrease sensibly as Tr changes (depending on whether the correlation is positive or negative) that is these deviations would be acceptable. The trends of the outputs and inputs for this envisaged scenario are shown in figures 4.29 and 4.30 below.

These plots support the assumption: virtually all the other outputs decrease or increase sensibly in the direction suggested in the step response plots of figure 3.30. However if an unmeasured disturbance (the effective coking factor, Ψ_f) is increased, the operators might observe controlled output trends shown in figure 4.31 below.

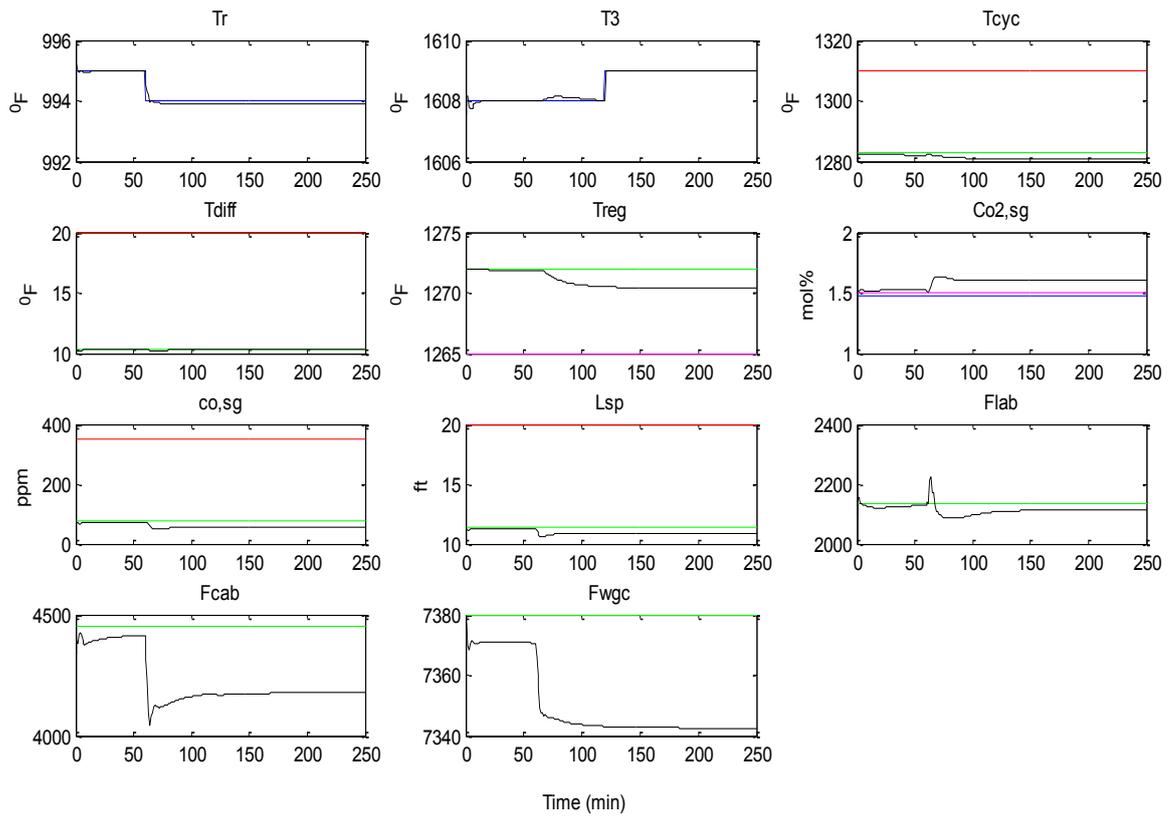


Figure 4.29: Controlled output trends with set-points on T_r and T_3 , no output constraints, no disturbances (black: trend; blue: set-point; green: nominal; red: upper constraint; magenta: lower constraint)

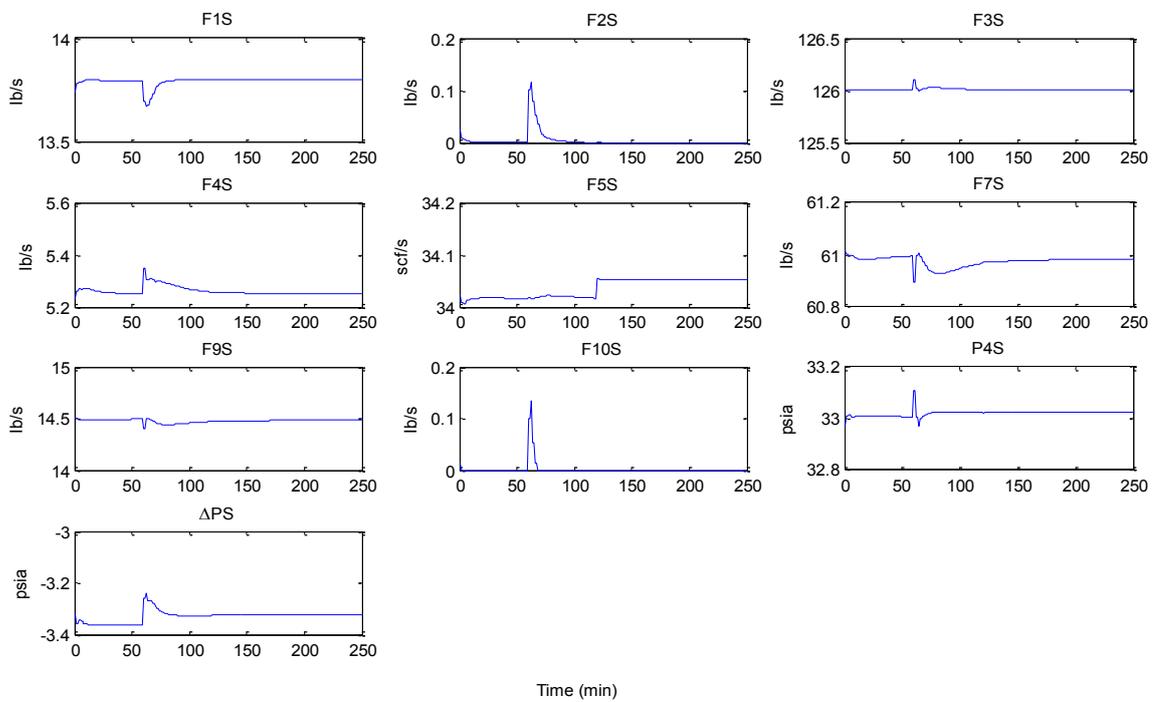


Figure 4.30: Manipulated input trends with set-points on T_r and T_3 , no output constraints, no disturbances

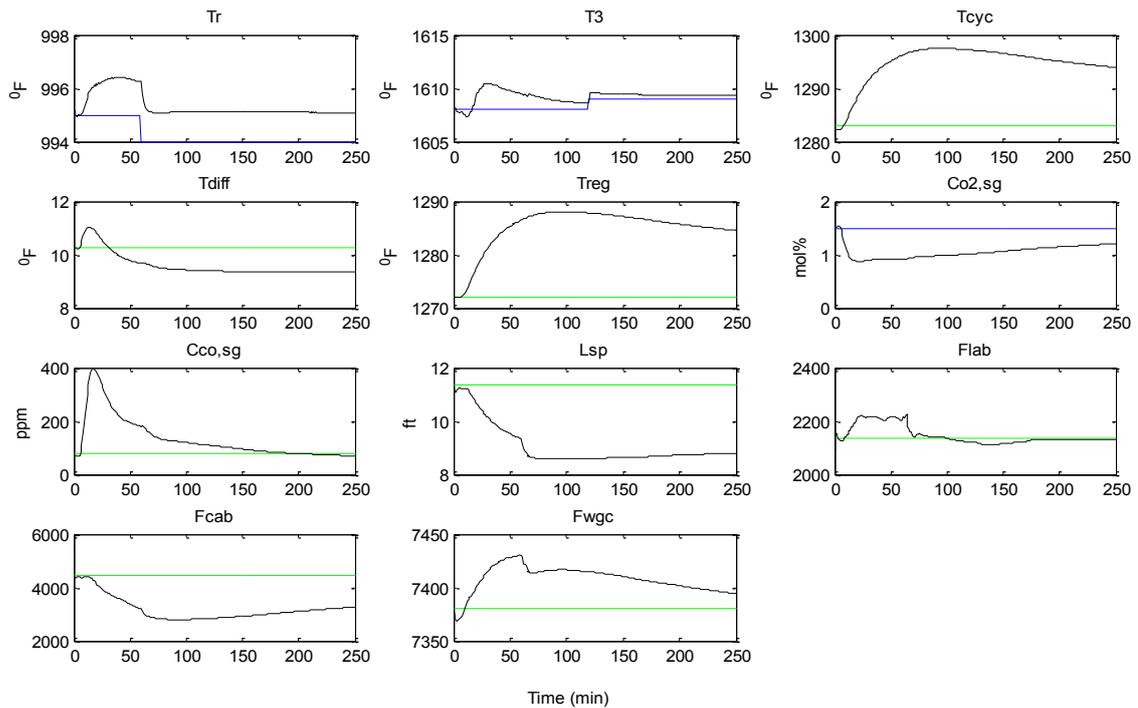


Figure 4.31: Controlled outputs trends with set-points on T_r and T_3 , no output constraints, but with input disturbance

4.8.2 The Operators' Perspective

The operators might observe that though T_r has not deviated much from its previous value, $C_{O_2,sg}$ in particular has deviated from its previous value both significantly and alarmingly, highlighting the need for constraints. It is assumed that nothing else had changed except for the appearance of the unmeasured disturbance.

4.8.3 Reasoning about the scene

In zone control, output constraints are very important to prevent outputs and inputs from wandering away from desired values. The fact that tell-tale signs of constraint activation that are seen in inputs (in terms of saturations) are usually not seen in output variables does not mean that output constraints are not essential in guiding input and output variables with set-point control to desired values.

The output constraints are now included in the simulation of the ‘real’ plant, in the presence of the input disturbance. The output trends for this implementation are shown in figure 4.32 below.

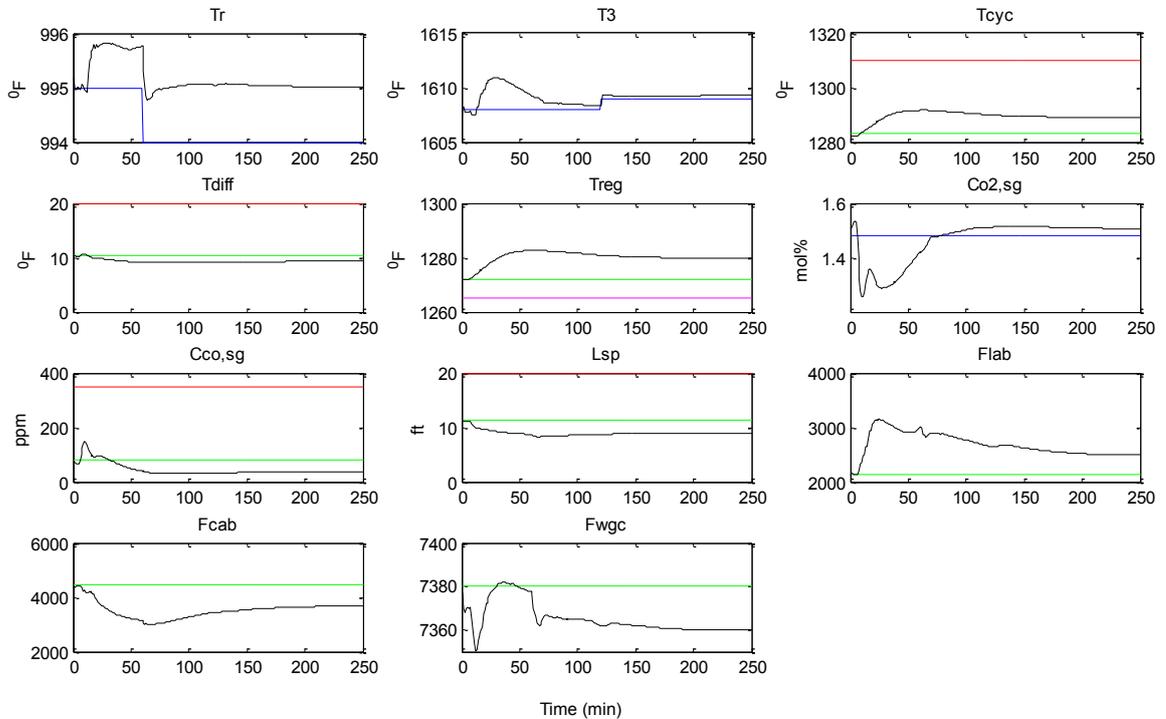


Figure 4.32: Controlled output trends with set-points on T_r and T_3 , with output constraints and input disturbance (black: trend; blue: set-point; green: nominal; red: upper constraint; magenta: lower constraint)

With output constraints, $C_{O_2,sg}$ was guided to settle at a value within its constraint band after about 75 minutes.

4.8.4. Case 6 Conclusion

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	NO	
	Partial set-points tracking?	YES	Disturbance
	Output constraints violation?	N/A	
	Constraint on output specified?	NO	
Observed MV Trends	Oscillations?	NO	
	MV saturates?	NO	
	Aggressive MV behaviour?	NO	
Actions			
Similar MPC on virtual Plant (this does not include effect of disturbances)	MV and CV trends similar to those of real plant?	NO	Include Output constraint on real MPC Investigate source possible disturbance
	Output constraints improved performance?	YES	

Result: Examine the properties of the unmeasured disturbance and the extent to which the operator can affect its value. Return to design value if possible.

Check reveals that an unmeasured disturbance has deviated from its nominal value. Operation at this new design condition would require re-evaluation of the MPC. Otherwise return to previous design value.

4.9 Case 7: Example relating to model plant mismatch

This case study is based on the CSTR example, in which MPC is implemented as a supervisory controller with V and C_A as the controlled variables (Section 4.3.1). The PI settings for the regulatory controllers are as given in table 3.5, and the MPC settings are the same as in table 3.8, except for the outputs weights which are 1 and 200 respectively for V and C_A . The case relates to a situation where the steady state gain for the model representing the $T_{jet}-C_A$ input output channel (table 3.7) is changed from -0.006 to 0.6 to study the effect of model plant mismatch.

4.9.1 The Scene

A situation is imagined where the operators observe deterioration in the tracking of variable C_A , with its offset greatest at the beginning and reducing gradually over time,

as shown in figure 4.33 below. The settling time of C_A has increased significantly even when it is operated at its nominal value of $0.245 \text{ Btu/hr ft}^2 \text{ } ^\circ\text{R}$.

4.9.2 The Operators' Perspective

Having noted the CV trends, the operators are likely to study the trends of the manipulated variables: the MPC outputs, which are the set-points to the two PI controllers (figure 4.34).

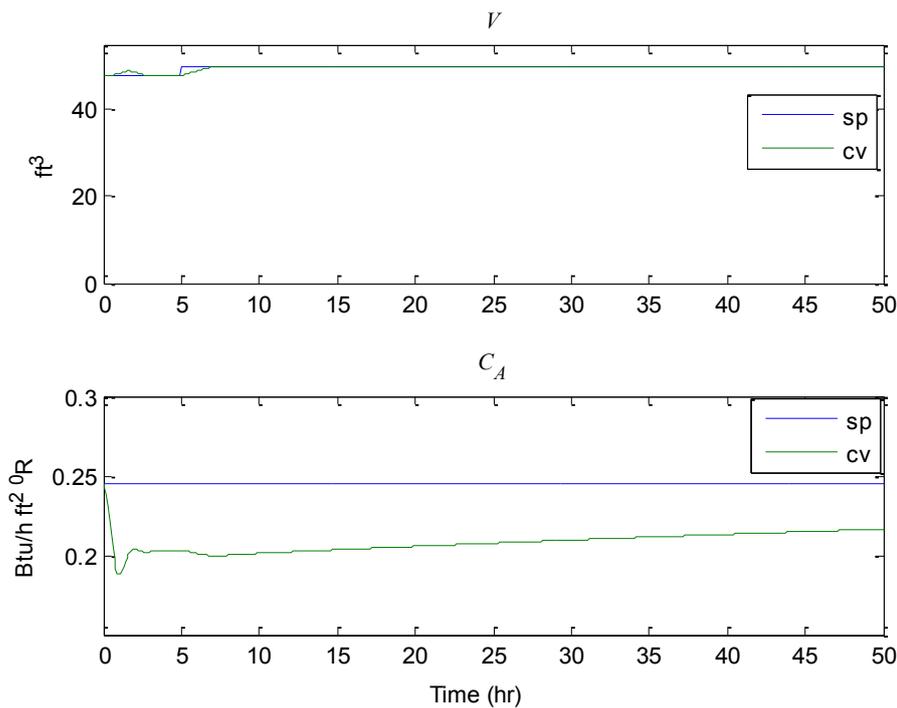


Figure 4.33: Case 7 CV trends

The operators might observe that of the two manipulated variables, the MPC degradation is more pronounced in T_{jset} . V_{set} moves quickly to enable V to track its set-point while movement in T_{jset} is very sluggish. The operators might note that the fact that the two MVs exhibit no oscillations eliminate causes associated valve sticking, deadband etc. They might also observe that the manipulated variables indicate no saturation or constraint activation, ruling out causes associated with constraints on inputs. They might now be well informed that manifestation of such an offset tracking may be due to poor tuning parameters, model-plant mismatch or an unmeasured

disturbance, since the two controlled variables have been shown to be an acceptable selection (Section 4.3.1), and other possible causes have been eliminated. Since there is no aggressive behaviour in the MVs, the operators might see no reason to alter the input weight settings. The operators might also realise that changes to the output weights would not improve the performance. The three disturbances of the system (T_0 , C_{A0} , T_{j0}) are not modelled so the operators are unable to experiment with their effect on the linear model of the process. The operators might now at the last resort decide to check for model-plant mismatch.

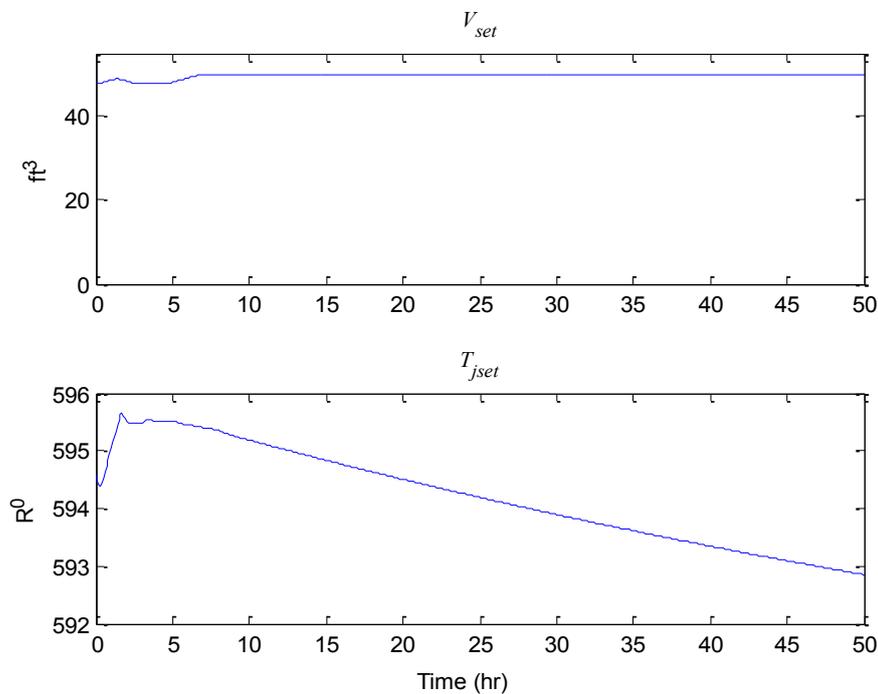


Figure 4.34: Case 7 MVs

4.1.3 Reasoning about the scene

A simple check for model plant mismatch involves putting the plant in open loop (disconnect the MPC) and applying excitation signals (step, random or PRBS) to the inputs and recording the response. The same set of signals is used to excite the linear model of the plant.

