PREDICTIVE MONITORING AND CONTROL

MOHAMMAD AMINUL ISLAM KHAN

Predictive Monitoring and Control

by

©Mohammad Aminul Islam Khan

A Thesis
submitted to the School of Graduate Studies
in partial fulfillment of the requirements for the degree of

Master of Engineering

Faculty of Engineering and Applied Science

Memorial University of Newfoundland

February 2013

St. John's Newfoundland

Abstract

This thesis investigates the effectiveness of model predictions in two important fields of process operation: process monitoring and process control. Both monitoring and control are essential for the safe and optimal operation of a chemical process. Process monitoring is necessary to notify the operator about an undesired condition, whereas a controller directs a process to desired condition. In Part I of the thesis, a novel model-based predictive technique is proposed for early warning generation to the operator. In Part II of the thesis, an in-depth simulation-based investigation is carried out to evaluate the performance between two control structures: 'model predictive controller (MPC) cascaded to the proportional-integral-derivative (PID) controller', and 'PID-free MPC'.

The proposed early warning generation method uses open-loop process model and disturbance model. Process feedbacks are used to correct prediction bias. This method exploits the controller limitations in dealing with process time delay and actuator constraints. An warning in form of alarm is issued if the open-loop predictions exceed the threshold within the process time-delay. Actuators also plays a major role in controlling processes. If at any point actuators do not have sufficient capacity, controller will fail to regulate the process. Using the process gain, measurements, and constraint information a linear programming algorithm is used to check the existence

of a feasible solution that meets all constraints. An alarm is generated if no feasible solution can be found (i.e., the actuator does not have sufficient capacity). The proposed method is applied to a single-input-single-output (SISO) and a multi-input-multi-output (MIMO) continuous stirred tank heater (CSTH) system. The method gave early warnings compared to the existing safety systems. Also the method demonstrated robustness under small disturbance in the process.

The comparative study between the 'MPC-cascaded to PID' and the 'PID-free MPC' structure is carried out on a simulated CSTH system. In the cascaded structure the flow-loops are regulated by the PID controller. On top of that a dynamic matrix controller (DMC) manipulates the set-points of the flow-loops to control tank temperature and level. The 'PID-free MPC' structure uses a dynamic matrix controller (DMC) to manipulate the valve positions directly. The study reveals that the PID-free MPC structure outperforms the cascade structure in both disturbance rejection and set-point tracking. However, the PID-free MPC structure demands more control action and has more computational load. Integrated square error (ISE) is used to quantify the performance.

Acknowledgements

I would like to take the opportunity to thank Dr. Faisal I. Khan and Dr. Syed A. Imtiaz for all their generous help, support and valuable suggestions. Since the first day, they have been steady guides on this difficult voyage. There were times of frustration during research, when they encouraged me and provided me with directions, spending hours with me solving problems.

Throughout my two years at Memorial University, I was lucky to have Professors Dr. Faisal I Khan and Dr. O. Dobre as my course teachers. They helped me to improve my analytical skill and enhanced my knowledge in control and safety and communication engineering. I am grateful for the generous support, guidance and suggestions from Dr. Salim Ahmed and all the help he gave me during my research work. The support I received from my fellow colleagues Musfiqur Rahman, Raihan Mallick is also greatly appreciated. My gratitude also goes to Mustafizur Rahman, Raju Hussain and Taufiqur Rahman for offering me help to understand control and dynamics.

I also want to take this opportunity to express my gratitude to a few friends whose company made my stay in St. John's a soothing experience. I would like to thank Taufiqur Rahman and his family, Sayeed Russel and his family, Tanmoy Das, Makamum Mahmood and Ridwan Hossain Razin for their help and companionship.

I would like to express the highest gratitude to my family in Bangladesh, especially my father, brother and wife, as well as my other siblings and relatives for their encouragement and support.

Finally, the financial support provided by the Research and Development Corporation, School Graduate Studies, Memorial University of Newfoundland and the Faculty of Engineering and Applied Science is much appreciated.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Predicting the future is an important part of the preparation for future events in our every-day life. Instinctively predictive models are used to predict the outcome of certain phenomena. For example, when someone is going out of the house and observes that, the sky is covered with black cloud, instinctively he would take either an umbrella or a rain-coat to negate the effect of a possible shower. This philosophy of pre-scientific prediction can be converted into a scientific prediction by satisfying some requirements. The phenomena have to be fully explainable using science, the outcome of certain phenomena have to be consistent and understandable, and finally phenomena can be expressed either through numbers or by logic. Satisfying these criteria, a prediction can be made scientific. Such scientific predictions have significant applications in many branches of science and engineering.

Predictions are also extensively used in the process industry for control purpose. Due to the extensive use of model predictive controller (MPC), the open-loop dynamic model of the process is usually known. In this thesis, the predictive power of these process model is used in two important areas of process operations: monitoring and control. In part I of the thesis an alarm system is developed based on open-loop dynamic predictions. In part II a simulation-based study is carried out to compare the purely predictive control structure (i.e., MPC manipulating actuator) with the existing hybrid (i.e, predictive cascaded to feedback) control structure.

Monitoring is essential for safe and uninterrupted operation of a plant, which is one of the primary goals in any process. On average 1500 process variables are continuously monitored in a typical process plant. For safe operation, each process variable is required to be inside certain limit values. If a variable violates these limits, an alarm is generated to alert the operator to take corrective actions. Failure of the operator to take the necessary actions before a process variable goes over the safety limit may cause severe consequences to equipment as well as to human and the environment. Detection of an abnormal situation in time is absolutely critical to avoid both human injuries and equipment damage. The state of the art univariate alarm systems generate alarms based on process measurements. As such, they often lack the ability of prediction. Advance multivariate monitoring methods use models and exploit correlations between different variables for fault detection and alarm generation. These methods have more success in issuing early warnings. However, some of these methods do not use the power of prediction fully and do not consider the impact of the controller and actuator explicitly on the alarm system. The proposed methodology considers the above factors and develops a truly predictive alarm generation system which is able to detect any impending fault at a very early stage. Using the proposed methodology the information of a possible abnormal situation can be sent to operators early on, which provides more time for the operators to respond to an abnormal situation.

The second part of the study deals with the predictive control. The prediction of variables is already in use in the field of process control in the form of MPC, but in most cases the existing MPCs are used as a supervisory layer over the base level PID controller. This structure does not allow the potential benefits of the MPC to be fully harnessed. This is why a PID-free control structure is proposed in the current study, where control valves are directly manipulated by the MPC. This structure offers several advantages: full use of valve capacity, handling of multiple feed-forwards, etc. Moreover, as no PID is present, updating is not required in MPC for changes in PID tuning parameters. A simulation-based comparative study is carried out to evaluate the performance of these alternative control structures. The study confirmed the advantages of the PID-free MPC structure over the MPC-cascaded to PID controller structure.

1.1 Objectives of the Current Study

This research is aimed to investigate and develop predictive methodologies for process monitoring and process control. In the process monitoring part, a novel model-based predictive early warning generation technique is proposed through a predictive alarm system. Proposed predictive alarm system uses open-loop predictions from process and disturbance models. To make the methodology robust, biases of the predictive signals are corrected using process measurement and a heuristic rule is used to generate alerts to the operator. The focus of this study is on generating early warning for operators and providing more time to respond in abnormal situations.

In order to achieve the objectives, followig specific tasks were set at the begining of this work.

- Develop a comprehensive theory for application of the open-loop dynamic model for early warning generation.

- Develop the methodology and algorithm for the predictive alarm generation system.

- Demonstrate the performance of the proposed methodology using process system in a simulation environment.

The predictive control study is the second focus of this thesis. The aim of this work is to perform a comparative study between two competing control structures: 'MPC cascaded to PID' and 'PID-free MPC'. A two-layer PID structure is initially used to control a process. The PID controllers of this structure are gradually replaced by MPC, and the process is made PID-free. Comparative performances of the different control structures are studied discussing the advantages and limitations. The specific tasks for this study are to

- Implement two control structures: MPC-cascaded to a PID controller and PID-free MPC on a continuous stirred tank heater (CSTH) system.

- Investigate the performance of the alternative control structures for set point tracking and regulatory control using quantitative measures.

- Investigate actuator demand and the effect of actuator non-linearity on control performance of these alternative structures.

1.2 Thesis Organization

The first chapter of this thesis briefly describes the motivations for this research and objectives of the study. The thesis is thematically divided into two parts. Part I deals with predictive alarm generations. Part I consists of Chapter 2, Chapter 3, Chapter 4 and Chapter 5. Part II documents the comparative study between the 'MPC cascaded to PID' and 'PID free MPC control' structures. Part II consists of Chapter 6, Chapter 7, Chapter 8 and Chapter 9.

Chapter 2 covers the extensive review of literatures on process monitoring. A brief introduction which discusses alarm management history and standards is also given.

The mathematical formulation of the proposed methodology is discussed in Chapter 3, mentioning the two limiting conditions.

Chapter 4 is devoted to describe two case studies to show the effectiveness of the proposed methodology. This chapter includes the detailed plant description and simulation results and discussions.

Chapter 5 documents the contributions of the predictive alarm generation system described in part I. Recommendations for more robust techniques are also provided in this chapter.

In Chapter 6, motivations to replace PID with MPC are stated. This chapter also covers the existing works to replace PID and a brief historical review of MPC.

The control algorithm used to design the MPC controller, is discussed in Chapter 7.

Chapter 8 describes the different control structures and gives the plant description. This chapter provides a detailed comparison of the performances of different control structures.

The contributions of the comparative study in part II are discussed in Chapter 9. Recommendations for future work are also provided in this chapter.

PartI:

Early Warning Generation through Alarm System

Chapter 2

Literature Review

2.1 Robust Alarm Management

Robustness is one of the main desired properties for alarms. Different surveys have been performed in process industries to identify the critical requirements needed to improve an alarm system. The major problems identified, are a lack of prioritization of the alarm, system rigidity, alarm flooding, a lack of well-designed alarm limits, and stress on the operators due to the high number of alarms. [Shahriari et al., 2006]proposed an ideal alarm system which emphasises the proper prioritization of alarms to make a well-designed system and ensure a less stressful work environment for operation. They also suggest that both the control system and monitoring system should be dynamic to make the alarm more reliable. However, the discussion remains limited in building the criteria of an ideal management system, which are the guidelines used to develop a robust alarm system to overcome the problems found in the survey.

Once the critical criteria for an alarm system is defined, it is recommended that the alarm system goes through a repeating cycle of analysis, plan and countermeasures.

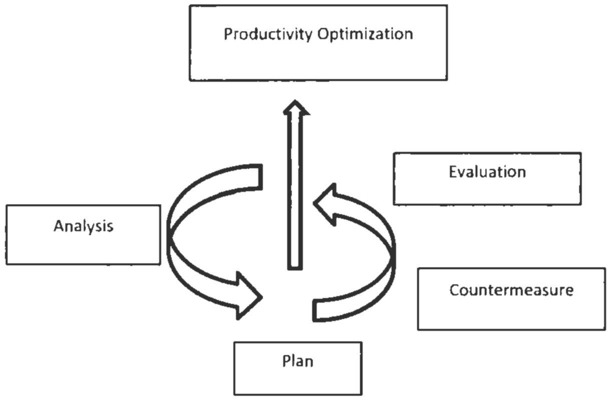


Figure 2.1: Spiral Improvement Cycle [Yuki, 2002]

As plant equipment and control systems changes frequently, the effect of alarm system improvement might not last long. [Yuki, 2002] recommends a three-step continuous effort to keep the alarm system optimized over time. In the first stage, a plan is formulated based on the analysis found from history. Based on the plan, counter-measures are developed for the plant. While counter-measures are being developed, evaluation is performed to observe the effectiveness of the counter-measures. Based on the data from the evaluation and counter-measures, a detailed analysis is performed to provide a plant for the next cycle.

[Chu et al., 1994] outlines a guideline for abnormal situation management. A framework is proposed that integrates some existing features available in the industry. User intent recogniser, a causal reasoning diagnosis system and an advanced graphical user interface is integrated together to assist the operator in taking the necessary action in abnormal situations. A detailed methodology for execution is described in this article.

Execution of the abnormal situation handling is performed in three stages: orientation, evaluation, and execution. The orientation stage refers to the task of focusing on the information relevant to a particular problem, whereas the evaluation stage includes the diagnosis and assessment of the situation. Finally, in the execution stage, the operator executes the necessary action based on their knowledge of the plant and abnormal situations. For successful execution of all these stages, the framework suggests different softwares. The user intention recogniser interprets the operator's goal based on his actions. It checks the current plant state with the plant history to provide feedback on whether the operator's action is consistent with the previous actions. Causal modelling is a way to identify the process disturbance early and predict the future effect of the disturbance. An advanced graphical interface is the medium of interaction used between the operator and the process plant. The effectiveness of the integrated framework is illustrated in this article through a case application to a distillation column.

[Ruiz et al., 2002] proposed an advanced framework using data history, the first principle plant model, and HAZOP analysis. The scheme used is provided in Figure 2.2. The fault diagnosis system (FDS) of the scheme is provided in Figure 2.3. Data initially is pre-processed to make it usable as the input to the FDS. The FDS is designed using an artificial intelligence system based on a neural network and fuzzy logic system.

Data pre-processing consists of various key tasks such as trend generation, principal component analysis, filtering, and data reconcialiation. The data is the input to the FDS which is a combination of a pattern recognition approach based on the neural network and fuzzy logic. Historical data is used by the neural network for recognition of the trend. On-line data from the system is used as the input for the neural network. Outputs of the neural network are used to generate the residual to diagnose the fault

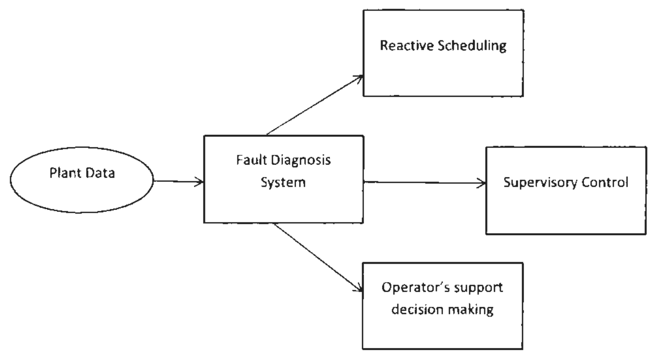


Figure 2.2: Alarm Scheme [Ruiz et al., 2002]

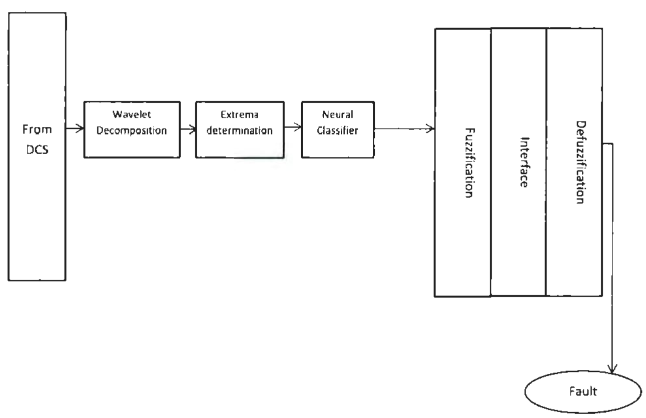


Figure 2.3: Fault Diagnosis System [Ruiz et al., 2002]

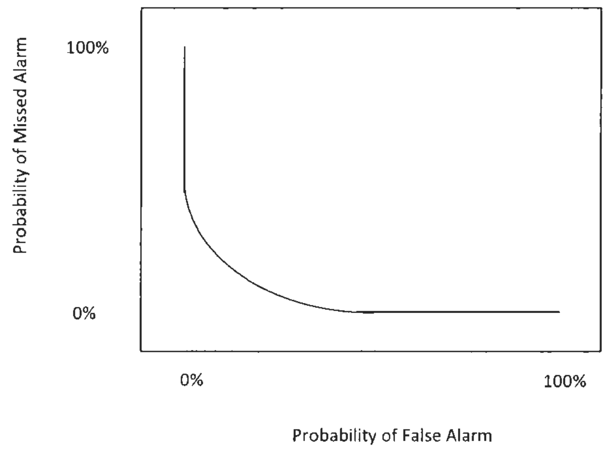


Figure 2.4: ROC curve [Izadi et al., 2009b]

using the historical data as reference. Usually abnormal conditions do not occur too frequently in the plant. Therefore, data history may not show a good trend for the abnormal situation. This is why a plant model is required to generate input output data for abnormal situations and generate a trend for the abnormal condition. HAZOP analysis is used to determine the root cause and IF-THEN rules are used to direct operators about the state of the plant and the necessary action to take. The complete methodology is demonstrated using a industrial case study.

An elaborate discussion on alarm system analysis and design is provided in [Izadi et al., 2009a]. The alarm flooding problem is approached here with different possible solutions. The key concept of false alarm and missed alarm is discussed, which are two of the major concerns in the process industry. A missed alarm leads to a severe consequence and is a major concern from the safety perspective, whereas a false alarm causes distraction for the operator and limits the efficiency of the operator's actions to clear the alarm. Considering the trade-off between false alarms and missed alarms an optimized alarm system framework is proposed in [Izadi et al., 2009b]. Three techniques (e.g. filtering, dead-band, and delay timers) to optimize the process alarm system are discussed. Alarm optimization is performed considering the probability of false alarms and missed alarms as optimization parameters. For an illustration of probability of a false alarm and a missed alarm, a graphical representation receiver operating characteristics (ROC) curve is discussed. A ROC curve is illustrated in Figure 2.4. The two axes of the curve represent the probability of false alarm and probability of missed alarm. Alarm optimization is performed based upon the minimum distance of the operating point of false alarm and missed alarm from the origin point of the ROC curve. The effectiveness of the ROC-based design is demonstrated through process data. Filtering is considered to be the second technique that can be used to reduce the false alarms and missed alarms caused by the process noise. The moving average and moving variance filter are the recommended methodologies. The compatibility of these filters in different scenarios is illustrated through process data. Delay timers and dead-band are also effective methods to reduce false alarms and missed alarms. The techniques described here can cause a significant reduction in false alarms and missed alarms. One common shortcoming for all the methods is that each of them introduces detection delay to the alarm system. These concepts make the alarm system more reliable, but the process remains vulnerable when the early detection of fault is necessary. This is because detection delay was not considered as an optimization parameter in these techniques.