The operator might, as described above, apply simultaneous unit step signals to the inputs of the plant and the model, and obtain a comparison plot of the responses, as shown in figure 4.35. From the figure, it should be obvious to the operators that there is model-plant mismatch, especially with regards to variable C_A . This suggests that only input-output transfer function models related to C_A (similar to the one in table 3.7) may require more detailed attention.

The operators might now apply unit steps in turn (step on an input one at a time) so as to gain further insight into whether one or more input-output pairs give rise to the model-plant mismatch.

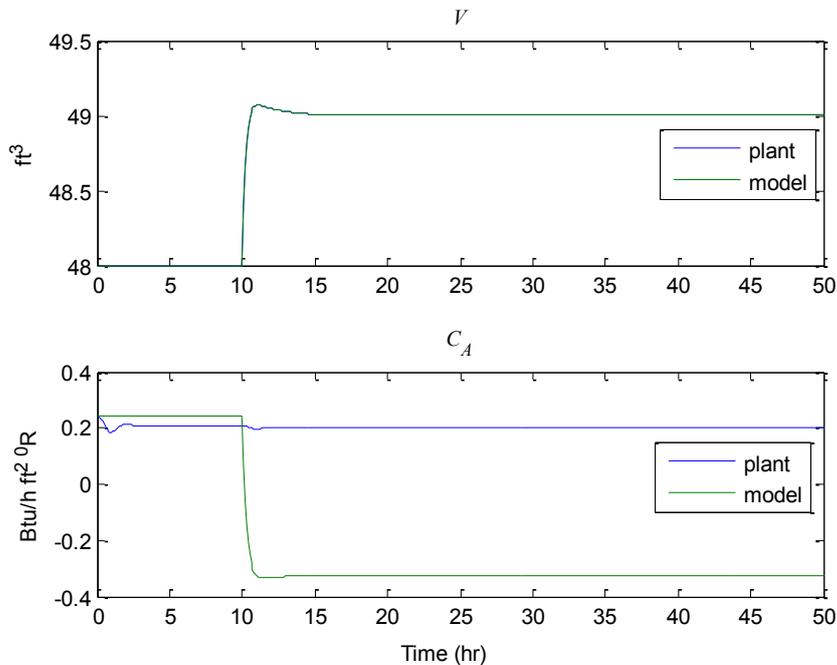


Figure 4.35: Responses of plant and model due to simultaneous application of step signals showing model-plant mismatch

The plots of figures 4.36 and 4.37 below are the response trends when unit steps are applied in turn. By studying the plots of figure 4.35 together with those of figures 4.36 and 4.37, the operators might see that the input-output channel when the model-plant mismatch occurs is $T_{jser}-C_A$. Since the transient dynamics of that input-output channel is not complicated, the operator might simply adjust the steady state gains of that channel

and see what happens. They might indeed discover that appropriate gain specification for the model for the $T_{jset}-C_A$ input-output channel would restore the MPC to good performance.

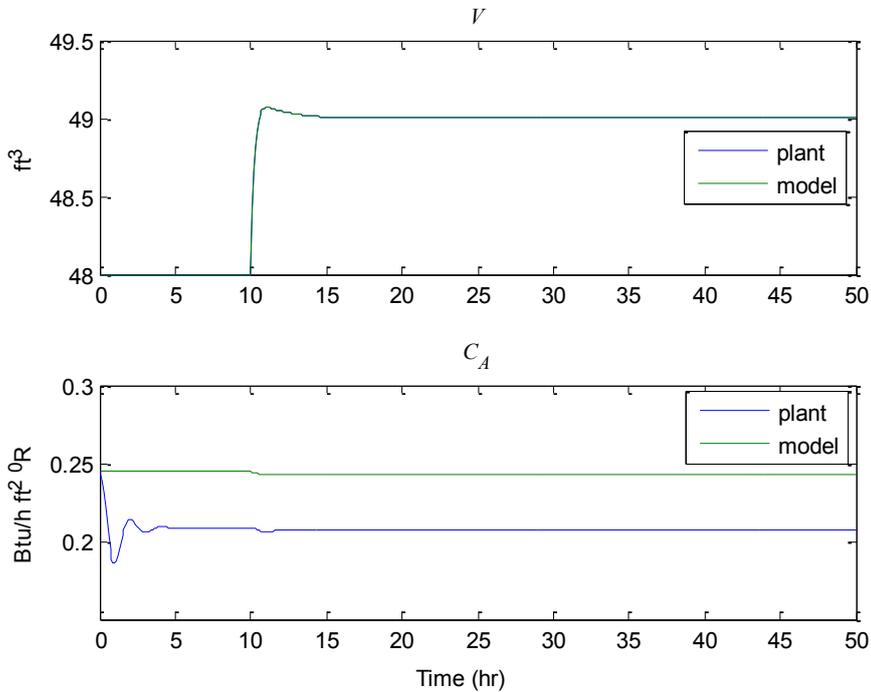


Figure 4.36 Responses of plant and model due to application of step signal to V_{set} only

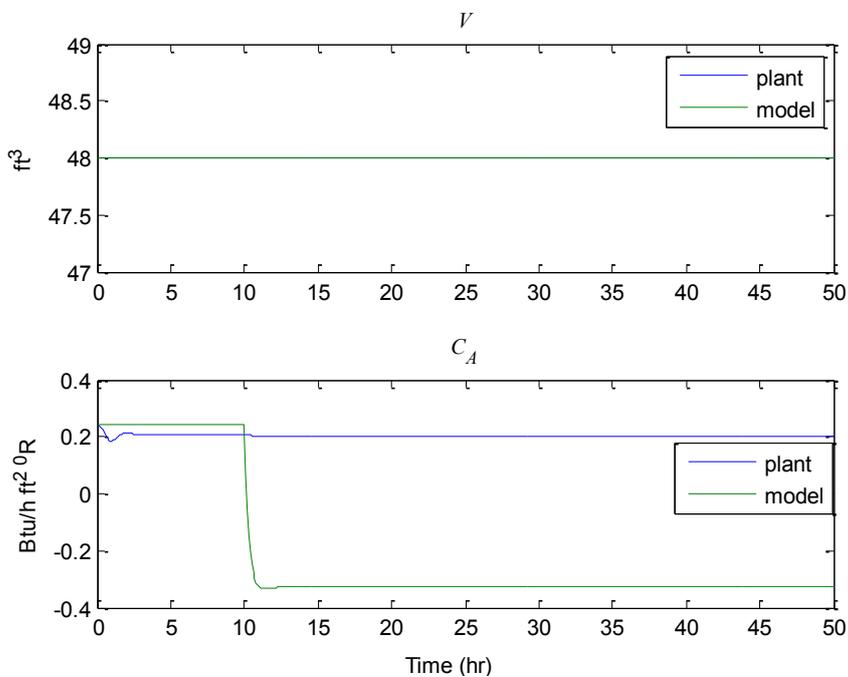


Figure 4.37 Responses of plant and model due to application of step signal to T_{jset} only

4.9.4. Case 7 Conclusion

The following table summarises the path that might be taken to conclude and isolate a case of model/plant mismatch.

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	NO	
	Partial set-points tracking?	NO	
	Zero set-points tracking?	NO	
	Increased settling time	YES	Model Mismatch Unmeasured disturbance Variable selection MPC parameters
	Unbounded outputs?	NO	
Observed MV Trends	Oscillations?	NO	
	MV saturates?	NO	
	Aggressive MV behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CV selections?	YES	Good variable selection
Examine CV and MV selections using RGA	Reasonable for given MV?	YES	Good variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	YES	Good output weights
Similar MPC on virtual Plant	MV and CV trends similar to those of real plant?	NO	Model mismatch
Switch off MPC. Apply excitation signals to the inputs. Plot the step responses	Do the step responses match those of the model?	NO	Comparison of the step responses reveal model/plant mismatch
Result: Model/plant mismatch diagnosed. Adjustment to the model improved performance			

4.10 Case 8: Example relating to PID degradation

This case study is based on the evaporator example, in which MPC is implemented as a supervisory controller (Section 3.4.3). The MPC settings are given in table 4.7. The proportional gain and integral time constant for the regulatory controller used are -10

and 10 seconds respectively. The case relates to a situation where the MPC degrades due to improper tuning of the lower level regulatory controller.

Table 4.7: MPC parameters for case 8

Parameter	Symbol	Value
Sampling Time	T_s	1
Prediction horizon	P	30
Control horizon	M	3
Output weights	ow	[100,10 10]
Input weights	iw	[0 0 0]

4.10.1 The Scene

A situation is imagined where the operators observe oscillations in the controlled and manipulated variables even as the evaporator is at close to designed equilibrium point. The trends of the controlled and manipulated variables as observed by the operators are shown in figures 4.38 and 4.39 below.

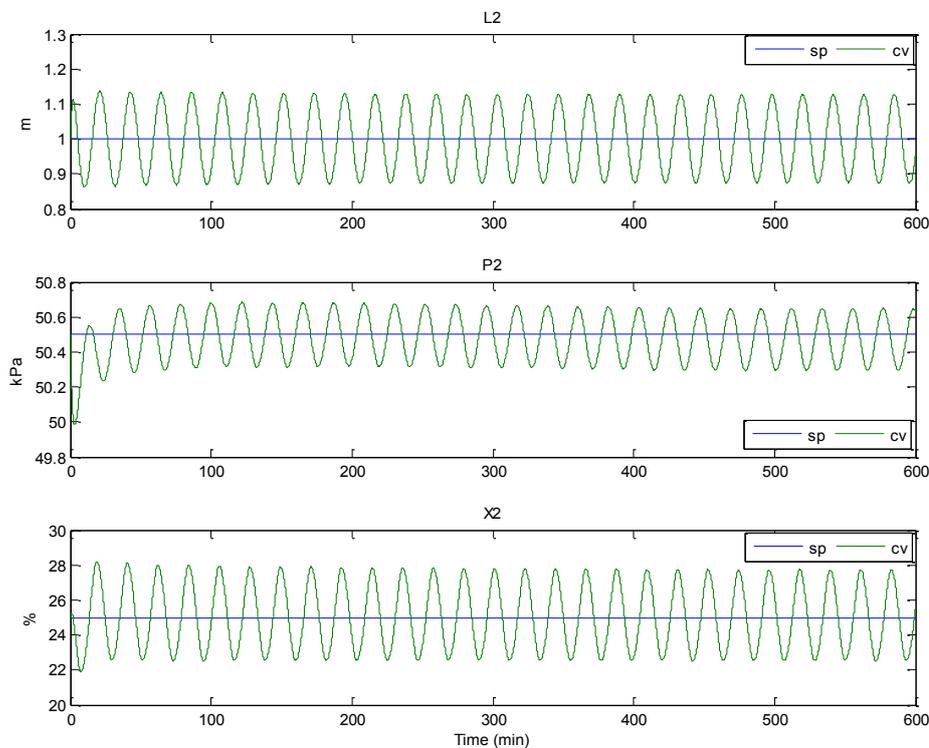


Figure 4.38: Trends of the evaporator controlled variables due to MPC degradation

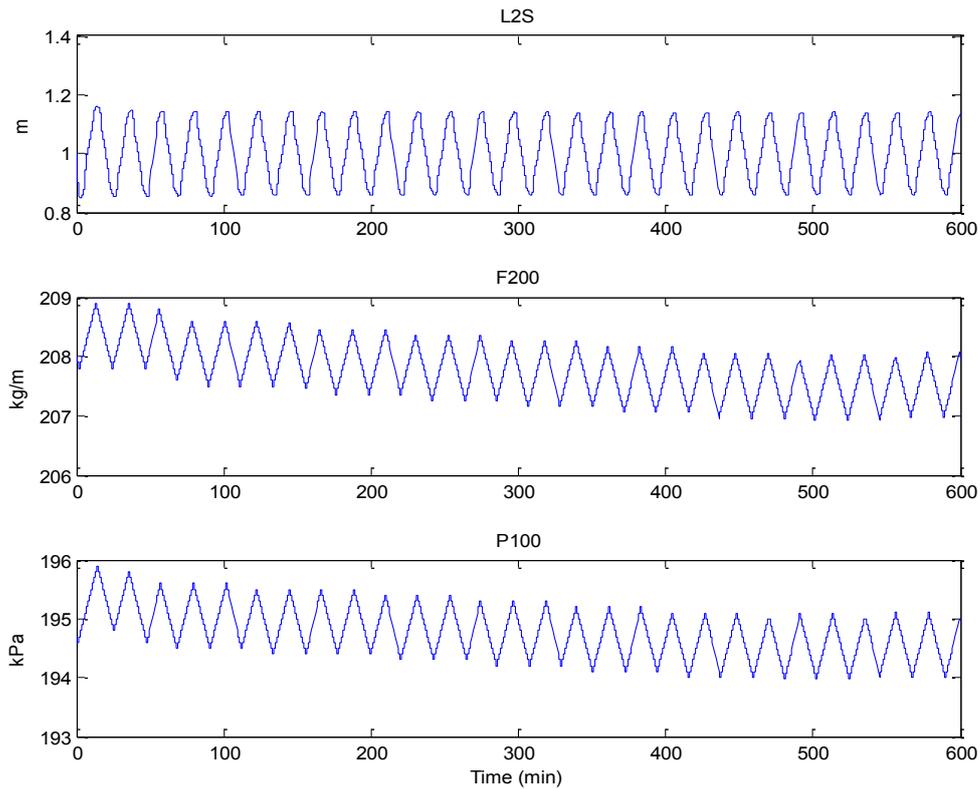


Figure 4.39: Trends of the evaporator manipulated variables due to MPC degradation

4.10.2 The operators' perspective

Having seen trends, the operators might guess that the oscillatory trends in the CVs and MVs are likely due to equipment degradation (valve sticking, deadband etc) or PID degradation; purely MPC performance issues do not normally manifest this way. The operators might then decide to switch off the MPC momentarily and control the plant only with the regulatory controllers on. The operators might use a value of 1.0 m as the set point of L2-L2S PI loop and 208.0 kg/m and 194.7 kPa as the set-points of the other two local servo controllers respectively. If this was done, the operators might see that the controlled output trends look like as shown in figure 4.40 below.

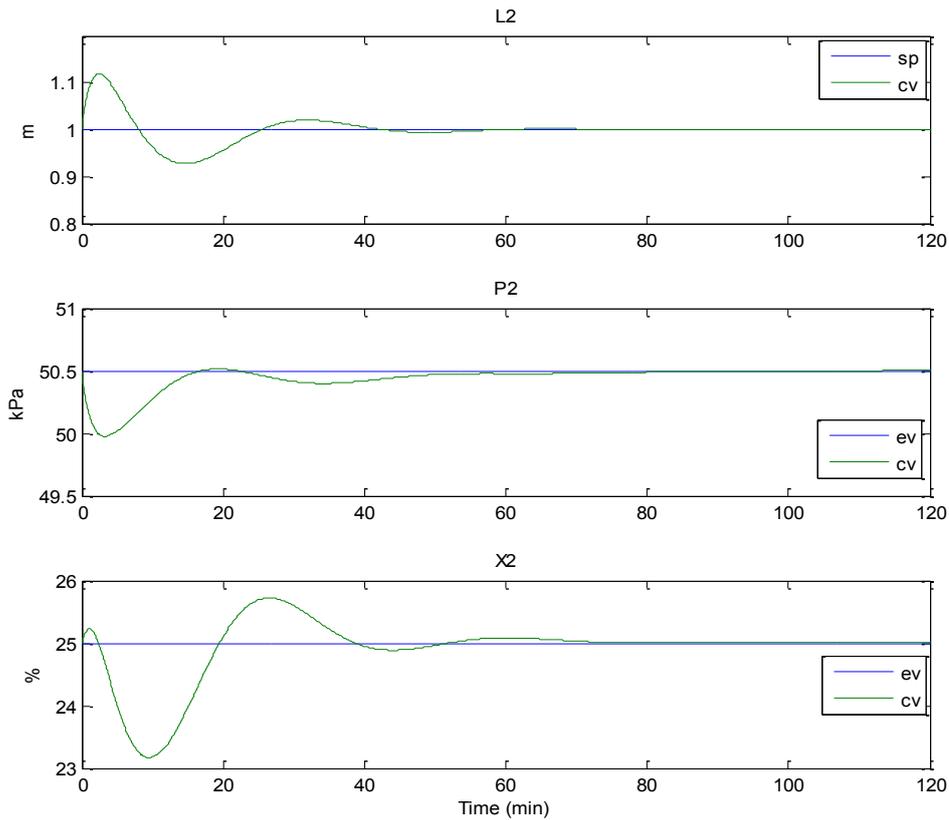


Figure 4.40: Controlled variable trends of evaporator with regulatory control only

In figure 4.40, sp (in the top plot) is the set point of the L2-L2S PI loop while ev (in the middle and the bottom plots) are expected values when the set-points of the PI controller and the other local servo controllers are as used in this case. As can be seen in figure 4.40, the operators might observe that their settling times appear larger than usual. At this point the operators might decide to take a closer look at each of the regulatory controllers.

4.10.3 Reasoning about the scene

When the plant is operated with regulatory controllers only, the operators might now obtain plots of the regulatory controllers' set-points (sp) against the controllers' measured outputs (pv) (figure 4.41). They might then obtain a comparison plot of measured controllers' outputs (pv_m) and with benchmark controllers' outputs (pv_{bm}) obtained during commissioning or during a period of good performance, with the

assumption that the same set of set-point values are used in both cases. Such a comparison plot is shown in figure 4.42 below.