Detection delay is considered as one of the optimization parameters to design threshold in [Adnan et al., 2011]. An optimum threshold design is very important in the process industry to optimize false alarms and missed alarms. Too high a threshold may cause missed alarms, which can lead severe consequences, whereas a low threshold increases the number of nuisance alarms, which can interrupt operators' attention and thus degrades the alarm system's reliability. In this design procedure, the threshold is optimized with an objective to minimize false alarm, missed alarm and detection delay in the alarm system. Markov processes are used to calculate the detection delay for those techniques. False alarm and missed alarm are first optimized using an ROC curve. Then, for different thresholds, false alarm rate (FAR) and missed alarm rate (MAR) are plotted with a different number of expected detection delay (EDD). The optimized value of a threshold can be detected for a given number of EDDs. The design procedure is illustrated using two industrial case studies which show superior performance. However one basic limitation of the methodology is that, it assumes that probability density function of processes for fault free and faulty case are completely known. But, in the real process, it is hard to define a probability density function for faulty cases. For processes that changes dynamics rapidly, performance of this design procedure would not be as good as is depicted in case studies.

[Chang et al., 2011] discussed a risk-based approach to design warning to the operator. A standard alarm response cycle is used to define process safety time, as is illustrated in Figure 2.5. Process safety time was considered to be absolutely crucial in designing the warning system, as an operator must respond to the abnormal situation within this period. Risk is assessed based on three parameters: process safety time, the probability (P) of the potential hazard, and the severity or impact (I) of the consequences.

Initially, a voting system is applied to reduce alarms that are generated based on raw sensor measurements. Hazards of the system are studied and the probability and severity of the impact is mapped. Considering the process safety time t, for a potential hazard risk is calculated Equation 2.1

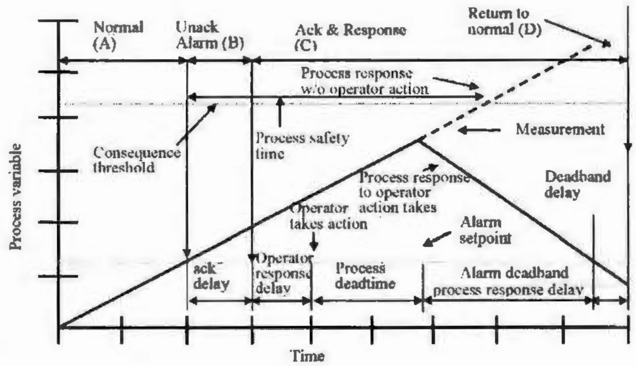


Figure 2.5: A standard alarm response cycle [Chang et al., 2011]

$$R = PI100^{(1-t/60)} \tag{2.1}$$

where P is the probability of the hazard occurrence, I is the severity of the impact of that hazard, t is the process safety time, and R is the final risk.

A risk threshold for the process is defined based on the expert knowledge of the system. As process safety time and impact are considered in the risk equation, hazards with higher importance provide more risk and, thus, the method is useful for thr proper prioritization of an alarm.

2.2 Univariate Alarm Generation Methods

Control chart, also known as Shewhart chart is a statistical process monitoring tool which gives an indication of the process state at any instant. Historically, this has been used as an indicator for a trusted alarm management system. Construction and detailed statistical aspects of this chart is discussed in [Shewhart, 1931].

A control chart is built using the points representing the measurements of a quality

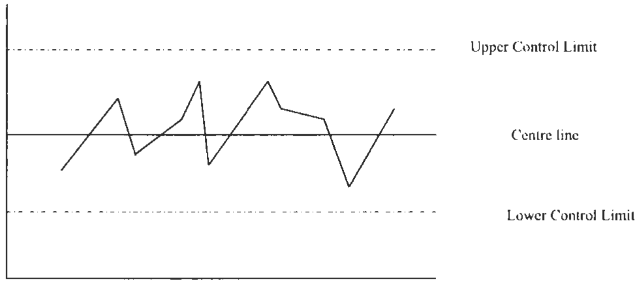


Figure 2.6: Control Chart

characteristics taken from the process at any given time. For any particular variable, the mean and standard deviation are calculated from the data points. The centre line is drawn using the mean of the measurements, whereas standard deviation is used to define the upper and lower control limits. Typically, three-sigma values to both side of the mean are used as upper and lower control limits. As $\pm 3\sigma$ covers 99.70 percent of the normal distribution, when data points lie inside the control band limit, process is believed to be in normal state. An observation outside the limit value indicates the introduction of a new source of variation and defined as special-cause variation. Special-cause variation requires immediate investigation to keep the process at an optimized level.

After detecting a change its cause should be identified and, based on the objective, further action is taken. In the case of a good change, a detected cause should be considered as a new way of working, whereas in the case of a bad change, the detected cause is required to be eliminated. Decision making may appear difficult when the process operating conditions or set-points vary frequently. In such case the method cannot distinguish between normal operational changes and an abnormal condition.

2.3 Advanced Process Monitoring Techniques

Though classical univariate methods are more popular due to their simplicity and robustness, they are unable to provide an in-depth diagnosis of the fault. Due to the availability of large number of sensors, there is a probability that the number of alarms can be triggered from a single abnormal cause of the process. [Kresta et al., 1991] introduced statistical process control (SPC) charts which are analogous to conventional Shewhart charts. only with the additional multivariate nature. Multivariate methods can compress the information down into lower dimensional spaces retaining significant part of the process information. Significant work has been performed in the field of advanced monitoring, resulting in a wide range of methodologies. [Venkatasubramanian et al., 2003b] describes a good classification on these methodologies. On broader scale, these methods can be classified into two major classes: model based approach and historic data based approach. The model based approach can be further divided into two categories based on the nature of the models used for monitoring: quantitative and qualitative. Quantitative monitoring relies on accurate quantitative models, whereas the other type is based on qualitative models. History based method uses historical data to extract features. This extraction of features can be either qualitative or quantitative. Expert system, and trend modelling are two common methods of qualitative feature extraction. Quantitative extraction methods can be further classified as non-statistical (e.g. neural network based), and statistical (e.g. Principal component analysis (PCA)/partial least squares (PLS) based) methods [Venkatasubramanian et al., 2003b].

Quantitative model based approaches are suitable for building monitoring systems for small process units. Models can be built either by using first principles or frequency response. The most important class of models that are frequently used is input-output or state-space models. Actual system behaviour is checked with the system model for inconsistency which is termed as residual. From the residual values decisions are made whether a system is faulty or not. [Isermann, 2005] describes the advantages of the model based approaches over the classical trend checking fault monitoring. Advanced methods provide early detection of fault, good supervision under close loop, and supervision facilities under transient state of the process. Advanced process monitoring systems consist of two parts. One is fault detection, the other is fault diagnosis. A brief introduction of different fault detection and diagnosis methods is provided in the following sections. The most frequently used model based FDI approaches include diagnostic observer, parity relation, Kalman filters and parameters estimation. [Isermann, 2005]

[N.Clark, 1979] and [Massoumnia, 1986] are some of the pioneering works on diagnostic observers. [Frank, 1990] decoupled the effect of fault and presented a more robust fault detection method. This work also considered a solution for a non-linear system in the form of a diagnostic observer. A non-linear observer is designed for non-linear system in the [Dingli et al., 1995]. In the parity relation approach, consistency of the plant model is checked with the sensor output. In this case, the plant model is formed by rearranging the parity equations. Dynamic parity relation was first introduced by [Willsky, 1976]. It was further extended by [Gertler et al., 1990], [Gertler et al., 1995] and [J. Gertler and Monajemy., 1995]. Some other significant works using parity relations can be found in [Ben-Haim, 1980], [Ben-Haim, 1983] [Chow and Willsky, 1984]. Among model based approaches, Kalman filter based approaches are used most frequently. It uses a recursive algorithm for state estimation and has a wide application in the field of process monitoring. The Kalman filter was first introduced by [Willsky and Jones, 1976] for fault detection and it was further advanced through the work of [Basseville and Benveniste, 1986] and [Basseville and Nikiforov, 1993]. [Fathi et al., 1993] used an extended Kalman filter (EKF) in designing local detection filters. [Chang and Hwang, 1998] proposed a suboptimal EKF to provide computational efficiency.

A model based approach for the handling of abnormal situations during the process transition is discussed in [Bhagwat et al., 2003a]. The motivation for this work is the inability of automatic control systems to cope with large changes in process variables during the transition. For this reason, transitions are executed manually by operators. Thus, a process is more vulnerable to the faults during the transition phase. Fault detection and identification (FDI) systems in practice, are usually designed assuming that, the process is at steady state, which is not suitable for process monitoring during the transition phase. As such, sudden changes due to discontinuous phase changes are detected wrongly as fault. Also, proper diagnosis cannot be done in the case of operator errors. These issues were addressed in this article. For the off-line development of the model, first standard operating point (SOP) and process knowledge are used to break the transition into different phases. Model components are extracted from the process knowledge for different phases. Based on the model components, different types of filters are designed. For non-observable components an open-loop observer is used, whereas a Kalman filter is used for observable linear components. For non-linear components an extended Kalman filter is designed. For on-line implementation, a phase is detected based on phase definition and the phase model component is selected initially. Based on the phase model component, a suitable filter is activated and residuals are generated. Residuals are then passed to a fault detector and, based on the fault definition, fault is detected. After detection, logical analysis is performed using a fault map before sending a notification to the operator. The main challenge for this method is developing the non-linear model and filter for the system which is often difficult and costly.

Keeping the main scheme similar, [Bhagwat et al., 2003b] used a set of linear models along the transition trajectory to build the monitoring system for transient systems.

In quantitative models, a priori knowledge of the process is expressed in terms of mathematical functional relation between input and output. In contrast qualitative models, understanding of the process knowledge is expressed in terms of qualitative functions. There are two major strategies used to develop qualitative models: causal models and abstraction hierarchy. Causal models can be formed using different strategies such as sign directed graph (SDG), fault-tree analysis, and qualitative physics. Primary requirement for qualitative model is to develop an expert system that mimics the behaviour of a human expert to solve problems. Usually, it is comprised of large sets of if-then-else rules and an inference engine which makes decisions based on the process knowledge. [Venkatasubramanian et al., 2003a] provides an excellent review on qualitative fault diagnosis methods.

The model based approach proved to provide better performance compared to classical methods. The model based approach is normally limited to processes with a small number of variables due to difficulties in building models for complex systems. In the case of handling a large number of variable data based multivariate process monitoring techniques are more useful. For the successful execution of a quantitative model based approach, the adaptability of the methods to the physical property of the process is required. In the case of non-linear process model formation becomes a bit complicated.

In contrast to the model based approach, history based methods require only a large historical data set of the process, and an explicit model of the system is not required. Features can be extracted using both a qualitative approach and quantitative approach. Qualitative data extraction uses an if-then-else structure similar to a qualitative model based approach. The only difference is that instead of using the input output functional relationship, an input-output trend formed from historical data is used. The qualitative approach can be largely divided into two types: expert systems and qualitative trend analysis. A comprehensive list of the methods can be found in [Venkatasubramanian et al., 2003b].

Quantitative feature extraction can be largely classified into non-statistical and statistical. Among non-statistical feature extraction approaches, a neural network (NN) is widely used in the field of fault diagnosis of chemical processes. Both supervised and unsupervised learning strategies have been used. Back propagation algorithms are most popularly used for supervised learning strategies. Some earlier work in the field of fault diagnosis using neural networks are [Venkatasubramanian, 1985], [Watanabe et al., 1994b], [Venkatasubramanian and Chan, 1989], [Ungar et al., 1990], [C.Hoskins et al., 1991]. A more detailed and thorough analysis of NN for fault diagnosis in steady state is presented in [Venkatasubramanian et al., 1990]. This work was later extended for a dynamic process in [Vaidyanathan and Venkatasubramanian, 1992]. A hierarchical neural network architecture for multiple fault detection was proposed by [Watanabe et al., 1994a]. Standard back propagation is improved for better performance by introducing explicit features to NN. [Fan et al., 1993], [Farell and Roat, 1994], [Tsai and T.Chang, 1995] presented the idea of the improvement of the back propagation algorithm. [Leonard and Kramer, 1990] suggested the use of a radial basis function network. Some other significant works using neural networks can be found in [Holcomb and Morari, 1991], [Kavuri and Venkatasubramanian, 1994], [Bakshi and Stephanopoulos., 1993].

The second type of quantitative feature extraction methodology includes multivariate statistical process monitoring which typically uses only a few feature variables to monitor plants' performance. Since the pioneer paper by [Kresta et al., 1991] PCA and PLS have been used extensively to monitor chemical processes. Some earlier works on fault detection and diagnosis using PCA and PLS are [MacGregor et al., 1994] and [MacGregor and Kourti, 1995]. [Qin and McAvoy, 1992] presented a neural net PLS approach to deal with non-linearity. [Dong and McAvoy, 1996] utilized a nonlinear PCA method to handle non-linearity more efficiently. One of the more recent techniques are discussed in [Raich and Cinar, 1996] combining PCA and discriminant analysis techniques.

PLS and PCA can project the information down to lower dimensional space. With the use of PCA or PLS, the dimensionality of the process is reduced. Highly correlated large data sets are reduced to a few latent variables that contain the most process information. Projections of new process observations over time on low dimensional planes, are plotted to detect an abnormal process variation. A square prediction error (SPE) is used to detect the major change of the process caused by new events. The methodology is simple in nature and it can be said as only the extension of the statistical control chart for a large number of variables. When a larger number of latent variables is required to capture the process information, it would be difficult to monitor a process successfully using this method.

The major limitation of PCA based monitoring is that, model is steady state time invariant. As most of the real time process is time varying PCA may not be as effective in dynamic state as it is in steady state. Therefore, the PCA model is required to be dynamically updated. A recursively updated PCA is proposed in [Li et al., 2000]. In this work, an adaptive monitoring approach was developed which is capable of robust monitoring of a dynamic state. Two different algorithms based on rank-one modification and Lanczos tridiagonalization are proposed. The number of principle components and the confidence limits for process monitoring are calculated recursively. A case study on rapid thermal process is presented to demonstrate the effectiveness of the methodology.

[Nomikos and MacGregor, 1994] extended the use of multivariate projection method to batch process by developing multiway-PCA (MPCA). The objectiv of batch process monitoring the trajectory is to monitor a new batch process with the past good batch runs. This gives rise to a three dimensional data matrix (i.e., time,batch number, variable). A methodology was developed to unfold a three dimensional data matrix to a two dimensional data matrix. Subsequently using PCA new batch trajectory can be composed with a trajectory band based on past good batches. If a significant deviation is detected a warning is generated.

In [Cherry and Qin, 2006], a recursive PCA algorithm is combined with a multi-way PCA for fault detection and diagnosis. The described methodology used a combined index incorporating the information from the SPE and Hotelling's T^2 for fault detection. This work facilitates the diagnosis procedure for processes with a large set of variables.

In [Misra et al., 2002] a multi-scale PCA based method is proposed which is a combination of PCA and wavelets simultaneously. Wavelets can capture the autocorrelation of a sensor, whereas PCA captures correlations across the sensor. These two techniques are combined together for their complementary strengths and maximum information from multivariate sensor data is extracted. First, each variable is decomposed using discrete wavelet transformation. PCA is applied to each of the matrices to extract the cross correlation across the sensors. SPE can be monitored based on the process objective. The moving window approach is used for dynamic monitoring. SPE is calculated for each level of wavelet decomposition. From the SPE values of each level, information about fault can be extracted. Multi-scale PCA is widely used for monitoring rotating equipment, such as compressors, pumps etc.

A recently developed technique independent component analysis (ICA) is used in [Lee et al., 2004] for statistical process monitoring. ICA is used to reveal the hidden factor that underlies a set of non-Gaussian measurements. Unlike PCA, ICA does not assume independence of the measurements in the temporal domain. Dynamic ICA can be used to monitor a process with auto correlated and cross-correlated variables. The results from case studies show that ICA clearly outperforms conventional PCA and dynamic PCA.

However, there are some limitations of ICA based monitoring. First, it is not easy to fix the number of independent components that are required to be extracted for building an ICA model. Moreover, the proper order of ICs cannot be determined as the ICA does not arrange the IC is any order. A modified ICA algorithm to overcome these problems is proposed in [Lee et al., 2006]. This methodology works in two steps. In the first step, the variance of dominant ICs and the direction is detected using PCA.

In the second step, conventional ICA is performed to update the dominant ICs. The method also provides information about the location of fault using a contribution plot.

A comparative study for data driven process Monitoring methods is performed in [Yin et al., 2012]. The literature discusses five basic data-driven algorithms. These are PCA, PLS, ICA, Fisher discriminant analysis (FDA) and subspace aided approach (SAP). These methods are implemented on benchmark Tennessee Eastman (TE) process and their performances are compared. Standard PCA shows a relatively lower fault detection rate (FDR) as it cannot handle dynamic data. DPCA shows better FDR compared to standard PCA. Two variants of PLS provides much improved FDR compared to lower FDR in the standard approach. ICA-related methods provide significantly improved performance compared to standard PCA. However, computational complexity is far greater in the ICA related methods which is a concern. SAP provides a superior performance for most of the cases. Again, as the practical process is large in nature there is a low probability that it would follow a Gaussian distribution. Though ICA provides a solution for this problem, it is physically unexplainable that non-Gaussian distributed process variables can be described as the linear combination of the ICs. This work recommends SAP to be the method which should be given more attention due to its higher FDR.