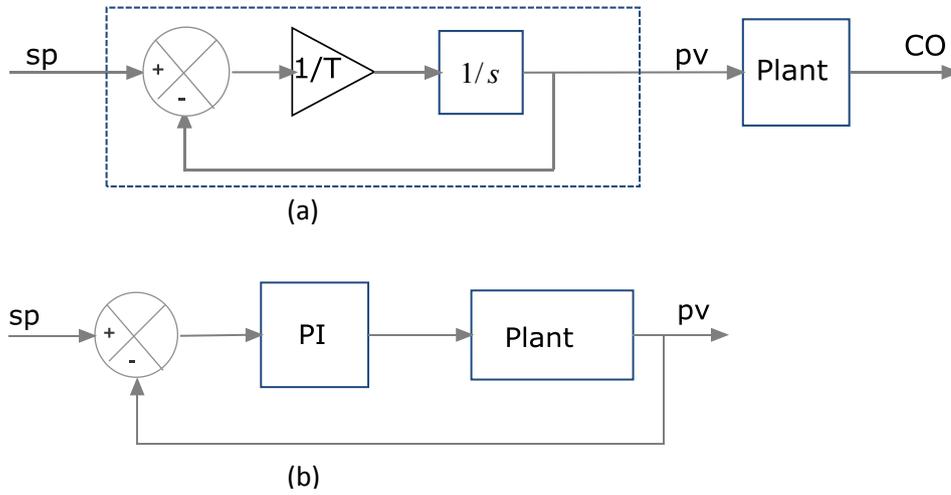


Figure 4.41: sp and pv for regulatory controllers (a) local servo controller, (b) PI controller

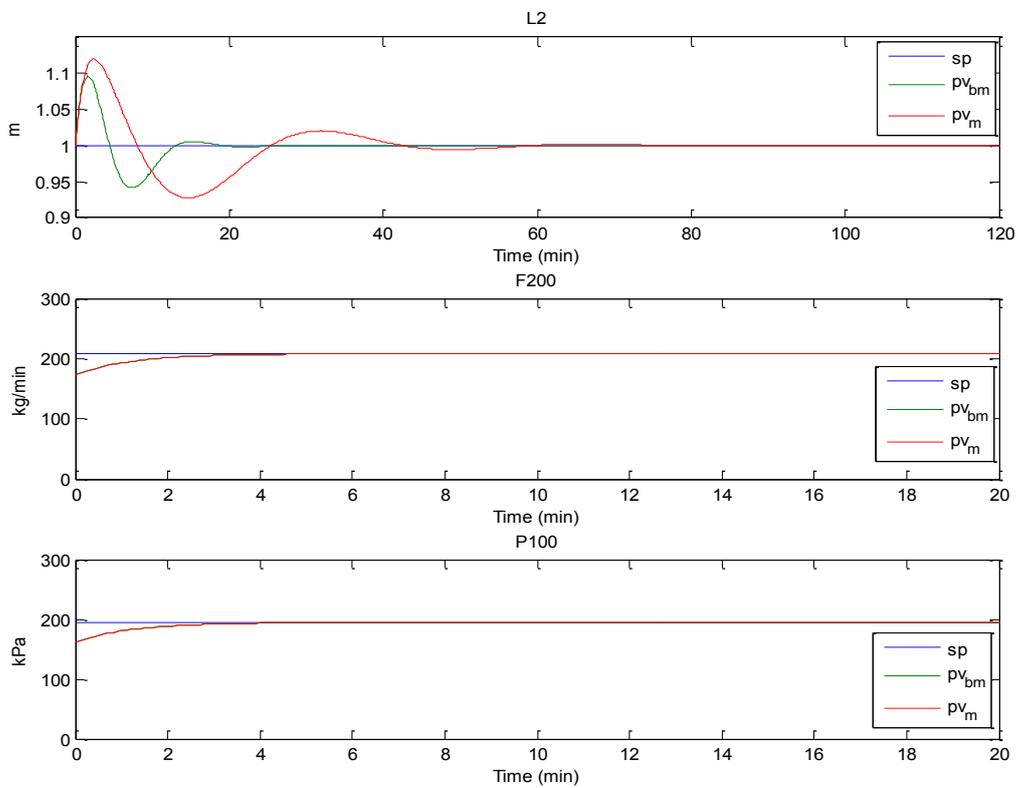


Figure 4.42: sp and pv plots for the regulatory controllers

The trends of figure 4.42 should reveal that p_{v_m} for loop L2-L2S differ substantially from $p_{v_{bm}}$, while there are no discrepancies in the p_{v_m} and $p_{v_{bm}}$ for the other two loops. The p_{v_m} for L2-L2S loop has higher settling time and overshoot valued compared its $p_{v_{bm}}$ which can be due to low PI gain value or high integral time constant.

The operators might then discover that retuning this loop to conform with the benchmark trend ($p_{v_{bm}}$) in figure 4.42 would restore the MPC to good performance.

4.10.4. Case 8 Conclusion

The diagnosis of the root cause of the MPC degradation is given in the following table

	Questions	Answer	Suspicious
Observations			
Observed CV Trends	Oscillations?	YES	Equipment failure PID degradation
	Partial set-points tracking?		
	Zero set-points tracking?		
	Unbounded outputs?	NO	
Observed MV Trends	Oscillations?	YES	Equipment failure PID degradation
	MV saturates?	NO	
	Aggressive MV behaviour?	NO	
Actions			
Examine linear Model dynamics (step response plots and transfer function matrix)	Reasonable CV selections?	YES	Good variable selection
Examine CV and MV selections using RGA	Reasonable for given MV?	YES	Good variable selection
Examine CV and MV pairings using RWA	Reasonable output weights for CVs	YES	Good output weights
Similar MPC on virtual Plant (virtual plant does not include effect of PID or equipment)	MV and CV trends similar to those of real plant?	NO	<ul style="list-style-type: none"> • Good MPC performance • Investigate equipment failure
	MPC Tunings improved performance?		
Run the plant in open loop (No MPC)	Performance improved	NO	Poor PID control PID loop degradation suspected
	Retuning the PID improved performance?	YES	
Result: PID loop degradation diagnosed. Retuning improved performance			

4.11 Summary

This Chapter simulated eight different cases of degradation in MPC performance. The CSTR and the evaporator were used for three cases each while the FCCU was used for the other two. Descriptions of how an average operator might reason if the degradations were to manifest on real plant were given, followed by assumed perspectives of technically minded operators about the symptoms and the suggested appropriate diagnostic procedure, taking cognisance of the dynamics of the systems. By observing the trends and making use of various forms of information (the process dynamics, process step response plots, the transfer function matrix, steady state gains, the relative gain array and a novel measure derived from the steady gains, the relative weight array) the reason for the degradations, the diagnoses and faults isolation, as well as the recovery procedures were explained transparently. Each case study was concluded by outlining the diagnostic steps that were taken towards investigating, isolating and recovering from the MPC degradation simulated.

Chapter Five

The Maintenance Tool Development

5.1 Perceptions of the Maintenance Tool and the operators

This chapter outlines a vision of what a maintenance tool might look like. In practice any package that would be suitable for operator use, would have to be produced by systems developers who would need to have graphical user interface (GUI) and perhaps ergonomics expertise. The outline described here should be viewed as a mechanism for engaging with other researchers, end-users and vendors, and little else.

The MPC operators that are envisioned here are not the high end advanced automatic control process engineers (who are not prevalent anyway), but the non-specialist engineers tasked with keeping most process plants running. The maintenance tool envisioned here is founded, in part, on a brief industrial visit that was made to Warri Refining and Petrochemical Company (WRPC), located in the southern part of Nigeria. At the time of this industrial visit, the company did not have MPC installed (there are plans to install one in the near future). A Yokogawa Electric Corporation Centum XL series distributed control system (DCS) was installed instead. This control system was installed in the early 1990s. The same operators will be around when MPC is eventually installed on the plant. So the understanding obtained during the visit, of how an engineer 'thinks' will carry-forward including their perceptions and limitations into this new approach. A DCS would still be installed as a regulatory control system, because this is still an integral part of most processes with MPC installations.

At WRPC the operators had a very good understanding of their plant, but were not well supported when it came to assessing performance in terms of plant efficiency. Although

they had access to current DCS input and output variables, they had no way of converting the data to historical graphical plots or benchmark plots, to see whether performance was improving or degrading. On a more positive note, DCSs are founded on individual PID loops making it easier to isolate a faulty loop.

5.2 Proposed Scope and Components of the Maintenance tool

The maintenance tool envisioned in this thesis is completely data-driven. Based not on complex statistical analysis of process data, but on observing process data and their trends, the tool seeks to help the operator make sensible judgements about performance degradation, the form and direction of diagnosis and fault isolation, and possibly, the recovery procedure. A particular feature of the tool is the ability to compare the actual performance trends for a given length of time with historical trends for approximately the same length of time. Operators might observe the following types of abnormalities in actual performance trends as evidence of MPC degradation:

- a) offsets in controlled variables;
- b) an unresponsive controlled variable or manipulated variable;
- c) saturation of manipulated variables;
- d) oscillations in controlled and/or manipulated variables;
- e) an over aggressive manipulated variable;
- f) output and/or input constraint violation;
- g) increased settling times of controlled variables;

Issues which might lead to these observed abnormalities in performance trends include the following:

- (i) inappropriate variable selection
- (ii) model/plant mismatch

- (iii) improper constraint specification
- (iv) PID degradation (PID tuning)
- (v) sensor/actuator failure
- (vi) poor MPC tuning
- (vii) poor MPC design

The maintenance tool is envisaged to have many assessment windows through which the operator might progress to detect MPC degradation, and carry out certain activities to confirm or dispel suspicions about the cause of the abnormality. The assessment windows are in five major groups, as shown figure 5.1 below. The first is the trends comparison group, after which comes the diagnostic questions group. After the diagnostic questions group are the suspected symptoms, the symptoms investigation and the MPC recovery groups. Central to the choice of and arrangement of the assessment windows is the role of the operator, who is expected to have the basic understanding needed to take appropriate actions at every stage of the assessment. The roles contained in the windows are outlined below.

5.3 Trends Comparison Assessment Group

The trends comparison group consists of two split windows (figure 5.2): the reference graphical performance (RGP) window, and the actual graphical performance (AGP) window. The RGP window remains static most of the time. It displays trends of data obtained well before degradation is suspected, perhaps at commissioning and during periods of excellent performance. The window would be generated and updated by technical operations personnel, who would re-evaluate its status regularly. The AGP window displays similar trends to the RGP, but now the trends are periodically updated during actual plant operation.

The rationale behind the trends comparison group is that the operator would become familiar with trends associated with normal operation. If unusual features were to appear in one or more of the actual trends, the operator would be encouraged to move to the diagnostic questions windows (figure 5.1). Sections 5.3.1 and 5.3.2 describe the type of trends that the RGP and AGP windows might display.

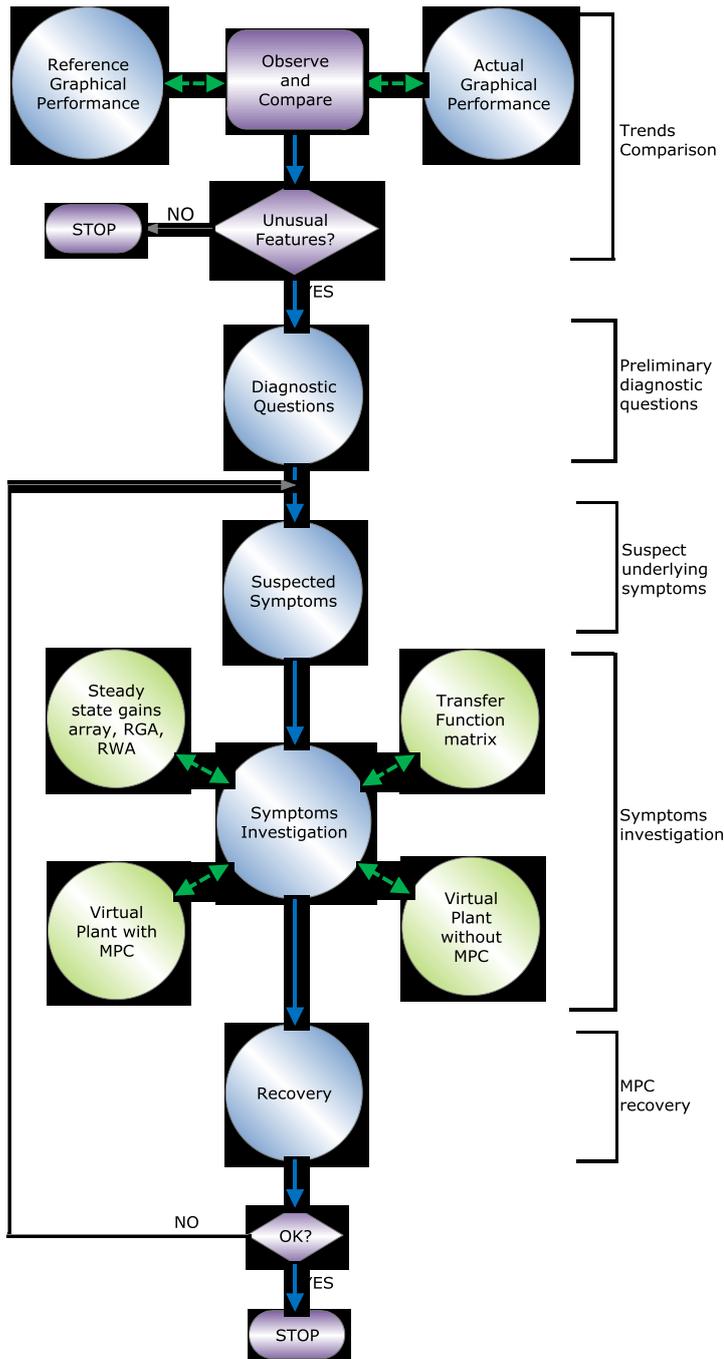


Figure 5.1: Structure of the Maintenance tool

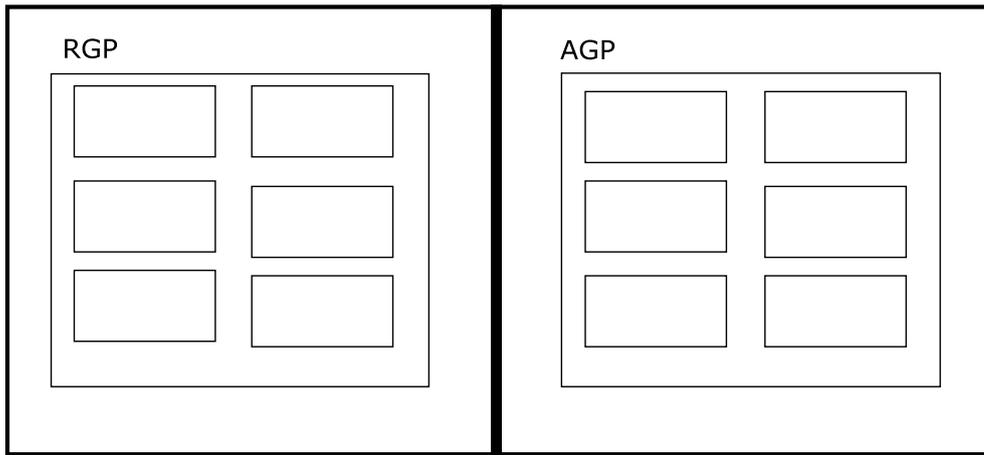


Figure 5.2 Arrangement of windows in the trends comparison assessment group

5.3.1 Reference Graphical Performance Window

The reference graphical performance window displays trends associated with good performance. The underlying data for the trends may be updated, for example if there is a change of operating condition. With reference to figure 2.4, the reference graphical performance window might display the following trends:

Plot 1) rv against time and cv vs time (on the same plot);

Plot 2) sv against time (sv are the outputs of the model predictive controller, which are the set-points for regulatory control);

Plot 3) mv against ov . This is to benchmark the performance of valve actuation. For a perfectly functioning valve the plots should have constant positive slopes at a steady state operating point.

Taking the CSTR case of Section 3.3.2 as an example, plot 1 would be as shown in figure 5.3, plot 2 would be as shown in figure 5.4, and plot 3 would be as shown in figure 5.5.

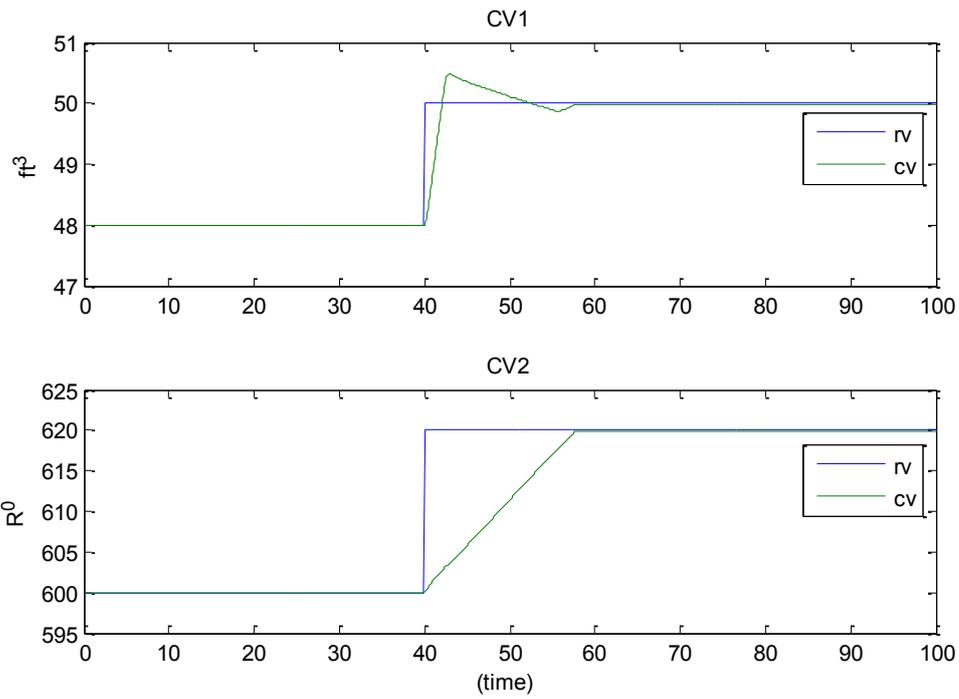


Figure 5.3: Sample rv and cv trends

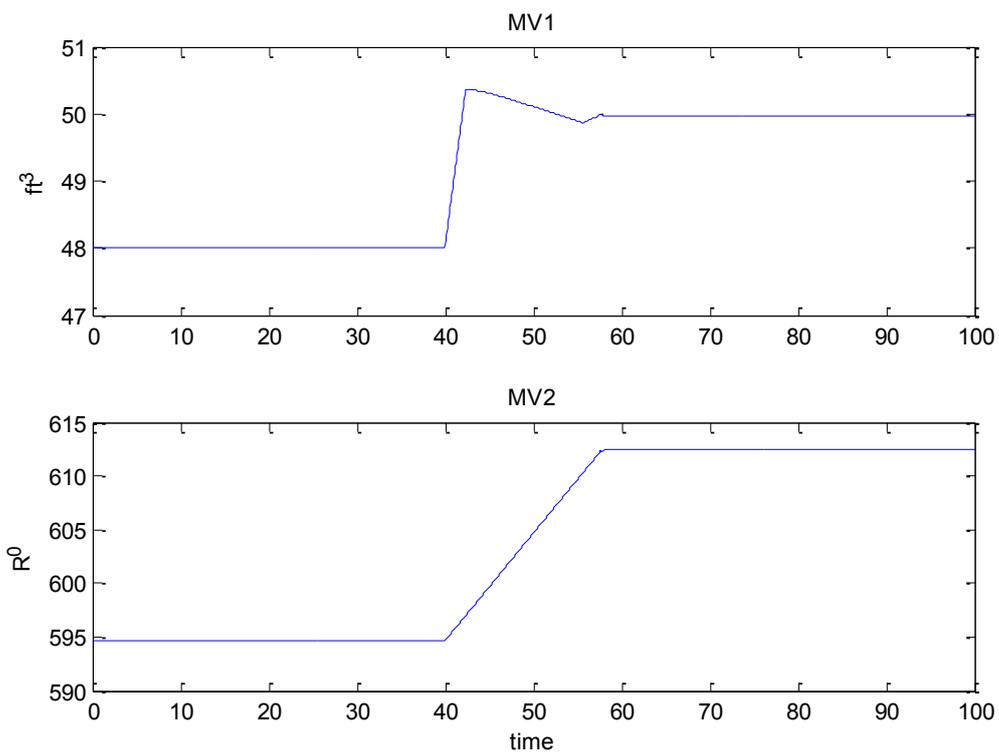


Figure 5.4: Sample sv (MPC manipulated variable) trends

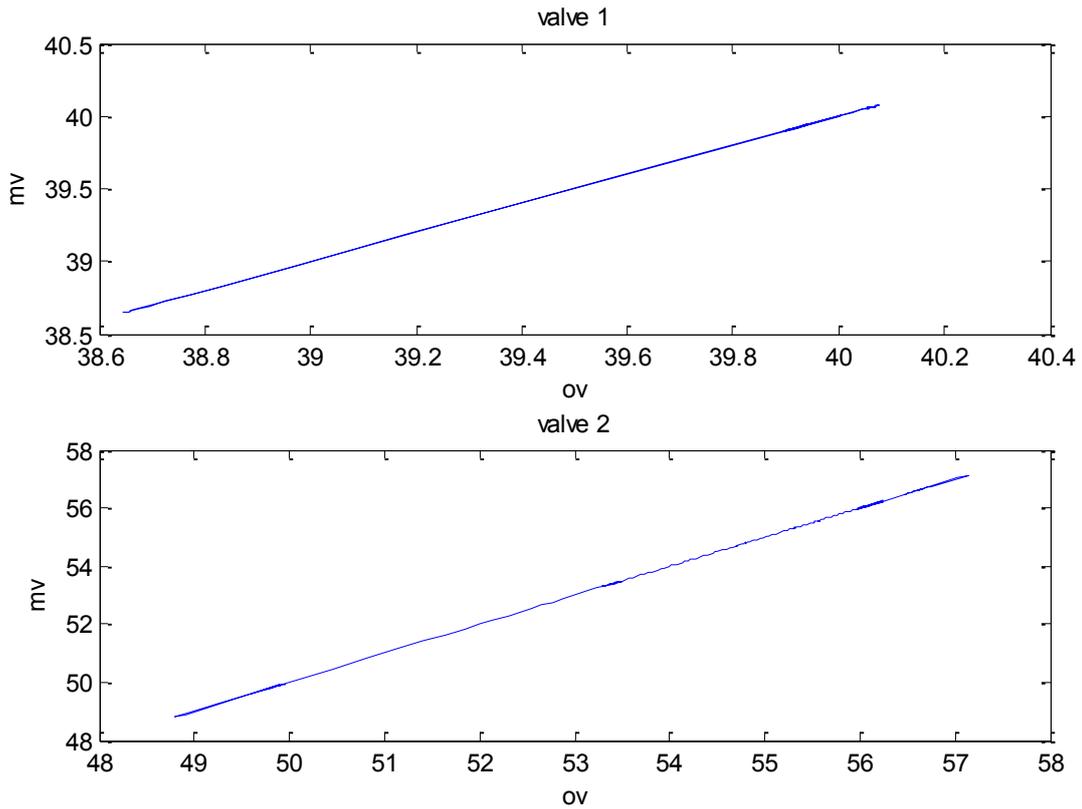


Figure 5.5: Sample mv against ov plot

During commissioning but also during periods of normal operation in which regular maintenance is carried out, it is possible that open loop tests would be performed on regulatory control loops. Plots 4 and 5 (explained below) exploit information generated during these activities.