A comparative study between a causal model based monitoring system and statistical multivariate system is presented in [Yoon and MacGregor, 2000] by implementing the two monitoring methods on a simulated CSTR system. The fundamental and practical differences of the two methods are described with their respective strengths and weaknesses. The parity-relation approach is presented as a representative causal model-based method whereas PLS/PCA is used to build a multivariate statistical process control model. This work concludes with the difference of the two methods. The causal model approaches are generally limited to small, well-defined systems, as the implicit model of the process is required to be known. On the contrary, MSPC can handle ill-defined and large processes very easily. Again, the parity equation approach provides the more direct isolation of known fault through process knowledge, whereas a statistical approach is much more indirect, having no causal information.

[Yang et al., 2010] discusses an alarm limit design procedure by taking the multivariate nature of the process into account. Correlation maps of process physical variable and their alarm history are compared to suggest the alarm threshold settings. Information on the process connectivity is required for this case. One of the shortcomings of this work is that, this methodology is not applicable for a large number of states.

2.4 Predictive Alarm Generation

The early detection of alarms is a necessity in the process industry. However, in alarm system design, typically, more emphasis is put on robustness, as such very rarely process plants generate alarms based on prediction signals. Predictive monitoring can be an efficient tool for the successful forecasting of an abnormal situation.

[Juricek et al., 2001] presented a model based predictive alarm methodology that uses a state space model of the process. A Kalman filter is used to make dynamic predictions of the process variable. The analysis begins only on demand. In the first stage, a pseudo disturbance is estimated to compensate for the plant mismatch and deterministic disturbance. In the next stage, a Kalman filter is applied to predict the future values of the measured variable. Finally, a confidence limit is constructed and this limit along with T^2 statistics, are used to generate early alarm for the system. The design variables for the method are the confidence level, prediction horizon and the form of the pseudo-disturbance. The method is demonstrated using a simulated CSTR model.

[Zamanizadeh et al., 2008] used an extended Kalman filter for alarm generation. The philosophy of prediction and fault detection is similar to that described in [Juricek et al., 2001]. The only difference is, the prediction is made using an extended Kalman filter instead of a Kalman filter to tackle nonlinearity in the process.

A supervisory method to predict an abnormal situation is discussed in [Fernandez et al., 2005]. The method initially identifies the time at which a process variable reaches a critical value. When a process reaches a critical value, the monitoring system starts trending the input and output data using a neural network (NN). Several dynamic models between the input and output are estimated and the best fit model is selected. An optimization algorithm is used to estimate the parameters. The focus of this study was to find the most appropriate model to predict output in the abnormal situation.

[Varga et al., 2010] introduced a methodology that uses a dynamic model and hazard analysis to predict safety limits in the alarm system. One of the main motivations of this work was to guide the operator about the consequence of different hazards initiated at the time of operation. Based upon the consequences of the prediction, the operator could take the necessary actions. Early fault detection is enabled which provides early diagnosis and suggests a preventive measure corresponds to the abnormal situation. The most important function of this alarm system is that it helps the operator in decision-making when more than one variable in the process show abnormal behaviour. This alarm system is particularly necessary for an abnormal system like a thermal runaway, where the system goes from stable to unstable in a space of a very short time, and once an unstable state is initiated, the process cannot be taken into a stable state with the control action. Using a simulator, the last controllable point of the system is detected and the simulator queries in a particular trajectory on the lookout for a possible uncontrollable state of the system. Lyapunov's indirect stability analysis of the state variables, along with simulated trajectories, are used to detect the boundary of the controllable region of the process. A pictorial algorithm provided for finding the safety region is shown in Figure 2.7.

An innovative risk-based fault diagnosis methodology and its integration with SIS for process systems is proposed in [Bao et al., 2011]. In this methodology, risk concept is combined with SPC for fault diagnosis in the processes. The proposed methodology has been validated using a tank filling system and the steam power plant system in the G2 environment. The method is simple in implementation and does not depend on any process model. Also, its demand for historical data is minimal. The proposed methodology uses a control chart to distinguish an abnormal situation from a normal operation based on the three-sigma rule and linear trend forecasting. Time series moving average filters are used to perform real-time prediction and noise reduction of the signal. Based on the forecasting signal the probability of a fault is calculated. The consequences of the fault are identified. Risk is defined as the multiplication of the probability of fault and consequence. An alarm is generated only when risk of a fault exceeds the threshold. Since it considers the consequence of a fault, the method is able to filter out spurious faults from the alarm system and also the operators can

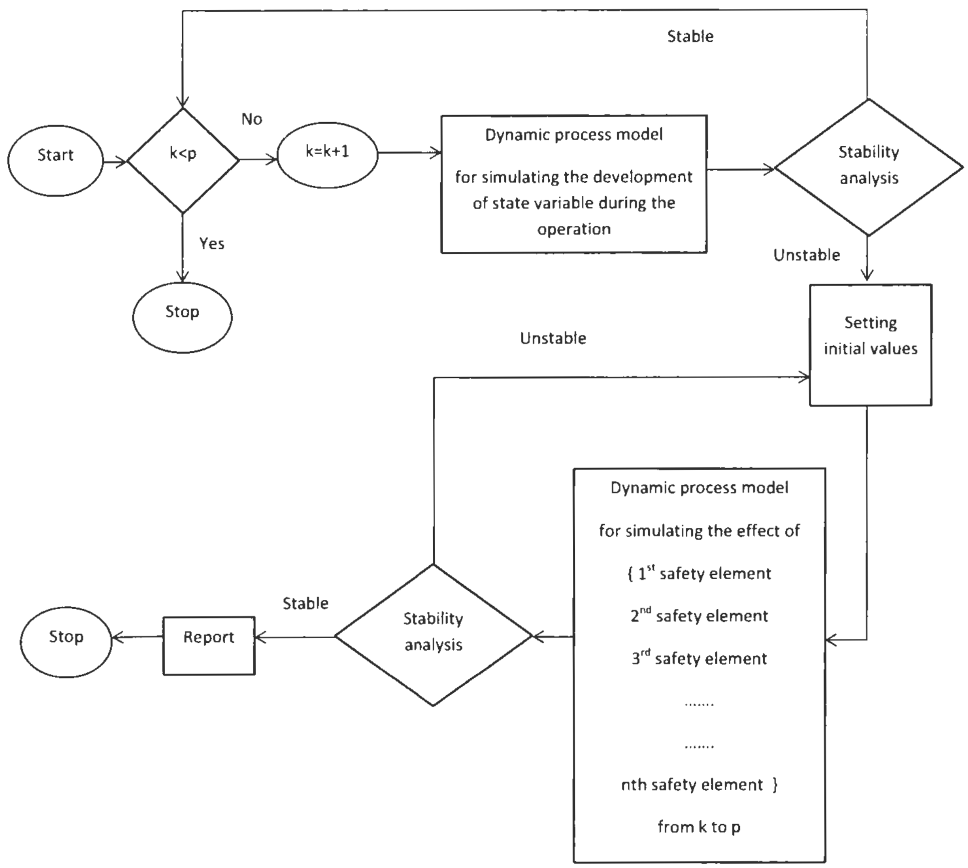


Figure 2.7: Detection of the last controllable state of reactor [Varga et al., 2010]

prioritize their responses based on the quantitative risk. The frame work of this work is represented in Figure 2.8 .

However, this methodology is based on the univariate technique. Hence, it has less power to distinguish between a process fault and an operating change. This limitation is addressed by extending the method to a multivariate model based framework described in [Zadakbar et al., 2011]. In this article, the probability of fault is calculated from the residual generated from a Kalman filter. As the method requires an explicit process model, this method is useful for small process systems when a process model can be easily built. For large process system with complex dynamic, finding an explicit model is a challenging job. In order to overcome this challenge, [Zadakbar et al., 2012] proposed a model-free risk based fault detection and diagnosis method. Instead of using an explicit dynamic model, process data history is used to capture process knowledge. From the historical data, a PCA model is built which projects data in the principal components' direction. Risk is calculated based on the scores of the principal components instead of the original signal.

From the above literature review the following conclusions can be made:

- Extensive work has been done on alarm generation focusing on robustifying the alarm.
- Most of the cases alarms were generated based on the process measured signals.
- Very little work has been performed on generating alarms on a predictive mode.
- None of the current methodologies consider the effect of the controller and the actuator limitations explicitly.

Figure 2.8: Framework of the methodology of risk based fault diagnosis and safety management for process systems. [Bao et al., 2011]

In the present work, early detection of the abnormal situation through model prediction is focused. The open loop predictions are used for generating the alarm. Also, due considerations were given to the effect of controllers and actuator constraints.