Plot 4) sv_o against time and pv_o against time (on the same plot). The subscript o indicates the values obtained when the regulatory controller is in open loop, and the controller is of the type shown in the upper loop in figure 2.4. The plot is obtained by applying a step to the input of a loop while the input to the other loops remains at equilibrium. This plot serves as a performance reference for that particular PID loop. A sample plot is shown in figure 5.6

Plot 5) sv_o against time and mv_o against time (on the same plot). The subscript o indicates the values obtained when the regulatory controller is in open loop, and

the controller is of the type shown in the lower loop in figure 2.4. The plot is obtained by applying a step to the input of a loop while the input to the other loops remains at equilibrium. Like plot 4, this plot serves as a performance reference for that particular PID loop.

Plot 6) mv_o against ov_o . The subscript o indicates the values obtained when the plant is in open loop. This will essentially be similar to that of figure 5.5 for perfectly working valves as shown in figure 5.7 below.

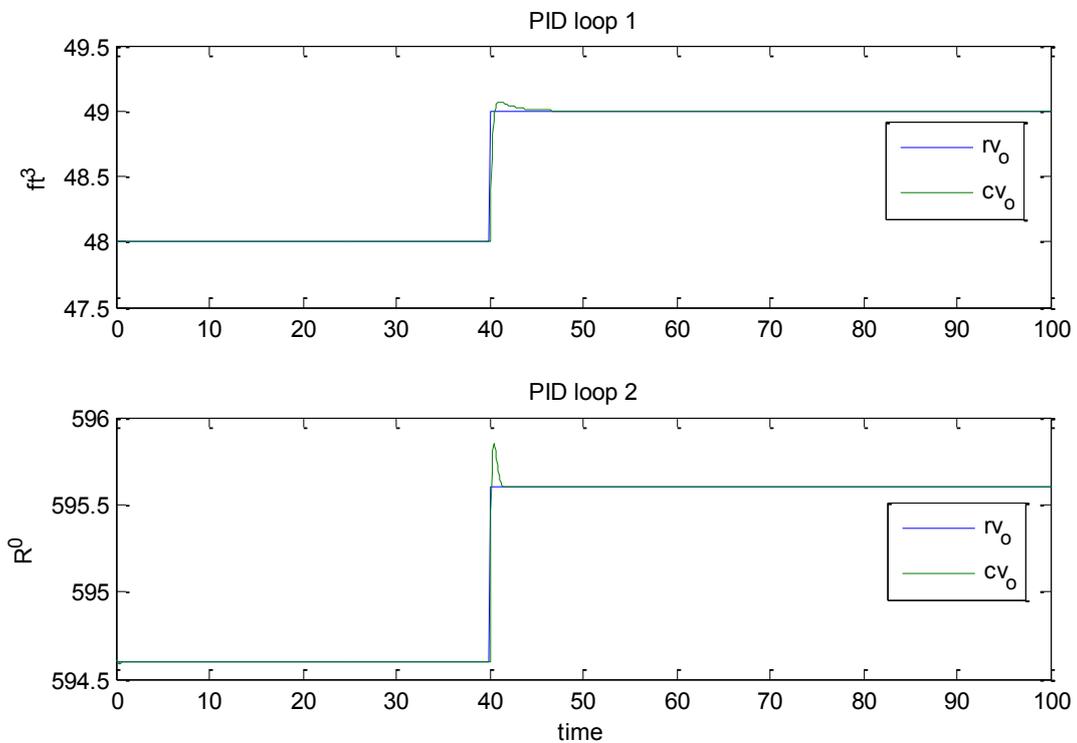


Figure 5.6: Sample r_{v_o} and cv_o trends (open loop)

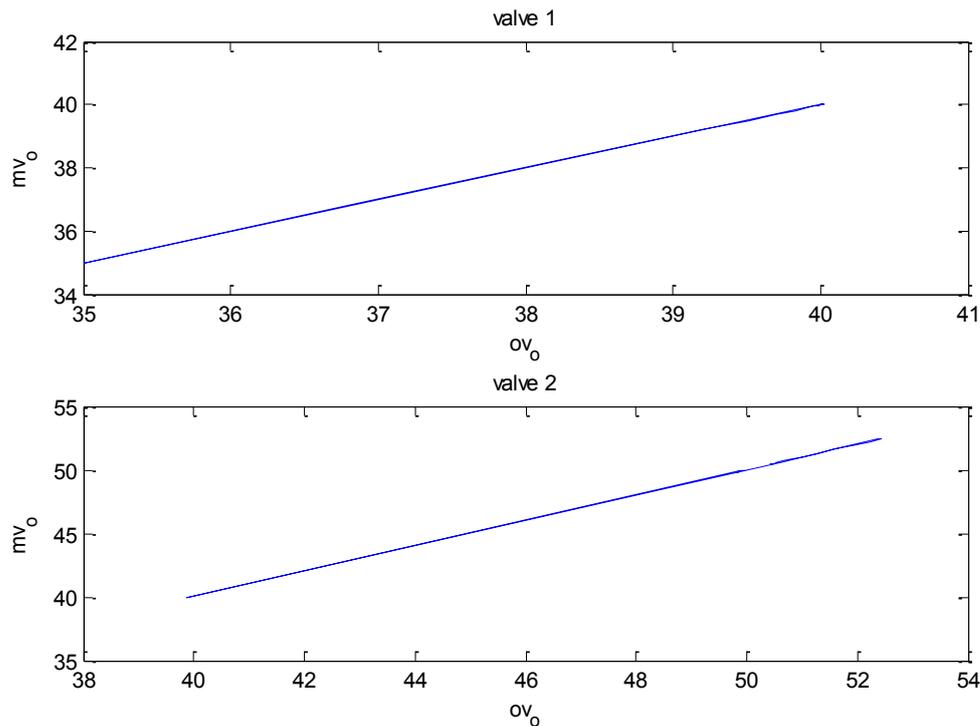


Figure 5.7: Sample mv_o against ov_o plot (open loop)

5.3.2 Actual Graphical Performance Window

This window displays performance trends similar to those described in Section 5.3.1 except that they are now updated at regular intervals. The actual performance data for updating the counterpart trends for plots 1 to 3 in Section 5.3.1 would be obtained from normal operating data captured by the data acquisition equipment, which should be an integral part of the control installations. The data for the counterpart trends for plots 4 to 6 in the same section would be obtained during maintenance only, since the regulatory controllers have to be in open loop for the operation. The data length used in both cases would be reasonably close to that of the reference graphical performance.

The operator compares these trends with those displayed in the reference graphical performance windows and concludes whether the observed trends are normal. If they are not, then the operator would move on to the diagnostic questions window (figure 5.1).

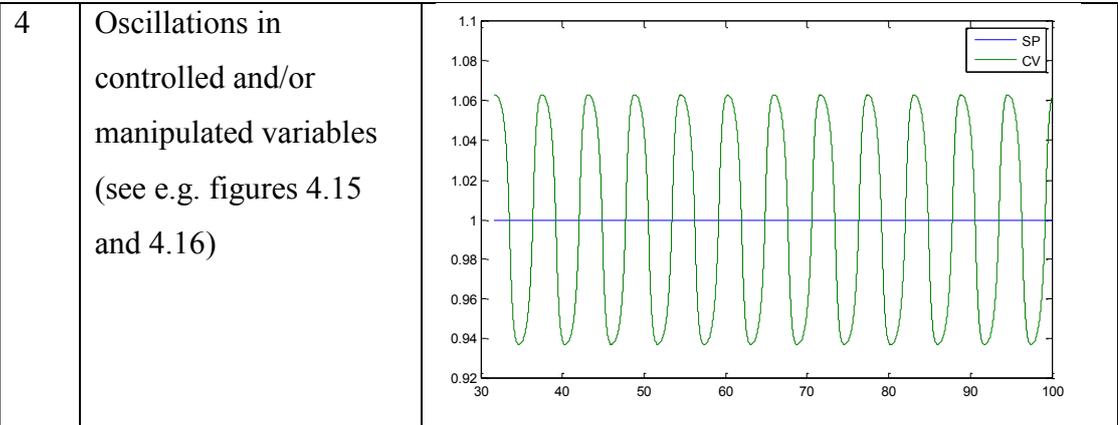
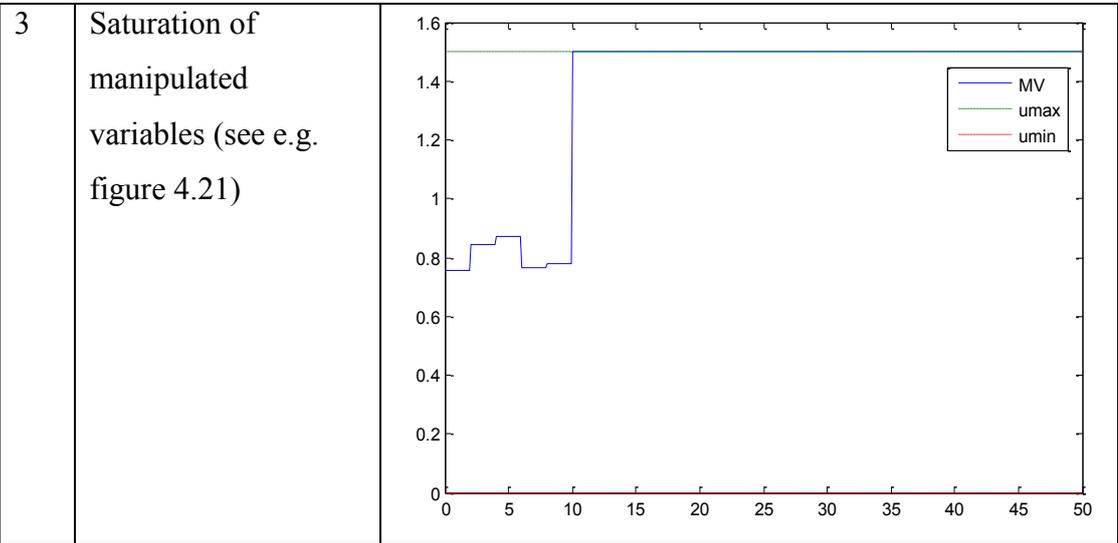
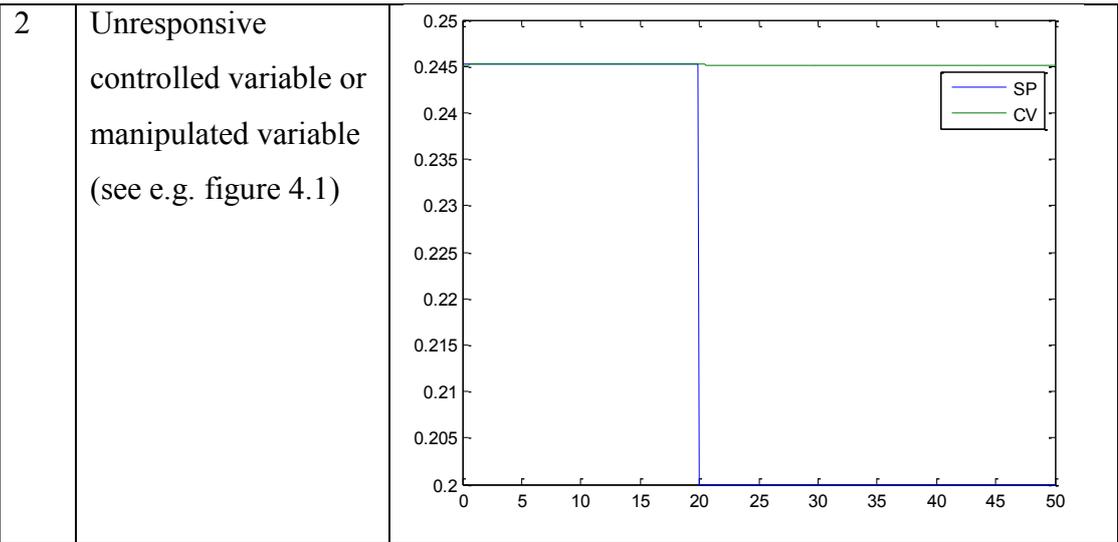
5.4 Preliminary diagnostic questions window

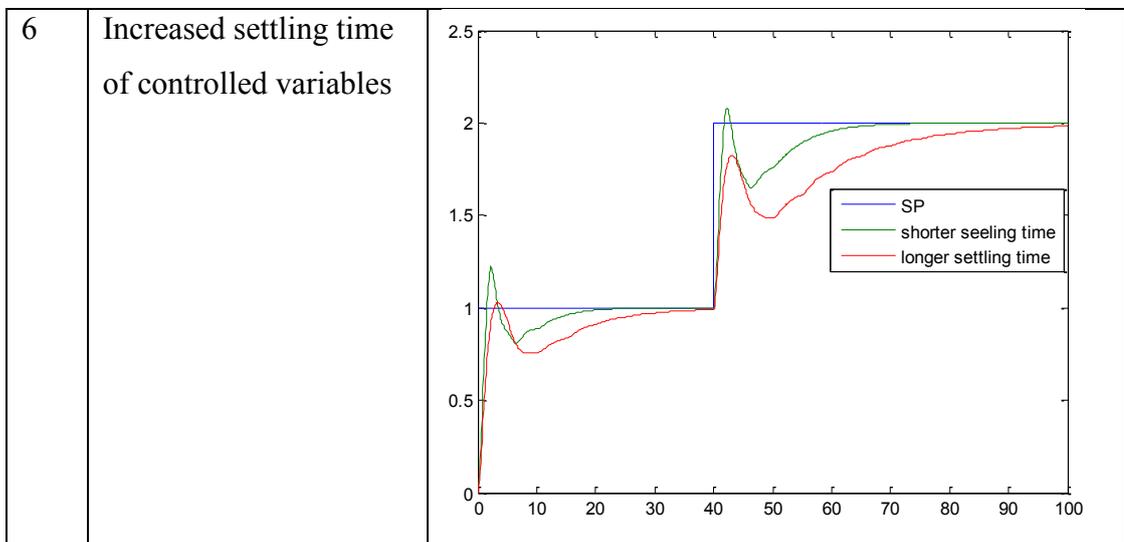
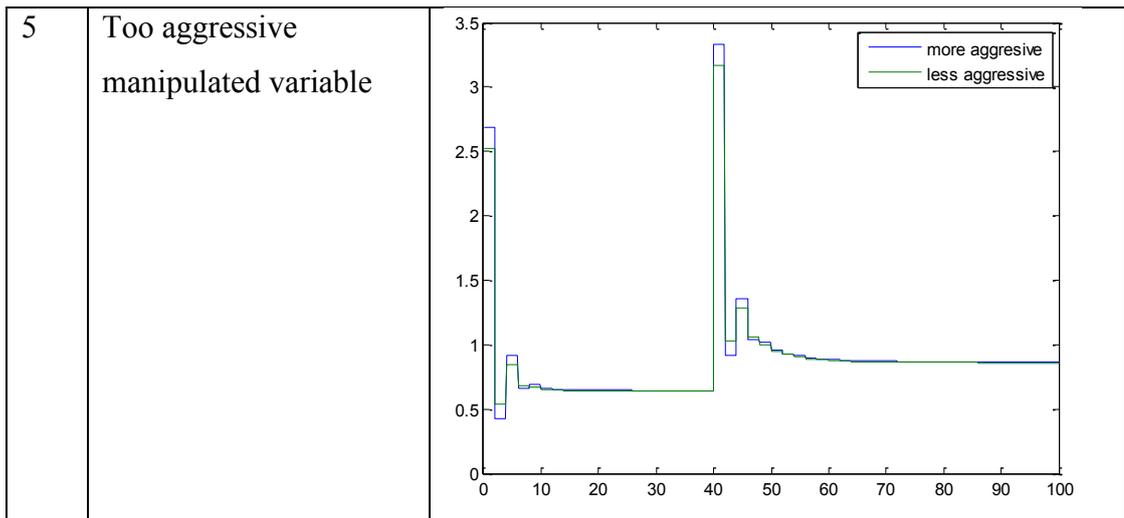
The window displays a list of diagnostic questions, each of which relates to a different type of abnormality that might manifest itself in a real MPC situation. The diagnostic questions displayed relate to the types of abnormality listed in Section 5.2. The questions are arranged in a form similar to table 5.1, where graphical sketches of the type of abnormality that the questions refer to are arranged in a column next to the questions. The operator is expected to select that type which best describes the observed abnormality in the performance trends. The type chosen by the operator then opens a different suspected faults window.

Note that table 5.1 below is split into six parts for ease of presentation

Table 5.1 Preliminary diagnostic questions

S/N	Option Questions	Sample trend displayed
1	Offsets in controlled variables (see e.g. figure 4.8)	





5.5 Suspected Faults Window

This window displays a list of causes that are likely to be responsible for the type of abnormality chosen by the operator in the diagnostic questions window in table 5.1. The list is arranged in order of diagnostic complexity; beginning with the one that might require the most basic diagnostic procedure, to the one that is likely to require the most engaging diagnostic procedure. The operators might choose to investigate in the order specified, or might jump around the list if new information leads them to do so. The list for each type of observed comparison trend is shown in table 5.2.

Table 5.2: Lists of causes associated with each underlying abnormality

S/N	Type	List displayed
1	Unresponsive controlled variable	<ul style="list-style-type: none"> • Inappropriate MPC tuning • Inappropriate variables selection • Model/plant mismatch • MPC design • Sensor/Actuator degradation
2	Non-oscillatory offsets in controlled variables	<ul style="list-style-type: none"> • Inappropriate MPC tuning • Process drift due to unmeasured disturbance • Model/plant mismatch • Inappropriate variables selection • MPC design
3	Unresponsive manipulated variable	<ul style="list-style-type: none"> • Inappropriate MPC tuning • Inappropriate variables selection • Model/plant mismatch • MPC design
4	Saturation of manipulated variables	<ul style="list-style-type: none"> • Inappropriate MPC tuning • Improper constraints specifications • Presence of unmeasured disturbance
5	Oscillations in controlled and/or manipulated variables	<ul style="list-style-type: none"> • MPC tuning • PID degradation • Sensor/actuator degradation • Model/plant mismatch
6	Too aggressive manipulated variable	<ul style="list-style-type: none"> • MPC tuning • Inappropriate variables selection
7	Output and/or input constraints violation	<ul style="list-style-type: none"> • MPC tuning • Improper constraints specifications • Inappropriate variables selection
8	Increased settling time of controlled variables	<ul style="list-style-type: none"> • MPC tuning • Presence of unmeasured disturbance • Model/plant mismatch

5.6 The Background Information Group

When a cause associated with an abnormality is chosen (table 5.2), the operator is lead to the symptoms investigation window (fig.5.1), where a systematic examination of the symptom is carried out. In carrying out the investigation of an underlying symptom, the operator might be guided by information contained in a number of windows collectively referred to as background information windows. These windows include the transfer function matrix window, the virtual plant without MPC window, the virtual plant with MPC window and the RGA and RWA window. The operator is expected to be able to navigate to and from any of these windows in search of information and guidance. The descriptions of the roles of each of the background information windows are given below.

5.6.1 Virtual plant Without MPC Window

This window houses the linear model of the plant (which may be in state space, difference equation or transfer function form). It should include provisions to apply different input sources (especially step deviation signals) to the virtual plant, and to display the step responses from such signals (figure 5.8). The step responses are of the type shown in many instances in Chapter Three. A sample of step response plots that may be displayed in this window is shown in figure 5.9. With good understanding of system dynamics as portrayed in the step responses, the operator may be guided in making appropriate selection of the variables used in the MPC, especially when operator wishes to consider alternatives. A simple example to illustrate the type of information that may be derived from system dynamics is given below.

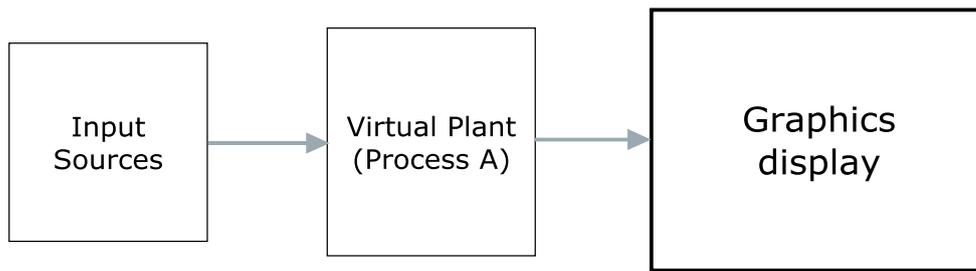


Figure 5.8: Virtual plant without MPC setup

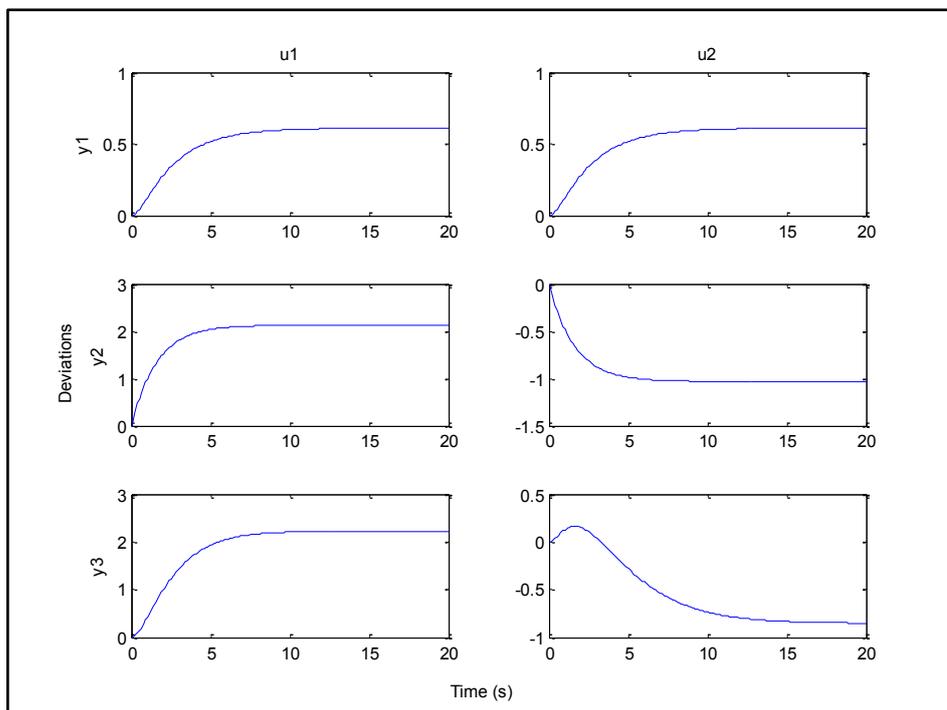


Figure 5.9: Sample step response plots from Virtual plant without MPC

The operator can investigate the appropriateness of a controlled and manipulated variable selection. The step response plots can also be used to check for suspected cases involving model/plant mismatch due to process drift.

A simple example is considered here to elaborate on how the step response plots may be used. Consider a process A with step response plots given in figure 5.9 with another process B having the step response plots shown in figure 5.10 below (taken from the CSTR example of Section 3.3.2). Considering the step responses of the two processes,

set-point control on any two of the output variables in process A may be used with the two input variables. But in process B, y_2 should advisedly not be used with y_3 or y_4 . Also y_3 should not be used with y_4 . In process A, all the outputs dynamics due to step in one input are not all similar to the outputs dynamics due to step in the other input. But in process B the dynamics of outputs y_2 , y_3 and y_4 due to step in the first input is very similar to the dynamics of the same outputs due to step on the second input. This indicates a strong degree of coupling among the three outputs which may make independent set-point control difficult.

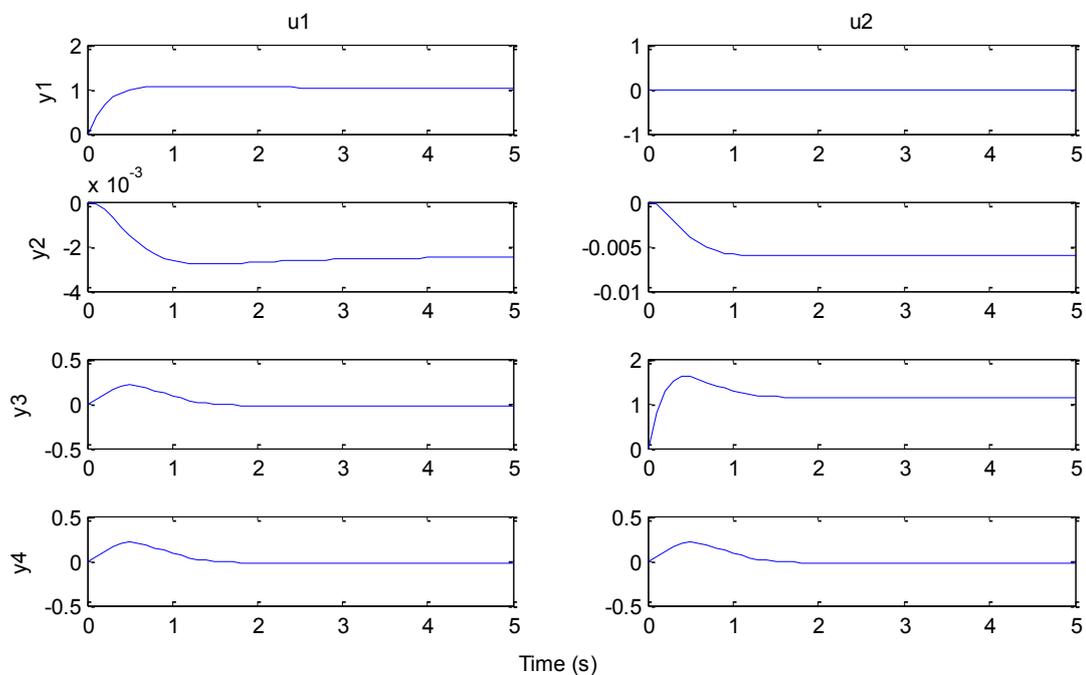


Figure 5.10: Step response plots of Process B

5.6.2 Transfer Function Matrix

An approximate low order transfer function matrix can be obtained by the method described in Appendix A. The transfer function matrix is used in conjunction with the step response plots of Section 5.3 to gain a deeper insight into the dynamics of the model and of the plant, which might help the operators to be more confident in their interactions with the plant and the MPC. An example of this is shown in table 5.3: here

the method described in Appendix A, was used to obtain the transfer function matrix with the linear model used to obtain step response plots of figure 5.9. Information like steady state gains and settling times would be transparent to the operator.

Table 5.3: Approximate transfer function matrix for process A

	u1	u2
y1	$\frac{0.61}{1.69s^2 + 2.86s + 1}$	$\frac{0.61}{1.69s^2 + 2.86s + 1}$
y2	$\frac{2.13}{1.5s + 1}$	$\frac{1.032}{1.5s + 1}$
y3	$\frac{2.21}{1.96s^2 + 2.86s + 1}$	$\frac{0.85(1 - s)}{7.29s^2 + 4.86s + 1}$

5.6.3 Steady state gain, RGA and RWA Window

The RGA and RWA calculations window displays the RGA and RWA for different combinations of inputs and outputs based on data of linear model steady state gains. The RGA can be used by the operators, in conjunction with the step response plots and the transfer function matrix, to make informed choices about merits and demerits of alternative MPC configurations that may be considered. The RWA is used to assist the operator in deciding the sensible starting point for controlled variable weight gain selection when tuning for a different configuration is required. The steady state gain, RGA and RWA for the virtual plant whose step response is given in figure 5.11 are shown below.

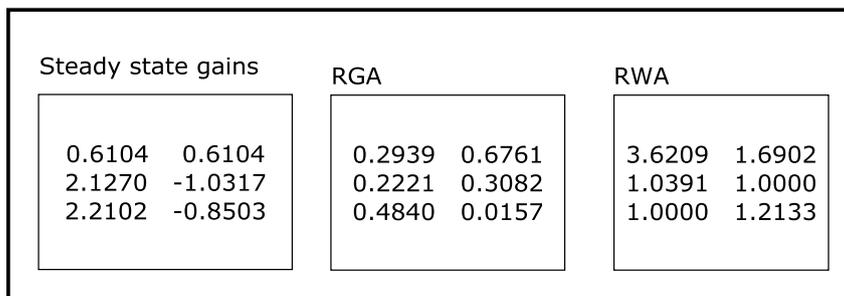


Figure 5.11: Display for steady state gains, RGA and RWA

In figure 5.11 for example, the operator may observe that for best set-point control, the RGA values suggest that outputs y_1 and y_3 are the pair more appropriate as controlled variables because their RGA values have highest numerical values, and are positive, indicating that the two outputs are together more affected by the two inputs than any other pair. With y_1 and y_3 chosen as controlled variables, their output weights should be closely equal, because according to the RWA values, the weight of y_1 can be from 1.6902 to 3.6209 and that for y_3 can be from 1.0000 to 1.2133, indicating that the numerical difference between the two output variables is not very high. This is unless the operators require preferential set-point tracking. This idea if RWA was used in the diagnosis of case 1 in Section 4.3.

5.6.4 Virtual plant With MPC Window

Here installed the virtual plant is a linear model of the real process under MPC. Both the linear model and MPC are similar to the one on the real plant (figure 5.12). The idea is that within the context of MPC, the virtual plant with MPC is normally expected to give better performance compared with the real plant with the same MPC. With this the operators can simulate different scenarios to see what is possible. This is particularly important in cases where the plant may be required to be operated at a new operating point, because certain MPC degradations can be ameliorated by retuning the MPC. The virtual plant with MPC provides a means of experimenting with proposed tuning strategies prior to implementation on the real plant. The virtual plant may also be used to experiment with different MPC configurations.

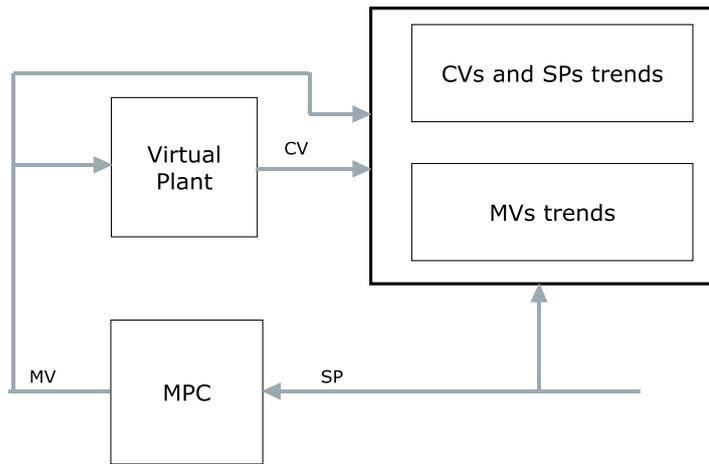


Figure 5.12: Virtual Plant with MPC setup

For example the virtual plant with MPC might be used to test the extent to which a current MPC might cope, if plant operation was to be transferred to a new operating point far removed from the current one. The idea is that if the MPC cannot cope with a required control action on the virtual plant, then it would be even more difficult on the real plant, since the virtual plant has no nonlinearities associated with the real plant, and the setup is devoid of any model/plant mismatch. This was demonstrated in the case study of Section 4.3.2 before.

5.7 MPC Investigation

The form of this window would depend on the outcome from the suspected faults window and hence from table 5.2. This window directs the operator to a number of investigative windows, each devoted to suggesting a number of actions, which may lead the operator to confirm or reject a particular reason for degradation. If a suspicion is confirmed then the remedial action that is required is likely to be obvious to the operator.

The procedures for investigating each suspected fault are given below. Some of the diagnostic procedures require an understanding of the system dynamics revealed by

RGA, RWA, transfer function matrix and step response plots, so it is assumed that the operators would have had prior access to these perhaps as part of a training programme.

5.7.1 Investigating inappropriate variables selection

If inappropriate variable selection is part of the list of causes associated with an abnormality (table 5.2) and the operator selects it, the Inappropriate Variables Selection window would then open. In it the operator would be asked to investigate along the lines of the following script. An example of this process was described in section 4.5 previously.

- 1) Simulate the virtual plant with similar MPC parameters, same variables and same operating conditions as in the real plant
- 2) Compare the trends from the virtual simulation with the real one. If the trends are similar, it is an indication that the initial assumption may be correct, and you should proceed to the next step.
- 3) Briefly navigate to the steady state gain, RGA and RWA window and the transfer function window. Using knowledge of the linear model dynamics (step response plots and transfer function matrix and RGA), examine if you can find more appropriate CVs for the given MV (or vice versa).
- 4) If other appropriate variables are discovered in step (3), experiment with these in the virtual plant with MPC window.
- 5) If the performance of the virtual plant in (4) improves as a result of the new variables, consider implementing the same on the real plant
- 6) If implementing the new variable combination in (5) on the real plant improves the performance of the MPC, then the source of the MPC degradation has been confirmed.

5.7.2 Investigating model/plant mismatch

If model/plant mismatch is part of the list of causes associated with an abnormality (table 5.2) and the operator selects it, the Model/plant Mismatch window would then open. In it the operator would be asked to investigate along the lines of the following script.

- 1) Run the plant in open loop (without MPC)
- 2) Apply step signals to each of the inputs in turn and record the controlled variables step responses.
- 3) Compare the trends of the step responses from (2) for each input-output pair with its counterpart in the virtual plant without MPC window.
- 4) If there are any remarkable differences in terms of the steady state gains, settling time or nature of response (e.g. whether inverse or non-inverse response), then model/plant mismatch is a suspect.
- 5) Contact the experts about this finding and request for further investigation.

5.7.3 Investigating improper constraints specifications

If improper constraints specification is part of the list of causes associated with an abnormality (table 5.2) and the operator selects it, the Improper Constraints Specifications window would then open. In it the operator would be asked to investigate along the lines of the following script.

- 1) Simulate the virtual plant with similar MPC parameters, same variables and same operating conditions as in the real plant.
- 2) Compare the trends from the virtual simulation with the real one. If the trends are similar, it is an indication that the initial assumption may be correct, and you should proceed with the next step.

- 3) While running the virtual plant with MPC, adjust the manipulated variables and the controlled variables constraints, and observe if there is an improvement in performance of the virtual plant.
- 4) If performance improvement is observed in (3), then check to ensure that the concerned valve can operate within the range (in the case of inputs) or that safety guidelines are not violated (in case of outputs).
- 5) Adjust the constraints of the real plant to reflect the settings for the virtual plant.
- 6) If the adjustment improves performance then the source of the MPC degradation has been confirmed.

5.7.4 Investigating PID degradation

Investigation of the performance of PID controllers has been covered extensively in Desborough and Harris (1992), Harris et al. (1996) and Thornhill et al. (1999), among many other publications. The procedure described here serves only to complement.

If PID degradation is part of the list of causes associated with an abnormality (table 5.2) and the operator selects it, PID Degradation window would then open. In it the operator would be asked to investigate along the lines of the following script.

- 1) Obtain the plots of actual sv_o and cv_o against time and mv_o against ov_o (open loop). (figure 2.4). While in open loop the data for mv_o and ov_o may be obtained by applying excitation signals to the inputs (step, PRBS).
- 2) See if cv_o does not track sv_o and exhibits an oscillatory trend for each of the plots of sv_o and cv_o against time. Also see if each mv_o against ov_o has not deviated from its counterpart in the reference graphical window. Proceed to step 3.

- 3) If cv_o is oscillatory in the plots of sv_o and cv_o against time and the plot of mv_o against ov_o has not deviated from its counterpart in the reference graphical performance window, then the suspicion is confirmed.
- 4) Embark on a recovery procedure involving the retuning of PID controllers.

5.7.5 Investigating actuator degradation

If actuator degradation is part of the list of causes associated with an abnormality (table 5.2) and the operator selects it, the Sensor/actuator Degradation window would then open. In it the operator would be asked to investigate along the lines of the following script.

- 1) Obtain the plots of actual mv against ov (under MPC operation) and mv_o against ov_o (open loop). (figure 2.4). While in open loop, the data for generating the plot of mv_o against ov_o may be obtained by applying excitation signals to the inputs (step, PRBS)
- 2) If any of the above plots are different from their counterparts in the reference graphical performance window, proceed to step 3.
- 3) See if any of the plots in (1) are similar to any of the common valve problems shown in figure 5.14 below.
- 4) If there is any match then the suspicion is confirmed.
- 5) Embark on a recovery procedure by repairing or replacing the faulty valve.