Chapter 3

Theory for Early Alarm Generation

3.1 Early Alarm Generation

In the process industry different types of fault can be triggered. Of the different types of fault we only dealt with the disturbance type faults, which is common in process industries. In our current study we focused on design an early warning generation system for this type of disturbance fault that cannot be handled by controllers. In a process, often the process model and the disturbance model are available, particularly in many plants where model predictive controller (MPC) applications are in use to control the process. The objective of this work is to use such models for early alarm generation. These models are open-loop models without any controller knowledge. As such, there are some limitations in using these models. However there are several advantages for using open-loop predictions: (i) an open loop model may be already available from an MPC application; (ii) typically, in a process controller are occasionally retuned in order to meet operational need for faster or slower dynamic response, as such closed-loop models need to be updated regularly. As open-loop models do not have any controller information, it remains valid for a longer period. Considering the above facts, open-loop predictions for alarm generation is used. However, it is understood that an open-loop prediction may show that a process variable will violate the threshold, but the controller action may actually keep the process within the normal operation limit. Therefore, due considerations need to be given to take the effect of the controller into account.

In this paper, we state two limiting conditions where the controller does not affect the open-loop prediction. Under these conditions, open-loop predictions can be used for alarm generation.

Condition 1 : In a system with time delay, any disturbance entering into the system will affect the measurements after the time-delay period has elapsed. Given that the controller has no feed-forward knowledge of the disturbance, a controller will take any corrective action only after the time delay period. As such, the open-loop predictions will be the same as the closed-loop predictions within the time delay period.

Condition 2 : At steady state, the ability of a controller to bring a process variable within the control limit will depend entirely on the available actuator capacity and steady state gain of the process.
Based on the above two conditions we develop two early alarm generation protocols for the process system.

3.1.1 Dynamic Alarm Generation

If the controller does not have the feed forward information (e.g., an MPC without a disturbance model) or a controller acting only based on feedback (e.g, PID), the controller will take action only after the disturbance effects are measured at the output.

Condition 1 states that, within the time delay period, the open loop prediction and the closed-loop predictions are the same. This provides a window where the open-loop prediction can be used to monitor the process. This is defined as *monitoring horizon*. Within the *monitoring horizon*, an alarm will be generated if an open-loop prediction exceeds the alarm threshold. The algorithm essentially uses the receding horizon prediction for the alarm generation. Receding horizon prediction is also the building block for model predictive controller (MPC). At each instant t, using the output from the controller and disturbance measurable to the process, the models predict the process output for the entire *monitoring horizon*, t_p. As new measurements become available, the predicted values are compared with the measurements; if a bias is observed, the predicted values will be corrected for the bias. Instead of simple bias corrections, a Kalman filter can be used to update the predictions. However, for most practical purposes, a bias correction should be sufficient. The detailed steps of the methodology are described below.

Let us consider a dynamic system with $\boldsymbol{u}^{(m\times1)}$ input to the system, $\boldsymbol{y}^{(n\times1)}$ measured output, $\boldsymbol{u_d}^{(p\times1)}$ disturbances affecting the system, m number of inputs, n number of outputs, $\mathbf{G(s)}$ process model, $\mathbf{D(s)}$ disturbance model, and e is the measurement error. System can be represented by the following transfer function equation

$$y^{(n\times1)} = \mathbf{G(s)}^{(n\times m)}u^{(m\times1)} + \mathbf{D(s)}^{(n\times p)}u_d^{(p\times1)} + e, \tag{3.1}$$

The first step is to develop a discrete, time-invariant, causal step-response model for the system given in Equation 3.2. The open-loop model predicts a process variable over a monitoring horizon, taking into account the disturbance model of the process. The monitoring horizon is chosen based on the time delay of the process. For a given process with time delay t_d, monitoring horizon $t_p \geq t_d$.

$$
\begin{aligned}
y^*_{i,t+l} = y_{i,t} &- \sum_{j=1}^{m}\sum_{k=1}^{h}(a_{ij,k} - a_{ij,l+k})\Delta u_{j,t-k} \\
&+ \sum_{j=1}^{r}\sum_{k=1}^{h}(d_{ij,k} - d_{ij,l+k})\Delta u^d_{j,t-k},
\end{aligned}
\tag{3.2}
$$

where, l=[*1,2,3,...,p*] and p is the horizon defined based upon the process knowledge. m is the total number of inputs; r is the total number of disturbances,i=[*1,2,3,...,n*], and n is the total number of outputs. Equation 3.2 predicts the i-th output over the horizon p. Here, m is the total number of input to the system, r is the total number of measured disturbance to the system and h is the number of history inputs, that are considered to predict the output.

In *step* 2, on-line output measurements are used to correct the predicted values. At every instant, the output is corrected by comparing the current measurement with the predicted value from the model. The difference in these two values gives the bias error. The bias error at time t can be calculted using Equation 3.3,

$$
b_t = y_t - y^*_t \tag{3.3}
$$

where y^*_t is the one step ahead prediction at time *t-1*. Based on the calculated error at time t, bias correction is done on all future predictions, as is given in Equation 3.4.

$$
\hat{y}_{t+l} = y^*_{t+l} + b_t \tag{3.4}
$$

where l= *1,2,... p*

The updated predictions show the effect of disturbances earlier than the process measurements, as they are based on both process and disturbance models.

Step 3 is alarm generation. An alarm limit is set for each variable based on process knowledge. At each instant the predicted values are checked against the limits. If the prediction exceeds the limit within the monitoring horizon, an alarm will be issued to alert the operator.

Step 4 improves the robustness of alarm. A single value can sometimes exceed the limit due to measurement noise. In order to make the alarm robust and avoid a nuisance alarm, a further heuristic rule is applied. If three consecutive predicted values cross the limit only at that point an alarm will be issued. However, this rule can be adjusted depending on the severity of the consequences and risk associated with the variable.

3.1.2 Steady State Alarm Generation

The steady state alarm generation algorithm is developed based on *Condition* 2, which was described earlier. Suppose that a process is at steady state and a disturbance enters the system; if there is no controller present in the system the steady state of the system will be disturbed and the system will become steady at a new state eventually. However, when a controller is controlling the system, it will take corrective actions and will try to bring the system back to its original state. Assuming the controller is perfect or very efficient, the ability of the controller to bring the system back to the original state is limited by the available actuator capacity of the system. Therefore, depending on the available actuator capacity, a controller will either bring the system back to the original state or the system will attain a new state which may or may not be within the safety limits. Steady state conditions are checked to see whether the variable can be brought back by the controller within the alarm limits using the available actuator capacity. If it appears that the actuators do not have sufficient capacity to bring the process within the safety limit an alarm is generated.

An alarm generation requires a prediction of the open-loop steady state values due to the disturbance, calculation of the capacity of actuators, and maximum possible control action on variables. The open-loop steady state value of a variable is predicted by adding the change in process variable due to the disturbances in the present steady state value. Consider a disturbance size of Δu^d enters into the system at time t. If there is no control action, the final value of the output is given by Equation 3.5,

$$y_i^{ss} = y_{i,t} - \sum_{k=1}^{p}(d_{i,k} - d_{i,l+k}) * \Delta u_{t-k}^d \tag{3.5}$$

where P is the number of the history inputs required to estimate a variable at steady state.

The minimum requirement from a controller is to make changes in the actuators such that the output remains within the control limits. Assuming that, the high and low limits for the i-th output are $y_{i,low}$ and $y_{i,high}$, respectively, the controller has to satisfy the following condition

$$y_{i,low} \leq y_i^{ss} + \Delta y_i^{ss} \leq y_{i,high} \tag{3.6}$$

where Δy_i^{ss} is the steady state change in the i-th output due to the input changes made by the controller. At steady state the input and output changes are related by the process gain as given below :

$$\Delta y_i^{ss} = \sum_{j=1}^{m} a_{ij}(t_{ss}) * \Delta u_j \tag{3.7}$$

where $a_{ij}(t_{ss})$ is the step response coefficient at steady state, which is equivalent to the process gain and m is the total number of input to the system.

$$a_{ij}(t_{ss}) = G_{ij}(0) \tag{3.8}$$

Equations 3.6 and 3.7 can be combined together to express the desired condition in terms of the input variable as given in Equation 3.9.

$$y_{i,low} - y_i^{ss} \leq \sum_{j=1}^{m} a_{ij}(t_{ss}) * \Delta u_j \leq y_{i,high} - y_i^{ss} \tag{3.9}$$

The capacity of an actuator is given by the difference between current steady state value of actuator (i.e. valve) and high and low limits known from the actuator range, which can be written in the following input constraint Equation.

$$u_{j,low} - u_{j,t} \leq \Delta u_j \leq u_{j,high} - u_{j,t} \tag{3.10}$$

where $u_{j,low}$ and $u_{j,high}$ are the low and high limit values of the actuator respectively.

The controller will be able to bring all the process variables within the desired limits only if Equations 3.9 and 3.10 are satisfied simultaneously. Therefore, Equations 3.9 and 3.10 give the desired condition for steady state alarm generation. If these two equations cannot be satisfied simultaneously an alarm will be issued. A linear programming (LP) algorithm is used to check the existence of a feasible solution for the output constraints arising from Equation 3.9 and input constraints arising from Equation 3.10. For example, for a system with m inputs and n outputs there will be m input constraints and n output constraints. An alarm is issued if there is no feasible solution that satisfies all $(m + n)$ constraints.