5.7.6 Investigating Poor MPC tuning

If poor MPC tuning is part of the list of causes associated with an abnormality (table 5.2) and the operator selects it, the poor MPC Tuning window would then open. In it the operator would be asked to investigate along the lines of the following script.

1. Simulate the virtual plant with similar MPC parameters, same variables and same operating conditions as in the real plant.

2. Compare the trends from the virtual simulation with the real one. If the trends are similar, it is an indication that the initial assumption may be correct, and you should proceed to the next step.
3. Adjust the MPC parameters (output and input weights, prediction and control horizon, sampling interval) of the virtual plant, noting that increasing the input weight increases the process settling time, and increasing the output weight reduces the process settling time. Also a smaller sampling interval leads to a smaller settling time. You may briefly navigate to the background information window and examine the RWA values as a guide to output weight selection.
4. If the performance of the virtual plant improves as a result of the tuning, consider implementing the same on the real plant.
5. If implementing the new MPC tuning on the real plant improves the performance of the MPC, then the source of the MPC degradation has been confirmed.

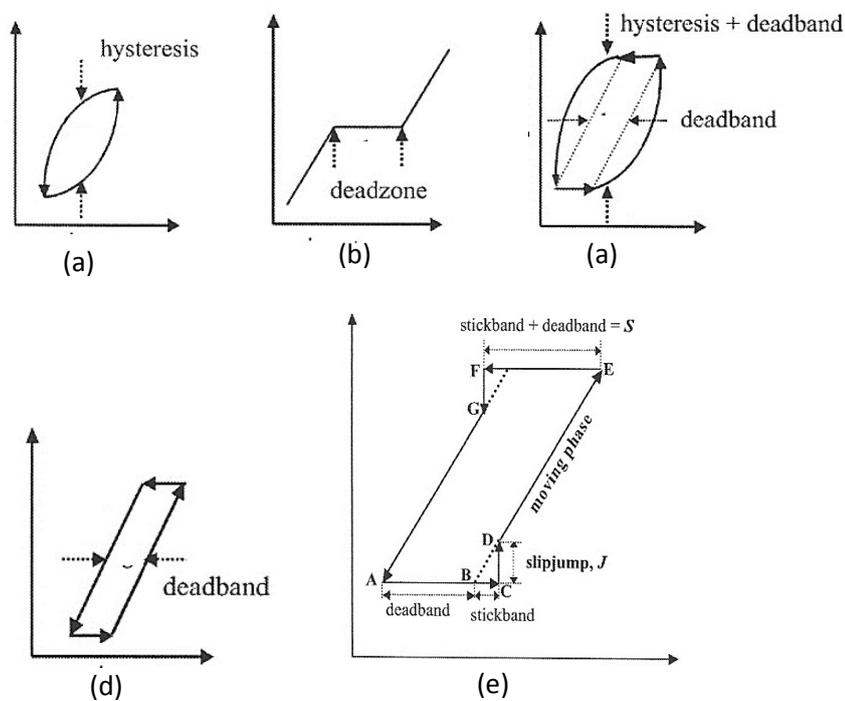


Figure 5.13: plot of mv against ov for common valve problems: (a) hysteresis, (b) deadzone, (c) hysteresis + deadband, (d) deadband (e) stiction. Adapted from (Shoukat Choudhury et al., 2005)

5.7.7 Investigating Poor MPC design

Investigation for poor MPC design is meant to be the last resort; after all other options have been exhausted. It involves using the guide contained in the background information windows to determine if an alternative design for the MPC may be recommended. It involves the following

1. Refer to the RGA, transfer function matrix and step response plots to ensure that the correct variables have been used;
2. Refer to the step response plots to see if there are input – output pairs that may require special attention like the presence of an integrating loop that may cause stability problems.

5.8 Summary

Chapter 5 discussed the development of the completely data-driven MPC maintenance tool. The maintenance tool described comprises of many assessment windows, grouped into five stages: trends comparison, diagnostic questions, suspected symptoms generation, symptoms investigation and MPC recovery. The chapter gave a description of the functions of the assessment windows in each group. It described the abnormalities that might occur when PMC degrade and provided detailed procedures of how the operator might use the maintenance tool to investigate the symptoms and recover the MPC to good performance.

Chapter Six

Conclusions and Recommendations

6.1 Progress to date

Model predictive control has found widespread use in the industries. Like traditional control systems, MPC performance must be regularly reviewed if it is to fulfil its potential. The need to maintain good performance has elicited control assessment research interests in many academicians and process control practitioners. While the studies have produced many methods and metrics by which poor performance may be detected, little has been written about how these methods and metrics are interpreted when MPC has actually degraded. Unlike traditional controllers where, when degradation occurs, the operator is usually able to understand why and initiate recovery procedures, the often reported situation in the case of MPC is for the operator to turn it off (Huang et al., 2000, Jiang et al., 2011). The main reason for this is that MPC technology and its performance assessment is still opaque to many operators.

The major reasons why many model predictive controller' performance deteriorate are also widely reported. These include process drift, unmeasured disturbances, poor tuning of lower level regulatory controllers, inappropriate constraints settings and poor design. A means of identifying tell-tale signs (i.e. symptoms) of what factors cause these symptoms to appear and how to recover from performance degradation related to them, without resulting to complex statistical analysis, will be invaluable to most operators, whose needs are more practical than theoretical.

This thesis contributes to this vision. Beginning with the notion that to proffer solution you must understand, Chapter 2 is devoted to reviewing the principles and algorithms of

MPC and its development and industrial applications over the years. The chapter also reviewed many of the control performance assessment methods, especially the ones pertaining to MPC, with an emphasis on those employing data-driven approaches in their analyses. The chapter discussed the role of the different internal model formats in the development of MPC algorithms, and why one format, the transfer function, offers a level of transparency to the operator. A brief discussion was given of how a linear internal model of a plant may be identified. The discussions on algorithms and model identification were especially important because the computer programs for implementing MPC in this thesis were written from scratch using MATLAB and SIMULINK.

Chapter 3 describes the results obtained when various MPC formulations were applied to three nonlinear process models. These then act as vehicles for case studies in Chapter 4. The three nonlinear process models are a non-isothermal Continuously Stirred Thermal Reactor (CSTR), an evaporator process and a Fluid Catalytic Cracking Unit (FCCU). Each model was selected because of its properties: the CSTR is integrating and is open loop unstable. The evaporator is marginally open loop stable, with integrating loops. The FCCU is very complex with many inputs and outputs, and is highly coupled. All the nonlinear models were simulated at or near their defined nominal operating points. Their individual properties were used to determine whether MPC was implemented directly or in a supervisory capacity. State space MPC was used throughout. Linear state space models, obtained by linearizing the nonlinear models about their operating points, were used as internal MPC models in the implementations. Approximate low order models in transfer function matrix format were also obtained from state space models, and used as internal models on the CSTR and the evaporator MPC implementations. The trends obtained from these implementations showed that

the low order models were good approximations of their full state space counterparts. The various simulations reveal details of systems dynamics, which might be useful for diagnosis if degradation occurs.

Eight cases studies that focus on various ways of degrading MPC performance are discussed in Chapter 4. Three used the CSTR, three the evaporator and two the FCCU. Each case study is preceded with a simulation, followed by a description of how an average operator might reason if the degradation were to manifest itself on a real plant. The assumed operator's perspective is followed by how an informed and technically minded operator might reason about the symptoms and the appropriate diagnostic procedure, taking cognisance of the dynamics of the system. Various forms of information about the dynamics assist in the investigation: process step response plots, the transfer function matrix, steady state gains and the relative gain array derived from it. A simple novel measure derived from the steady gains, called the relative weight array might also sometimes assist the operator in the investigations. In some instances virtual plants, which are linear replica of the nonlinear plant with similar MPC, are used in the investigations. Each case study ends by outlining the diagnostic steps that were taken towards investigating, isolating and recovering from the MPC degradation.

The knowledge gained from the MPC simulations in Chapter 3 and the case studies in Chapter 4 were used to develop an outline of a vision for the maintenance tool (Chapter 5). The tool is completely data-driven, aimed at helping the operator make sensible judgements about performance degradation, the form and direction of diagnosis and fault isolation, and possibly, the recovery procedure.

The maintenance tool comprises of many assessment windows, grouped into five stages as follows:

- trends comparison,
- diagnostic questions,
- suspected symptoms generation,
- symptoms investigation,
- MPC recovery.

The reference graphical performance (RGP) and the actual graphical performance (AGP) are the two windows in trends comparison. The RGP window, which remains static most of the time, displays trends of data obtained well before degradation is suspected, the underlying data would be generated and updated by technical operations personnel, who would re-evaluate its status regularly. The AGP window displays similar trends to the RGP, but now the trends are periodically updated during actual plant operation. If operators observe unusual features in the AGP, they are encouraged to move to the preliminary diagnostic questions window.

Preliminary diagnostic questions are displayed in one window. The window displays a list of diagnostic questions, each of which relates to a different type of abnormality that might manifest itself in a real a MPC situation. Here, the operator is expected to select that type of abnormality which best describes what is observed. The choice leads the operator to the suspected symptoms window.

This window displays a list of causes that are likely to be responsible for the type of abnormality chosen by the operator in the diagnostic questions. The operator chooses a cause, which leads to scripts that detail the systematic examination of each symptom listed. The symptoms that can be investigated in the symptoms investigation window are:

- inappropriate variables selection

- model/plant mismatch
- improper constraints specifications
- PID degradation
- sensor/actuator degradation
- poor MPC tuning
- poor MPC design

At this stage the operator can access windows collectively referred to as background information windows:

- virtual plant without MPC window
- virtual plant with MPC window
- transfer function matrix window
- steady state gain, RGA and RWA window

The operator may use the step response plots generated from the virtual plant without MPC window as a guide to the appropriateness of controlled and manipulated variables, especially when an operator wishes to consider alternatives. Step response plot information can also be used to check suspected cases involving model/plant mismatch due to process drift. The transfer function matrix, is used in conjunction with the step response plots to gain a deeper insight into the dynamics of the model and of the plant, which might help the operators to be more confident in their interactions with the plant and the MPC.

Steady state gain, RGA and RWA information may be used by the operator, in conjunction with the step response plots and the transfer function matrix, to make informed choices about the merits and demerits of alternative MPC configurations. The RWA is used to assist the operator in deciding the sensible starting point for controlled

variable weight gain selection when tuning for a different configuration is required. The RGA values help to guide the operator in associating controlled variables with manipulated variables, when alternatives exist. Both RWA and RGA are derived from steady state gain values.

The virtual plant with MPC window is a linear model of the real process under MPC similar to the one on the real plant. The operator may use this to simulate different scenarios; including experimenting with proposed tuning strategies prior implementation on the real plant, to see what is possible.

6.2. Critique

The maintenance tool development is meant to serve two purposes: to educate and to help to recover from degradation. Even if for some experts, some of the case studies used here may look simple or contrived, it cannot detract from its main purpose. Hopefully also, this tool will encourage new approaches to MPC maintenance: in particular an opportunity to inform and educate operators. Obviously this would benefit from keeping records of the symptoms, the degradation type and the diagnostic procedures, although issues of commercial confidentiality might limit this possibility.

As presented the maintenance tool is aimed at the operator who may not be expert in MPC. Some of the diagnosis required and the knowledge to diagnose problems may be viewed as being too advanced. The truth is that most plant operators now have sound technical education. Some of the operators at WRPC have degrees in engineering. The tool is presented as part of the learning required to become conversant with MPC.

Purposely great efforts have been made to avoid employing heavy statistical analysis in MPC diagnosis. This approach is supported by the fact that the reported MPC recovery procedures in industrial plants have been mainly about using knowledge revealed by

process trends and system dynamics. What was probably missing was the presentation of the diagnostic procedures in a way that the operators can relate the symptoms with the plant dynamics. This is partly what this study has been about.

6.3 Suggestions for further work

The development of the MPC maintenance tool still requires further research in order to address the wide range of scenarios that the operator may be faced with. Important outstanding issues are elaborated on below:

- 1) Not all MPC implementations in industries are set-point control. Many are instead designed to keep controlled variables within defined zone or region only, and they also include provision for ranking the controlled variables. For these the MPC designs usually include an additional objective function explicitly. The MPC programs used in this research do not have this provision, and as such could not adequately exploit the complexities of nonlinear plant like the FCCU. Future MPC programs should provide for this.
- 2) The case studies in chapter 4 did not address the simulation and manifestations relating to inappropriate feedforward control (measured disturbance) specification, which was reported by Huang et al. (2000) to be a source of degradation on an industrial MPC. The main difficulty was in obtaining or simulating a nonlinear process model in which a measured disturbance responds to changes in the manipulated variable, as opposed to the manipulated variable responding to changes in the measured disturbance variable. Future work should address this.
- 3) Dead times (time delays) are common features of most process industries. The nonlinear examples used in this research have no time delays. Improper specification of dead times of a process for the MPC application are believed to

cause oscillations and instabilities in the process. Future examples should demonstrate if and how input delays affect MPC performance.

- 4) The development of the application package and graphical user interface for the maintenance tool. The user interface should make the maintenance tool interactive and user friendly. Users should be able to induce defined degradations into the virtual world, observe the degradation trends, be given descriptions of the nature of degradation, and be guided through recovery procedures automatically.

References

- ABB. (no date). *Multivariable control for the process industries - 3dMPC* [Online]. Available:
[http://www05.abb.com/global/scot/scot296.nsf/veritydisplay/b788e3a9870e9a1ec1256a34007129d6/\\$file/3bus025043r0001_en_3dmpc_multivariable_controller.pdf](http://www05.abb.com/global/scot/scot296.nsf/veritydisplay/b788e3a9870e9a1ec1256a34007129d6/$file/3bus025043r0001_en_3dmpc_multivariable_controller.pdf) [Accessed May 2013].
- Agarwal, N., Huang, B. & Tamayo, E. C. (2007a). Assessing model prediction control (MPC) performance. 1. Probabilistic approach for constraint analysis. *Industrial & engineering chemistry research*, 46, 8101-8111.
- Agarwal, N., Huang, B. & Tamayo, E. C. (2007b). Assessing model prediction control (MPC) performance. 2. Bayesian approach for constraint tuning. *Industrial & engineering chemistry research*, 46, 8112-8119.
- Alghazzawi, A. & Lennox, B. (2009). Model predictive control monitoring using multivariate statistics. *Journal of process control*, 19, 314-327.
- Annus, P., Land, R., Min, M. & Ojarand, J. (2012). *Simple Signals for System Identification, Fourier Transform - Signal Processing, Dr Salih Salih (Ed.)* [Online]. InTech. Available: <http://www.intechopen.com/books/fourier-transform-signal-processing/system-identification-simplified> [Accessed June 2013].
- ASPENTECH. (2013). *Aspen DMCplus* [Online]. Available:
<http://www.aspentech.com/products/aspen-dmcpus/> [Accessed May 2013].
- Badwe, A. S., Gudi, R. D., Patwardhan, R. S., Shah, S. L. & Patwardhan, S. C. (2009). Detection of model-plant mismatch in MPC applications. *Journal of process control*, 19, 1305-1313.
- Bemporad, A., Morari, M. & Ricker, N. L. (2013). *Model Predictive Control Toolbox User's Guide*. The Mathworks, Inc.

- Bezergianni, S. & Georgakis, C. (2003). Evaluation of controller performance—use of models derived by subspace identification. *International Journal of Adaptive Control and Signal Processing*, 17, 527-552.
- Borrelli, F., Baotić, M., Bemporad, A. & Morari, M. (2005). Dynamic programming for constrained optimal control of discrete-time linear hybrid systems. *Automatica*, 41, 1709-1721.
- Bristol, E. (1966). On a new measure of interaction for multivariable process control. *Automatic Control, IEEE Transactions on*, 11, 133-134.
- Camacho, E. (1993). Constrained generalized predictive control. *Automatic Control, IEEE Transactions on*, 38, 327-332.
- Camacho, E. F. & Bordons, C. (2004). *Model predictive control*, Springer London.
- Chang, J.-W. & Yu, C.-C. (1990). The relative gain for non-square multivariable systems. *Chemical Engineering Science*, 45, 1309-1323.
- Charos, G. N., Arkun, Y. & Taylor, R. (1991). Model predictive control of an industrial lime kiln. *Tappi Journal*, 74, 203-211.
- Clarke, D. W., Mohtadi, C. & Tuffs, P. (1987). Generalized predictive control—Part I. The basic algorithm. *Automatica*, 23, 137-148.
- Cutler, C., Morshedi, A. & Haydel, J. An industrial perspective on advanced control. AICHE annual meeting, 1983.
- Cutler, C. R. & Ramaker, B. Dynamic matrix control—a computer control algorithm. Proceedings of the joint automatic control conference, 1980. University of Michigan Ann Arbor, MI, Wp5-B.
- Darby, M. L. & Nikolaou, M. (2012). MPC: Current practice and challenges. *Control Engineering Practice*, 20, 328-342.
- Desborough, L. & Harris, T. (1992). Performance assessment measures for univariate feedback control. *Canadian Journal of Chemical Engineering*, 70, 1186-1197.

- Ding, S., Zhang, P., Naik, A., Ding, E. & Huang, B. (2009). Subspace method aided data-driven design of fault detection and isolation systems. *Journal of process control*, 19, 1496-1510.
- Ender, D. (1993). Process control performance: Not as good as you think. *Control Engineering*, 40, 2812-2817.
- Eykhoff, P. (1974). *System identification: parameter and state estimation*, Wiley-Interscience London.
- Favoreel, W., De Moor, B. & Van Overschee, P. (2000). Subspace state space system identification for industrial processes. *Journal of process control*, 10, 149-155.
- Froisy, J. & Matsko, T. IDCOM-M application to the Shell fundamental control problem. AICHE annual meeting, 1990.
- Froisy, J. B. (1994). Model predictive control: Past, present and future. *ISA Transactions*, 33, 235-243.
- Froisy, J. B. (2006). Model predictive control—building a bridge between theory and practice. *Computers & Chemical Engineering*, 30, 1426-1435.
- Gabriele, P., Michele, B. & Lucac, A. D. (2013). Application of a Method to Diagnose the Source of Performance Degradation in MPC Systems. *Chemical Engineering Transactions*, 32, 118-1194.
- Gao, J., Patwardhan, R., Akamatsu, K., Hashimoto, Y., Emoto, G., Shah, S. L. & Huang, B. (2003). Performance evaluation of two industrial MPC controllers. *Control engineering practice*, 11, 1371-1387.
- Garcia, C. E., Prett, D. M. & Morari, M. (1989). Model predictive control: theory and practice—a survey. *Automatica*, 25, 335-348.
- Goodwin, G. C. & Payne, R. L. (1977). *Dynamic system identification: experiment design and data analysis*.

- Grosdidier, P., Froisy, B. & Hammann, M. The idcom-m controller. Proceedings of the 1988 IFAC workshop on model based process control, 1988. Pergamon Press. Oxford, 31-36.
- Harris, T. (1989). Assessment of closed-loop performance. *Canadian Journal of Chemical Engineering*, 67, 856-861.
- Harris, T., Boudreau, F. & Macgregor, J. F. (1996). Performance assessment of multivariate feedback controllers. *Automatica*, 32, 1505-1518.
- Harris, T. J. & Seppala, C. T. Recent developments in controller performance monitoring and assessment techniques. AIChE Symposium Series, 2002. New York; American Institute of Chemical Engineers; 1998, 208-222.
- Harris, T. J., Seppala, C. T. & Desborough, L. D. (1999). A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *Journal of process control*, 9, 1-17.
- Harrison, C. A. & Qin, S. J. (2009). Minimum variance performance map for constrained model predictive control. *Journal of process control*, 19, 1199-1204.
- Hokanson, D. A. & Gerstle, J. G. (1993). Dynamic matrix control multivariable controllers. *Practical Distillation Control*. Springer.
- HONEYWELL. (no date). *Profit Controller* [Online]. Available: <https://www.honeywellprocess.com/en-US/explore/products/advanced-applications/operations-excellence/process-optimization/Pages/profit-controller.aspx> [Accessed May 2013].
- Huang, B. & Kadali, R. (2008). Model Predictive Control: Conventional Approach. *Dynamic Modeling, Predictive Control and Performance Monitoring*. Springer.
- Huang, B., Kadali, R., Zhao, X., Tamayo, E. C. & Hanafi, A. (2000). An investigation into the poor performance of a model predictive control system on an industrial CGO coker. *Control engineering practice*, 8, 619-631.
- Huang, B. & Shah, S. L. (1999). *Performance assessment of control loops: theory and applications*, Springer.