Chapter 4

Case Studies

In order to clearly explain the implementation steps, the proposed methodology is first demonstrated on a single-input-single-output (SISO) system. Subsequently the methodology is applied to a multiple-input-multiple-output (MIMO) continuous stirred tank heater (CSTH) system. A schematic diagram to express the alarm generation process is provided in Figure 4.1. This alarm generation procedure is discussed for the two case studies.

4.1 A Simple SISO Example

Consider a simple SISO system with a disturbance input, as described in Equations 4.1, 4.2 and 4.3.

$$y = G(s)u + D(s)u^d + e \tag{4.1}$$

$$G(s) = \frac{e^{-14.7s}}{21.3s + 1} \tag{4.2}$$

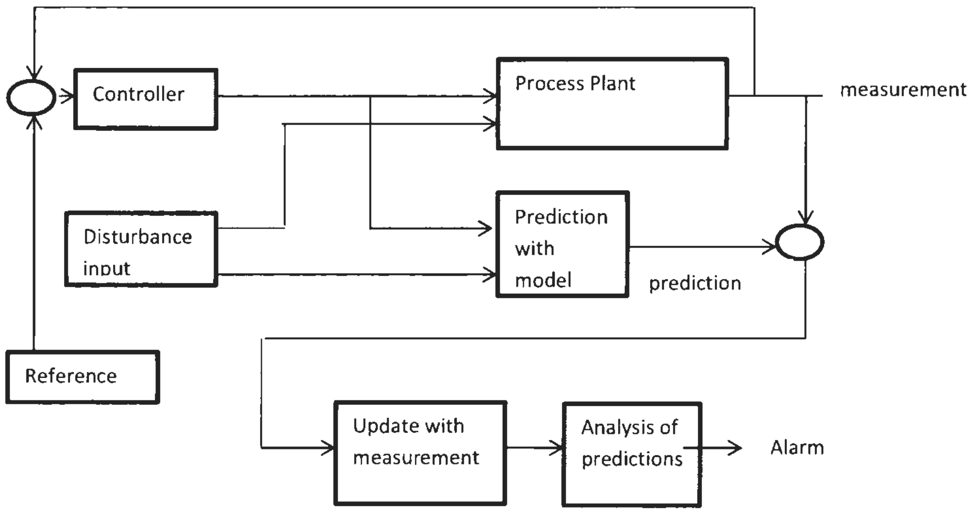


Figure 4.1: Schematic diagram of alarm generation process using open-loop model

$$D(s) = \frac{1}{25s + 1} \tag{4.3}$$

where, y is a output variable, u is an input to the process and u^d is a measured disturbance. The system is controlled using a dynamic matrix controller (DMC). DMC controller is designed using the step response model of the simple SISO system. Different design parameters for designing the DMC is stated in 4.1. The controller only utilizes the process model and does not have any knowledge of the disturbance. This mimics a feedback controller, which is the most common industrial scenario. There is a system time delay of 14.7 s. As such, when a disturbance enters the system, the controller does not take action immediately. After the time delay period (14.7 s) has elapsed, the disturbance starts affecting the output y. At that point, the controller takes action to reject the disturbance and to bring the process back to the desired set point.

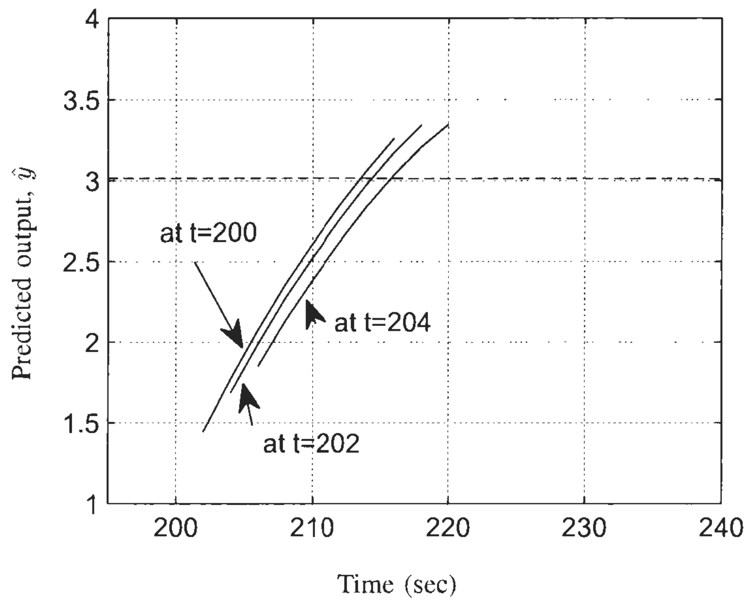
Table 4.1: Design parameters of the DMC controller

Variable	Value
Prediction horizon(p)	15
Control horizon (m)	5
Weighting factor (Q)	1

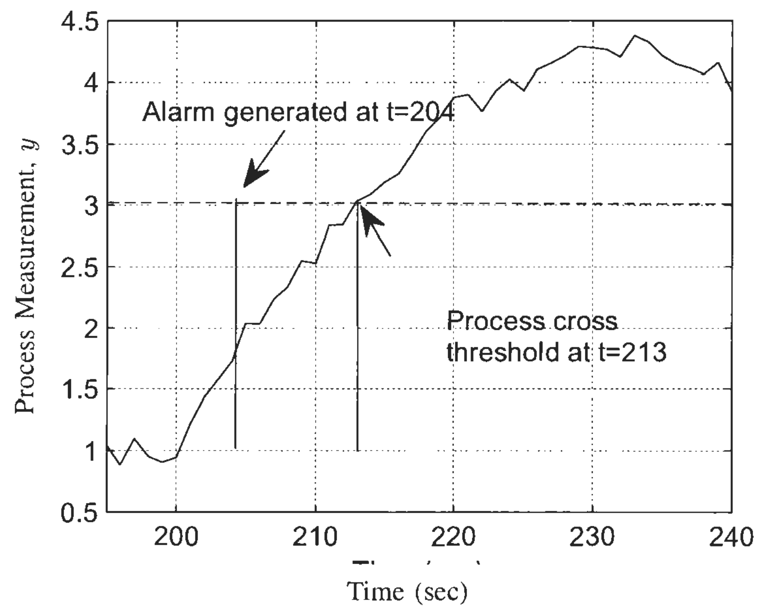
4.1.1 Dynamic Alarm Generation

The purpose of dynamic alarm generation is to monitor the process in the time delay period or *monitoring horizon* when the controller does not have any influence on the system. In this case, 16 sec is chosen as the monitoring horizon which is slightly greater than the system time delay. Using the finite step response models of the system and disturbance as given in Equations 4.2 and 4.3, the output over the monitoring horizon is predicted. The sampling interval for the system is 2 sec: therefore, at each instant predictions are made for the next 8 samples.

At time t=200, a disturbance u^d of step size 5 is applied to the process. In this case, we considered the alarm threshold to be at 3. The predictions are shown at different time instants in Figure 4.2a. At 201 sec the prediction first showed that the output will exceed the threshold at 213 sec; however, in order to have more confidence, the alarm was issued at 204 sec when three predicted values exceeded the threshold. In Figure 4.2b the closed loop process measurement validates the predicted system response. The measured output exceeded the threshold at 213 sec. If an alarm was generated solely based on the process measurement, the earliest an alarm can be issued is at 213 sec. The proposed scheme generate alarms 9 sec early compared to conventional alarm generation.



(a) Predicted output over horizon at t=204



(b) Measured Output with threshold

Figure 4.2: Predictions over horizon at the time of alarm generation and process measurement

4.1.2 Steady State Alarm Generation

Given the current steady state conditions, process gain, and the safety limits, the proposed scheme checks whether the controller has enough capacity to keep the process within the safety band. At steady state the following relationship exists between the input and output

$$\Delta y^{ss} = 1 * \Delta u \tag{4.4}$$

where Δy^{ss} is the change in measured output at steady state, Δu is the maximum available capacity of the input and steady state gain is 1 for this process. High and low limit values for output are 2 and -2, respectively, whereas for the input variable capacity varies from -7 to 7. Thus, inequality constrains for this process can be rewritten in input space as

$$-2 - y_{i,ss} \leq \Delta u \leq 2 - y_{i,ss} \tag{4.5}$$

$$-7 - u_{j,t} \leq \Delta u \leq 7 - u_{j,t} \tag{4.6}$$

Two different disturbance scenarios were simulated to check the steady state alarm conditions. In the first scenario a disturbance of step size 10 is introduced to the system at $t=$ 200. Also given that the steady state values for input and output at t=200 are $u_{200}-1$ and y_{200}=1. The steady state value for the process at any instant can be predicted using Equation 3.5 which gives the open-loop steady state value for output, y_i^{ss} =11. Substituting these values in Equations 4.5 and 4.6, we get the following inequality constraints arising from output and input limitations

$$-13 \leq \Delta u \leq -9 \tag{4.7}$$

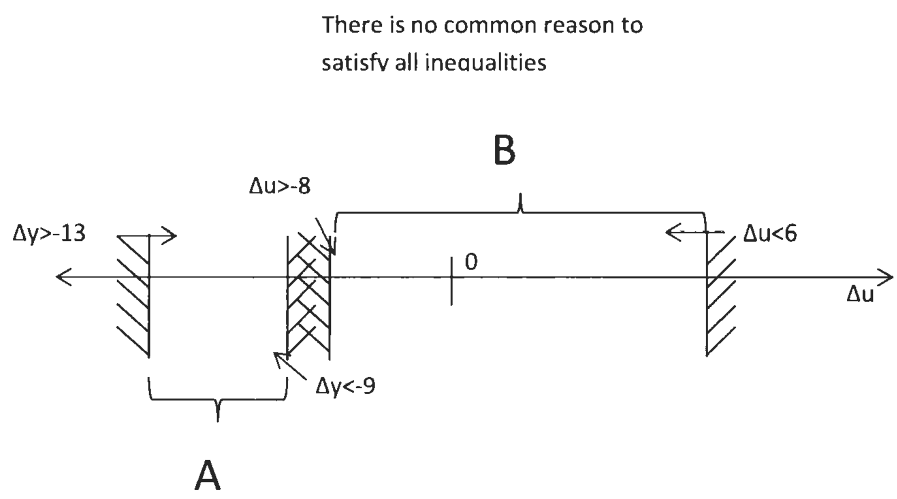


Figure 4.3: Constraints inequalities for the first scenario.