- Huang, B., Shah, S. L. & Kwok, K. How good is your controller? Application of control loop performance assessment to MIMO processes. Proceedings of 13th IFAC world congress, 1996 San Francisco. 229-234.
- Huang, B., Shah, S. L. & Kwok, K. (1997). Performance assessment of multivariable processes. *Automatica*, 33, 1175-1183.
- Hugo, A. (2000). *Limitations of model predictive controllers* [Online]. Hydrocarbon Processing. Available:
<http://www.controlartsinc.com/Support/Articles/LimitationsOfMBC.pdf>
[Accessed November 2013].
- Jelali, M. (2006). An overview of control performance assessment technology and industrial applications. *Control engineering practice*, 14, 441-466.
- Jiang, H., Shah, S. L., Huang, B., Wilson, B., Patwardhan, R. & Szeto, F. (2011). Model analysis and performance analysis of two industrial MPCs. *Control engineering practice*.
- Julien, R. H., Foley, M. W. & Cluett, W. R. (2004). Performance assessment using a model predictive control benchmark. *Journal of process control*, 14, 441-456.
- Kadali, R., Huang, B. & Tamayo, E. A case study on performance analysis and trouble shooting of an industrial model predictive control system. American Control Conference, 1999. Proceedings of the 1999, 1999. IEEE, 642-646.
- Katayama, T. (2005). *Subspace methods for system identification*, Springer.
- Ko, B. S. & Edgar, T. F. Performance assessment of multivariable feedback control systems. Proceedings of the American control conference, 2000 Chicago, Illinois.
- Ko, B. S. & Edgar, T. F. (2001). Performance assessment of constrained model predictive control systems. *AIChE Journal*, 47, 1363-1371.
- Kothare, M. V., Balakrishnan, V. & Morari, M. (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32, 1361-1379.

- Kozub, D. J. & Garcia, C. E. Monitoring and diagnosis of automated controllers in the chemical process industries. AICHE annual meeting, 1993.
- Lee, E. B. & Markus, L. (1967). Foundations of optimal control theory. DTIC Document.
- Lee, K. H., Huang, B. & Tamayo, E. C. (2008). Sensitivity analysis for selective constraint and variability tuning in performance assessment of industrial MPC. *Control engineering practice*, 16, 1195-1215.
- Lee, S., Yeom, S. & Lee, K. S. (2004). Methods for performance monitoring and diagnosis of multivariable model-based control systems. *Korean Journal of Chemical Engineering*, 21, 575-581.
- Lennox, B. (2005). Integrating fault detection and isolation with model predictive control. *International Journal of Adaptive Control and Signal Processing*, 19, 199-212.
- Li, S., Lim, Y. K. & Fisher, D. G. (1989). A state space formulation for model predictive control. *AICHE Journal*, 35, 241-249.
- Ljung, L. (1999). *System identification*, Wiley Online Library.
- Ljung, L. & Mckelvey, T. (1996). Subspace identification from closed loop data. *Signal Processing*, 52, 209-215.
- Loquasto III, F. & Seborg, D. E. Model predictive controller monitoring based on pattern classification and PCA. American Control Conference, 2003. Proceedings of the 2003, 2003. IEEE, 1968-1973.
- Luenberger, D. G. (1984). *Linear and nonlinear programming*, Addison Wesley Publishing Company.
- Luyben, W. L. (1989). *Process Modelling, Simulation and Control for Chemical Engineers*, McGraw-Hill
- Maciejowski, J. M. (2002). *Predictive control: with constraints*, Pearson education.

- Marquis, P. & Broustail, J. SMOC, a bridge between state space and model predictive controllers: application to the automation of a hydrotreating unit. Proceedings of the IFAC workshop on model based process control, 1988. 37-43.
- Mcfarlane, R. C., Reineman, R. C., Bartee, J. F. & Georgakis, C. (1993). Dynamic simulator for a model IV fluid catalytic cracking unit. *Computers & Chemical Engineering*, 17, 275-300.
- Muske, K. R. & Rawlings, J. B. (1993). Model predictive control with linear models. *AIChE Journal*, 39, 262-287.
- Newell, R. B. & Lee, P. L. (1989). *Applied process control: a case study*, Prentice-Hall.
- Patwardhan, K. & Jay, H. L. (1997). Diagnostic tools for multivariable model-based control systems. *Ind. Eng. Chem. Res.*, 36, 2725-2738.
- Patwardhan, P. & Shah, S. L. (2002). Performance diagnosis of model-based controllers. *Journal of Process Control*, 12, 413-427.
- Patwardhan, R. S., Shah, S. L., Emoto, G. & Fujji, H. (1998). Performance analysis of model-based predictive controllers: An industrial case study. *AIChE annual meeting*. Miami.
- Patwardhan, R. S., Shah, S. L. & Qi, K. Z. (2002). Assessing the performance of model predictive controllers. *The Canadian Journal of Chemical Engineering*, 80, 954-966.
- Peterka, V. (1984). Predictor-based self-tuning control. *Automatica*, 20, 39-50.
- Propoi, A. (1963). Use of linear programming methods for synthesizing sampled-data automatic systems. *Automation and Remote Control*, 24, 837-844.
- Qin, S. J. (1998). Control performance monitoring — a review and assessment. *Computers & Chemical Engineering*, 23, 173-186.
- Qin, S. J. & Badgwell, T. A. An overview of industrial model predictive control technology. *AIChE Symposium Series*, 1997. Citeseer, 232-256.

- Qin, S. J. & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control engineering practice*, 11, 733-764.
- Qin, S. J. & Yu, J. (2007). Recent developments in multivariable controller performance monitoring. *Journal of process control*, 17, 221-227.
- Rawlings, J. B. (2000). Tutorial overview of model predictive control. *Control Systems, IEEE*, 20, 38-52.
- Rawlings, J. B. & Muske, K. R. (1993). The stability of constrained receding horizon control. *Automatic Control, IEEE Transactions on*, 38, 1512-1516.
- Richalet, J., Rault, A., Testud, J. & Papon, J. Algorithmic control of industrial processes. Proceedings of the 4th IFAC symposium on identification and system parameter estimation, 1976. URSS September, Tbilisi, 1119-1167.
- Richalet, J., Rault, A., Testud, J. & Papon, J. (1978). Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14, 413-428.
- Ricker, N. L. (1991). Model predictive control: State of the art. *Proc. Chemical Process Control IV*, 271-296.
- Schäfer, J. & Cinar, A. (2004). Multivariable MPC system performance assessment, monitoring, and diagnosis. *Journal of process control*, 14, 113-129.
- Seborg, D. E., Mellichamp, D. A., Edgar, T. F. & Doyle III, F. J. (2010). *Process dynamics and control*, Wiley.
- Shell (2004). *Advancing Process Control* [Online]. Available: <http://s07.static-shell.com/content/dam/shell/static/globalsolutions/downloads/innovation/knowledge-centre/advancing.pdf> [Accessed May 2013].
- Shoukat Choudhury, M. A. A., Thornhill, N. F. & Shah, S. L. (2005). Modelling valve stiction. *Control engineering practice*, 13, 641-658.
- Söderström, T. & Stoica, P. (1988). *System identification*, Prentice-Hall, Inc.
- Tatjewski, P. (2010). Supervisory predictive control and on-line set-point optimization.

- Thornhill, N., Oettinger, M. & Fedenczuk, P. (1999). Refinery-wide control loop performance assessment. *Journal of process control*, 9, 109-124.
- Tyler, M. L. & Morari, M. (1996). Performance monitoring of control systems using likelihood methods. *Automatica*, 32, 1145-1162.
- Van Den Hof, P. (1998). Closed-loop issues in system identification. *Annual Reviews in Control*, 22, 173-186.
- Van Overschee, P. & De Moor, B. Closed loop subspace system identification. Decision and Control, 1997., Proceedings of the 36th IEEE Conference on, 1997. IEEE, 1848-1853.
- Wang, L. (2009). *Model Predictive Control System Design and Implementation Using MATLAB®*, Springer Verlag.
- Wismer, D. A. & Chattergy, R. (1978). *Introduction to nonlinear optimization: a problem solving approach*, North-Holland New York, Amsterdam.
- Xu, F., Huang, B. & Akande, S. (2007). Performance assessment of model predictive control for variability and constraint tuning. *Industrial & engineering chemistry research*, 46, 1208-1219.
- Yin, S., Ding, S. X., Haghani, A., Hao, H. & Zhang, P. (2012). A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of process control*.
- Yousfi, C. & Tournier, R. Steady state optimization inside model predictive control. American Control Conference, 1991, 1991. IEEE, 1866-1870.
- Yu, J. & Qin, S. J. (2008a). Statistical MIMO controller performance monitoring. Part I: Data-driven covariance benchmark. *Journal of process control*, 18, 277-296.
- Yu, J. & Qin, S. J. (2008b). Statistical MIMO controller performance monitoring. Part II: Performance diagnosis. *Journal of process control*, 18, 297-319.

Zanovello, R. & Budman, H. (1999). Model predictive control with soft constraints with application to lime kiln control. *Computers & Chemical Engineering*, 23, 791-806.

Zhang, Q. & Li, S. (2006). Performance Monitoring and Diagnosis of Multivariable Model Predictive Control Using Statistical Analysis. *Chinese Journal of Chemical Engineering*, 14, 207-215.

Appendix A

Low Order Approximation of Processes by Direct Method

The dynamic relationship between the input and output of many processes are usually defined in terms of measurable transient characteristics such as steady state gain, natural frequency, damping coefficients, time constant etc. For a multiple input multiple output system, a matrix of low order transfer functions can be used to present and explain these dynamics characteristics. This method of representing linear model of processes is still very popular in process industries because of the transparency that it offers.

The transfer function matrix describing a process is commonly identified from data obtained through excitation of the plant with step inputs. Advances in process identification have given rise to many identification methods that use different excitation signals (e.g. PRBS) and that present linear models in different compact formats (state space, difference equations). The approximate transfer function matrix of such models can be obtained, using simple techniques, as described in the following sections.

A1: The Direct Method

This method describes a simple, direct way of obtaining an approximate, strictly proper, first or second order transfer function for each input-output pair of a process. The method assumes that the step response data of the process exits (obtained either from the plant directly through step test, or from steps applied to an identified linear state space or difference equation of the process).

A1-1: First order transfer function approximation

The equation for the strictly proper, transfer function of a first order system is given as:

$$G(s) = \frac{Ke^{-T_d s}}{Ts+1} \quad \dots(A1)$$

where

K is the steady state gain of the system (calculated as the ratio of steady state deviation from equilibrium of the system response to the size of the step input)

T_d is the input delay of the system (deadtime, or time elapsed before system response after application of step input)

T is the time constant of the system (a measure of how fast the system approaches the steady state value)

The general shapes of a first order system are shown in figure A1 below

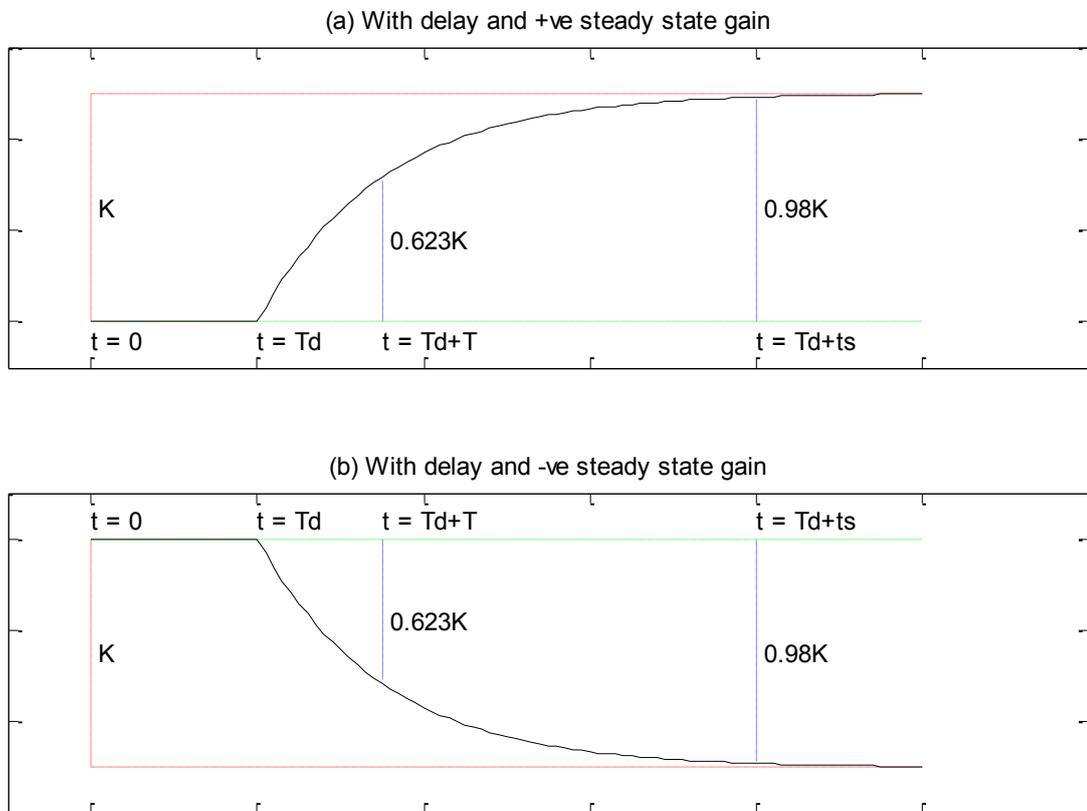


Figure A 1: general shape of a first order system

For a system that is subjected to step input, it is normally easy to read out the values of K and T_d from the resulting curve. T is much difficult to read, though it can be obtained as 62.3% of K

The simple direct method process for identifying first order system from step data is as follows:

- 1) Obtain a plot of the step response data. This serves as the original model m_0 (figure A.2)
- 2) Measure both K and T_d from the step response data
- 3) Substitute the values of K and T_d obtained in (2) into equation A.1 and with any arbitrary value of T . Superimpose the step response plot of this model on the one obtained in (1).
- 4) If the plot of the model from (3) is above m_0 , (that is m_1 in figure A.2), it indicates that the value of T used is smaller than that for the real model m_0 , so increase it. If the plot is below m_0 (that is m_2 in figure A.2), it indicates that the value if T used is bigger than that of the real model m_0 , so reduce it
- 5) Continue with the process of increasing and reducing T until the two plots coincide. The value of T at the point is the desired value, and the resulting model describes the dynamics of the system.

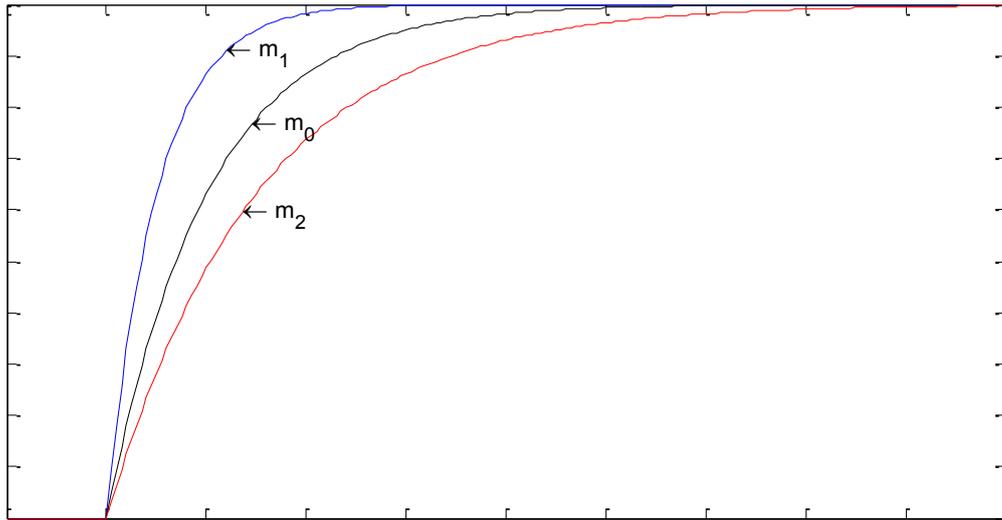


Figure A 2: First order systems with the same steady state gain, but different time constants

A1-2 Second order transfer function approximation

The equation for the transfer function of a non-minimum phase second order system is given as:

$$G(s) = \frac{K\omega_n^2 e^{-T_d s}}{s^2 + 2\omega_n \zeta s + \omega_n^2} = \frac{K e^{-T_d s}}{\tau^2 s^2 + 2\tau \zeta s + 1} \quad \dots(\text{A.2})$$

Where

ω_n is the natural frequency of the system

ζ is the damping coefficient of the system

$$\tau^2 = 1/\omega_n^2,$$

Both K and T_d are as defined before.

Depending on the numerical value of the damping coefficient ζ , second order system can exhibit different types of responses. The general shapes of a second order system are shown in figure A.3 below

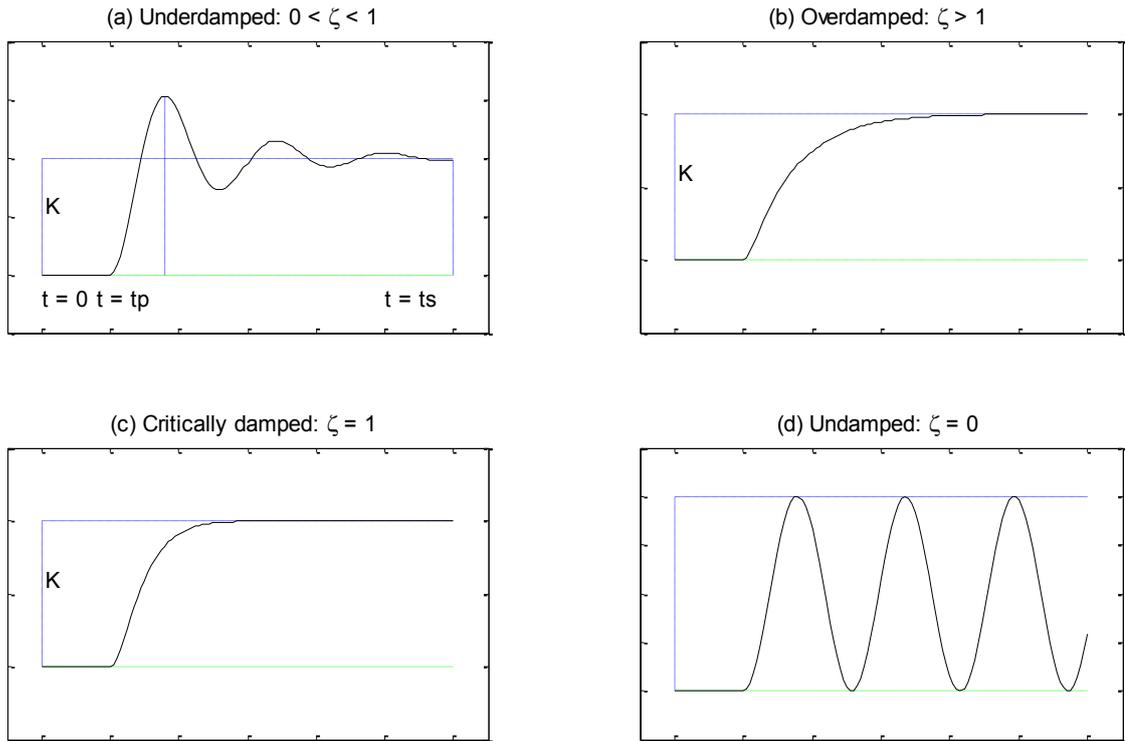


Figure A 3: Shapes of response of second order systems

The figure above shows that a second order system may exhibit dynamic responses similar to that of a first order system, as shown in the plots of figure A.3b and A.3c. Closed loop systems are usually underdamped, and the following figure shows how important parameters of such system such as settling time (t_s), response at maximum overshoot (y_{max}), and time maximum overshoot (t_p) may be calculated given ω_n and ζ . For an underdamped second order system's response (shown in figure A4), the values at key points of the response can be obtained from the equations below.

$$y_{max} = K \left(\frac{e^{-\zeta\pi}}{\sqrt{1-\zeta^2}} + 1 \right) \quad \dots(A3)$$

$$t_p = \frac{\pi}{\omega_n \sqrt{1-\zeta^2}} \quad \dots(A4)$$

$$t_s = \frac{4}{\zeta \omega_n} \quad \dots(A5)$$

where

y_{max} = maximum overshoot

t_p = time at maximum overshoot

t_s = settling time

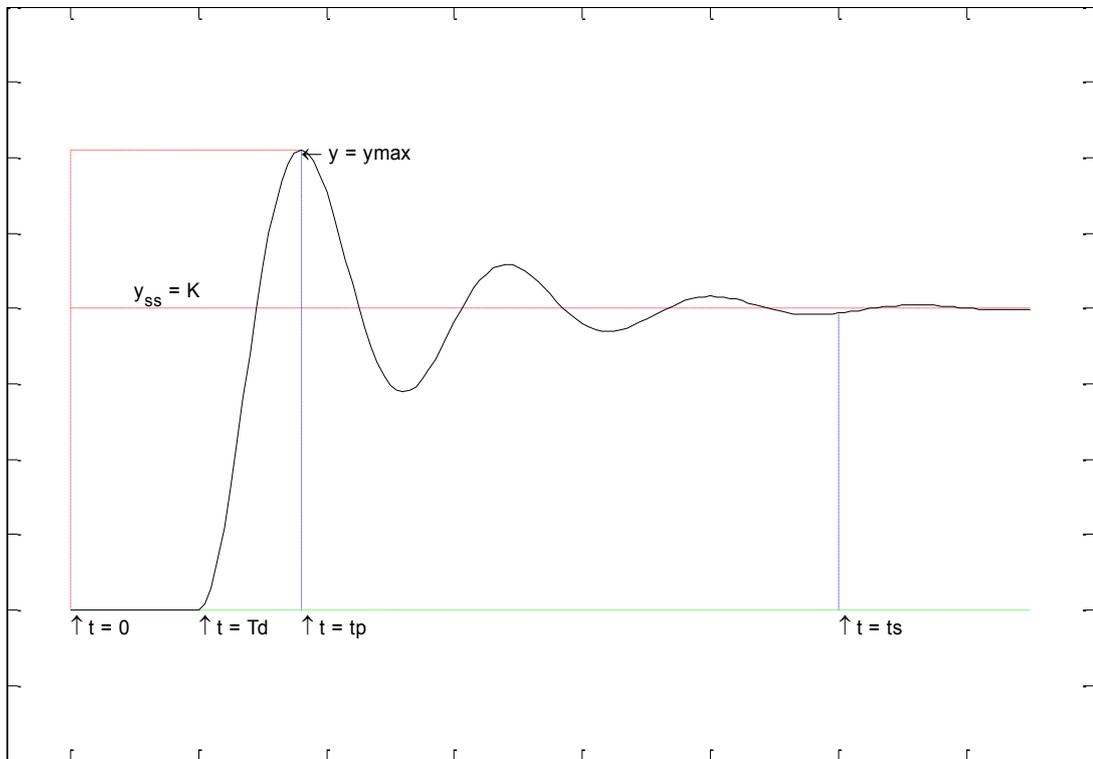


Figure A 4: Response of underdamped second order system showing key points

It can be seen from equation A.3 that the maximum overshoot depends on the damping coefficient ζ only. This is illustrated further in figure A5 below using three different systems. Underdamped second order systems with different natural frequencies ω_n ($1/\tau$) but with the same damping coefficient have the same peak overshoot. Their settling times differ.

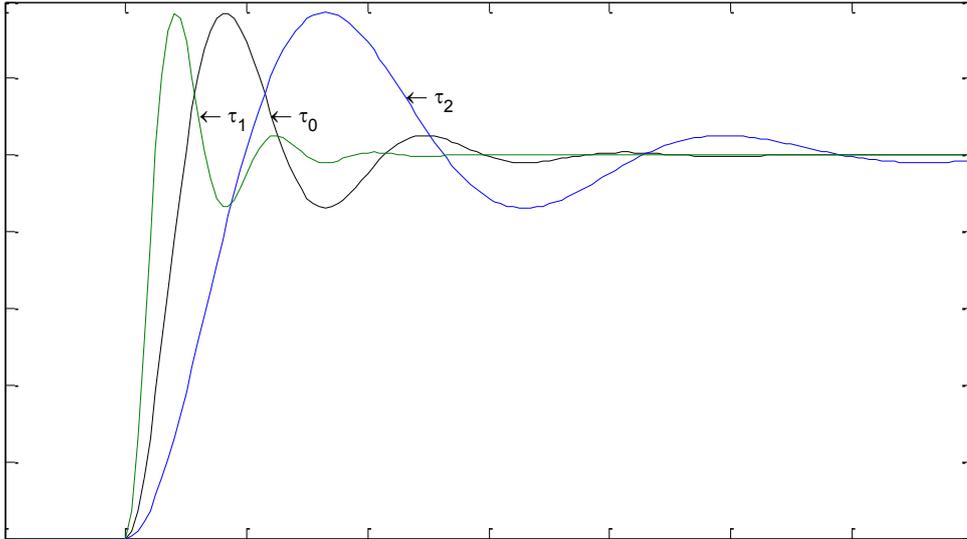


Figure A 5: Second order systems with the same damping coefficient but different natural frequencies

From the foregoing, given that the values K and T_d are known, the simple direct method process for identifying second order system from step data is as follows:

- 1) Obtain a plot of the step response data. This serves as the original model with τ_0 (figure A5)
- 2) Substitute the values of K and T_d obtained in (2) into equation A.2, using arbitrary values of ζ and ω_n ($1/\tau$).
- 3) Superimpose the step response plot of the model from (2) on the one obtained in (1).
- 4) Adjust the value of ζ in (2) until the peak overshoot of the plot from (3) matches that of the plot in (1). If the plot of the model from (3) is above ζ_0 , (that is ζ_1 in figure A6 below), it indicates that the value of ζ used is smaller than that for the real model ζ_0 , so increase it. If the plot is below ζ_0 (that is ζ_2 in figure A5), it indicates that the value of ζ used is bigger than that of the real model ζ_0 , so reduce it.

- 5) Continue with the process of increasing and reducing ζ until the two plots (real and model) have the same peak overshoot.
- 6) While maintaining this value of ζ , adjust ω_n ($1/\tau$) until the two plots coincide. High value of ω_n gives more oscillations before settling time (figure A5). The resulting value of ω_n is the desired value.

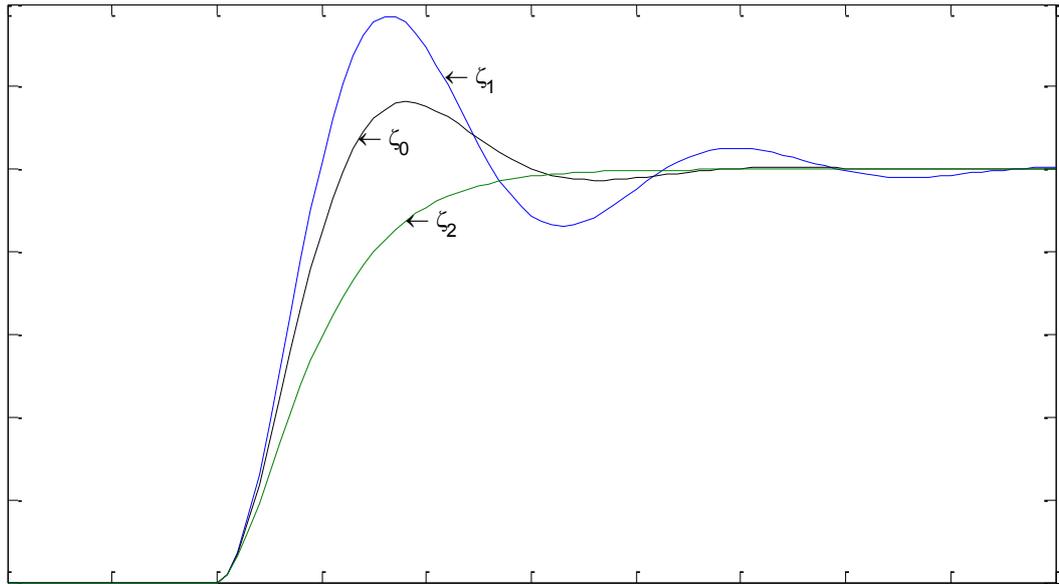


Figure A 6: Second order systems with the different damping coefficients but the same natural frequency

A1-3 Second order with inverse response transfer function approximation

A system exhibiting inverse response can be represented by a second order system with a zero, as given in equation A6 below.

$$G(s) = \frac{K(1+zs)\omega_n^2 e^{-T} a^s}{s^2 + 2\omega_n \zeta s + \omega_n^2} = \frac{K(1+zs)e^{-T} a^s}{\tau^2 s^2 + 2\tau \zeta s + 1} \quad \dots(A6)$$

The shape of the response depends on the location of the zero (left (negative) or right (positive) side of the s-plane), and on the sign of the steady state value, as shown in figures A7 and A8 below.

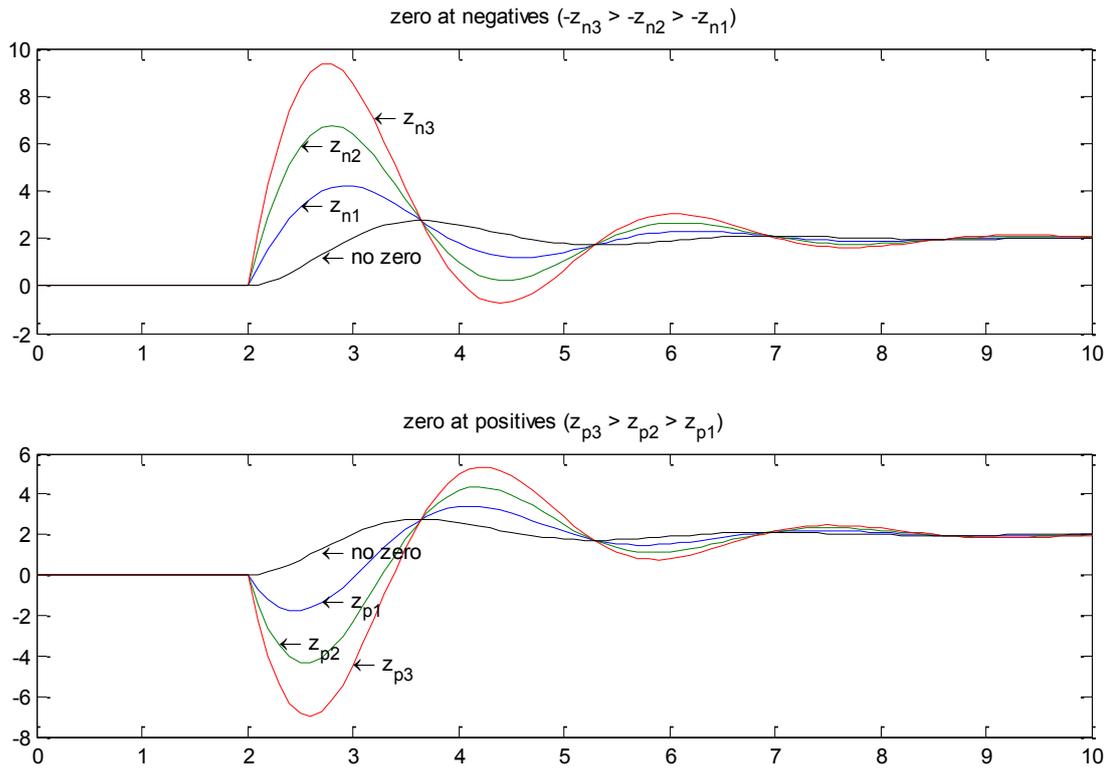


Figure A 7: Effect of zero location on second order system with positive steady state values

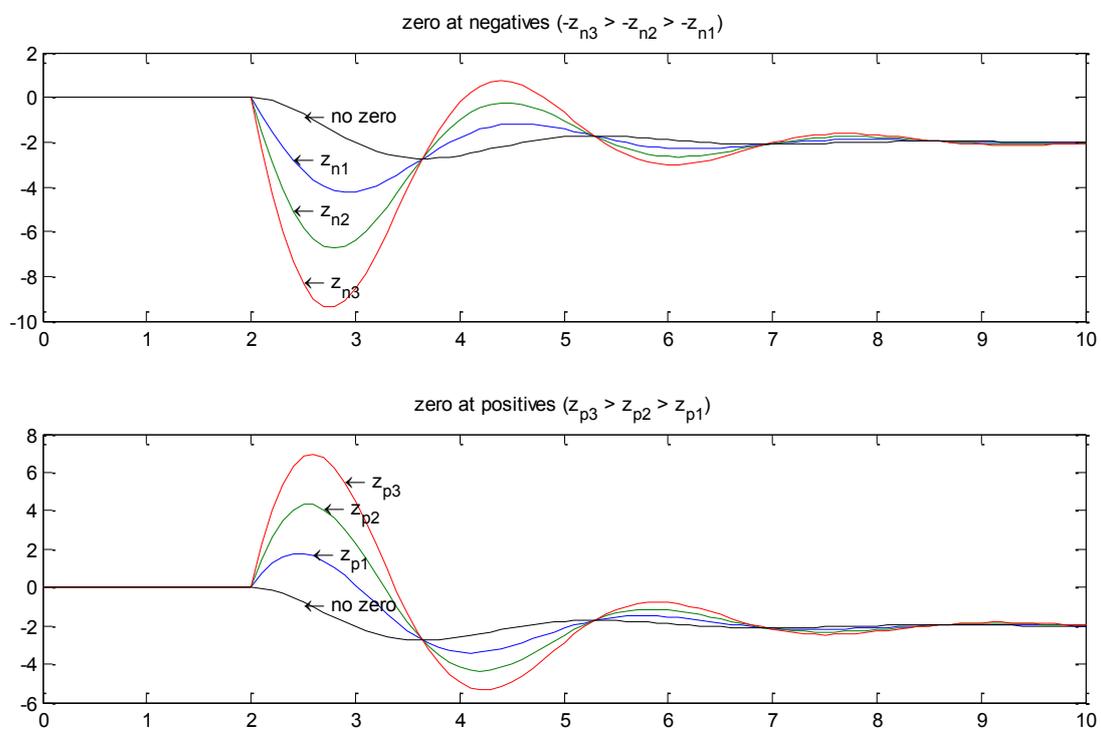


Figure A 8: Effect of zero location on second order system with negative steady state values

Also, the effect of damping coefficients and natural frequencies on second order system with inverse relations are shown in figure A9 below

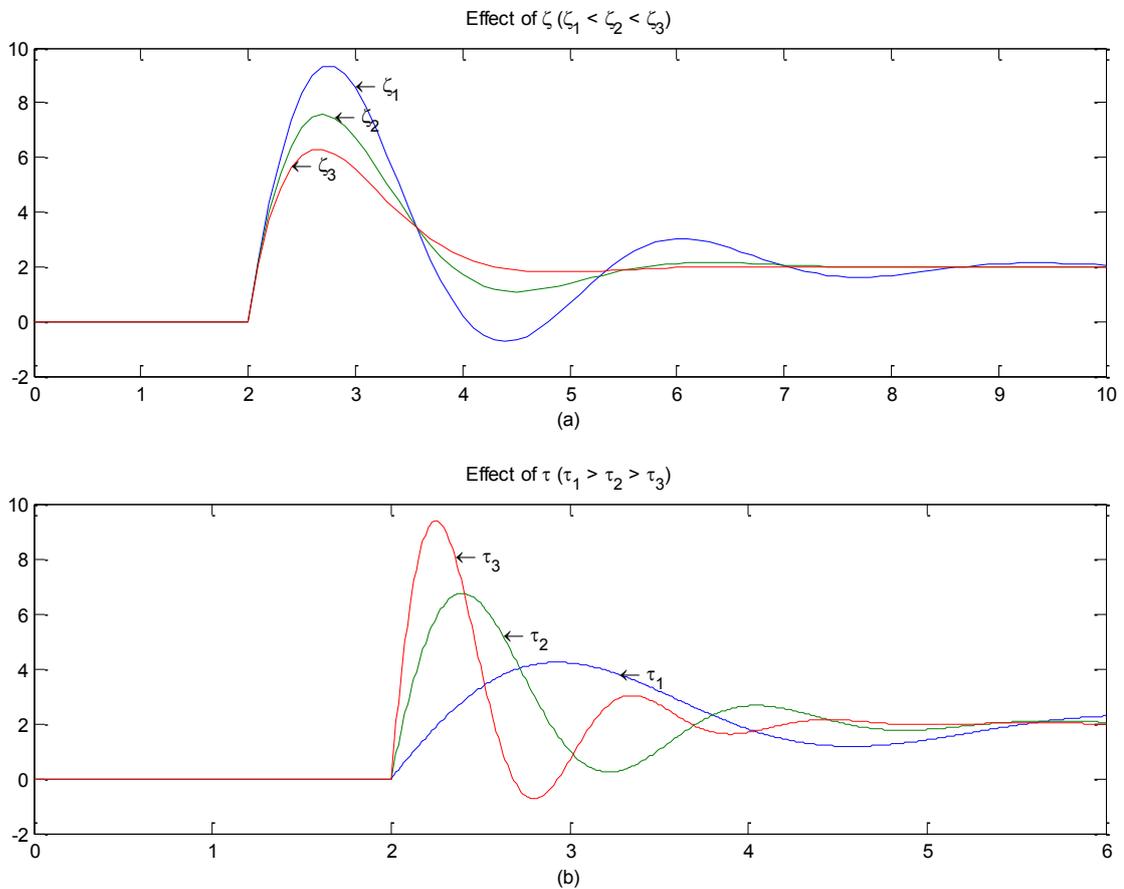


Figure A 9: Effects of damping coefficient and natural frequency differences on second order system with inverse response. (a) same τ , different ζ ; (b) same ζ , different τ .

From figure A7, A8 and A9, it can be summarised that:

- 1) Zeros affects the magnitude of the inverse responses (and by extension the overshoots and undershoots), but does not affect settling time (figures A7 and A8).
- 2) As before, damping coefficients affects overshoots and undershoots, but not the settling time

Natural frequency affects the settling time, and both the overshoots and undershoots