$$-8 \leq \Delta u \leq 6 \tag{4.8}$$

which are plotted in Figure 4.3. Clearly, it shows that there is no common space between these two inequalities ($A \bigcap B = 0$). So a feasible solution does not exist. It is easy to graphically plot and visualize the feasibility in a simple system; however for complex systems, it is not always possible to graphically represent the inequalities. In such a case linear programming (LP) can be used to check the existence of a feasible solution. For example, in this case the LP algorithm could not find a feasible solution as well confirming that there is not enough capacity in the actuator to bring the output within the limit. Therefore, an alarm will be issued at t=200.

For the second scenario, a disturbance of step size of -5 is applied to the system at t=600. Measured output and input at t=600 are y_{600}= 3, u_{600}=1 respectively. The predicted steady state value y_{ss} = -4. The inequality constraints for output and input

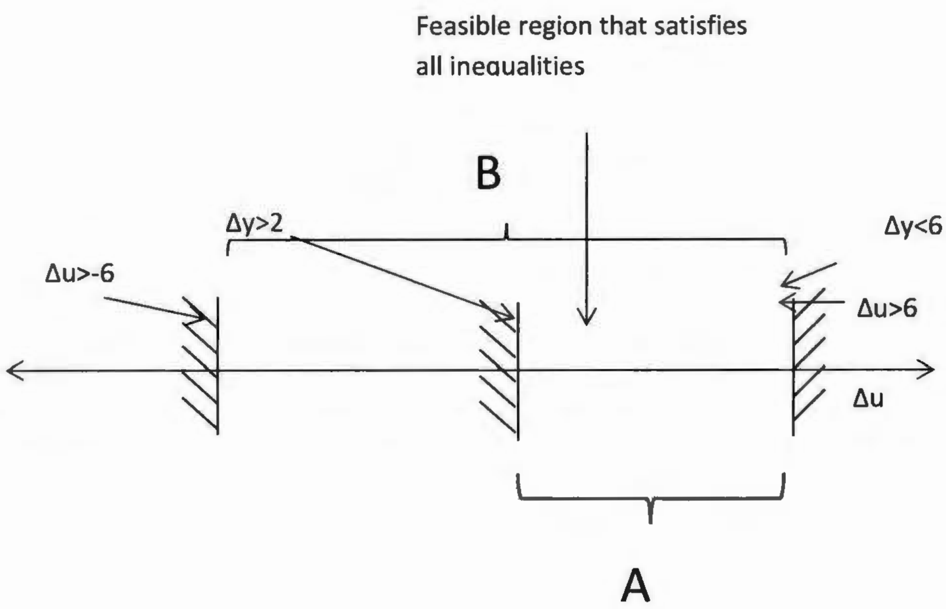


Figure 4.4: Constraints inequalities for the second scenario.

after expressing in the input space are

$$2 \leq \Delta u \leq 6 \tag{4.9}$$

$$-6 \leq \Delta u \leq 6 \tag{4.10}$$

which are plotted in Figure 4.4, which shows that there is a feasible solution so, $(A \bigcap B \neq 0)$. Therefore, no alarm will be issued at this instance.

These results are verified in Figures 4.5a and 4.5b, which show the closed loop process responses for these two scenarios. Figure 4.5a shows that the output remains outside the limit at steady state. The measured signal crosses the threshold at 350 sec. Therefore, based on the conventional method an alarm will be issued at 350 sec, whereas using the predictive approach the alarm will be issued at t=200 sec. Conversely, Figure 4.5b shows that the method is robust to false alarm; it does not issue an alarm when the controller is able to wither away the disturbance effect in this case.

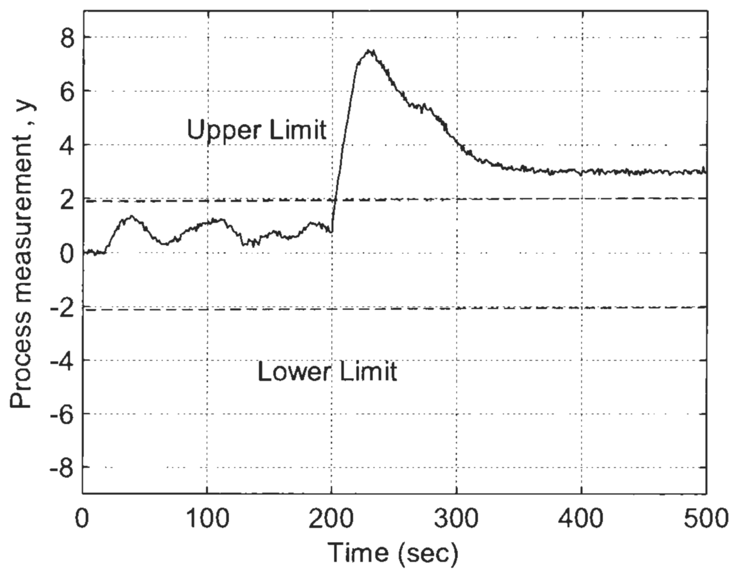
4.2 A MIMO Example

The proposed predictive alarm protocol is applied to a continuous stirred tank heater (CSTH) presented in [Thornhill et al., 2008]. The model is built using dynamic equations as well as experimental data of a pilot plant located in the Department of Chemical and Material Engineering at the University of Alberta. An equivalent simulink model for the plant is available. Even though it is a simulated model it is very real life like as it uses sensor noise obtained from real sensors. In this work the simulink model is used as the process plant considering that, the dynamic behaviour of the model will be close to the actual process.

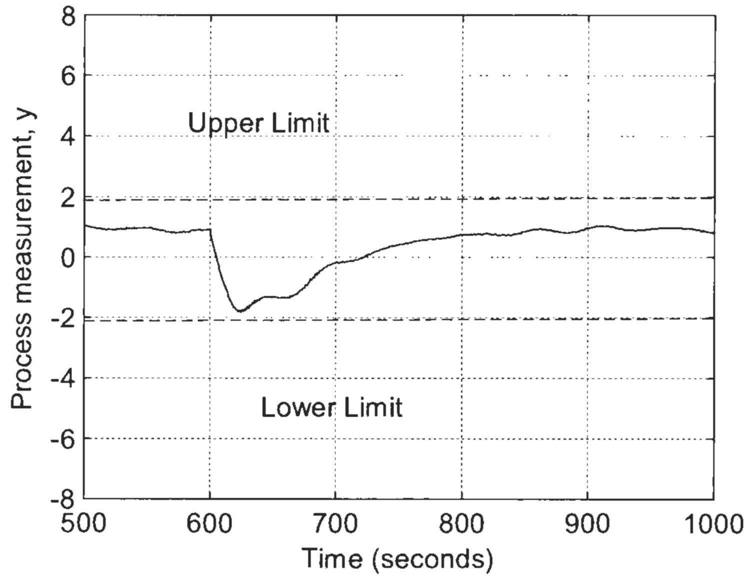
Figure 4.6 shows the schematic diagram of the CSTH plant. A steam and hot water supply is used to heat the cold water in a tank. A continuous flow of water comes from the cold water supply. The process dynamics of the plant are discussed in detail in [Thornhill et al., 2008]. The flow of steam, cold water and hot water can be manipulated using control valves. System can be represented by following Equation

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} G_{11}(s) & 0 \\ G_{21}(s) & G_{22}(s) \end{bmatrix} * \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} D_1(s) \\ D_2(s) \end{bmatrix} * u^d \tag{4.11}$$

where, y_1 is the level, y_2 is the temperature, u_1 is the cold water valve position, u_2 is the steam valve position, and u^d is the hot water valve position. Standard operating points for which the simulink model is developed, are stated in Table 4.2. A



(a) Closed loop process measurement where disturbance enters at t=200



(b) Closed loop process measurement where disturbance enters at t=600

Figure 4.5: Simulated results of process variable measurement with limit values for steady state

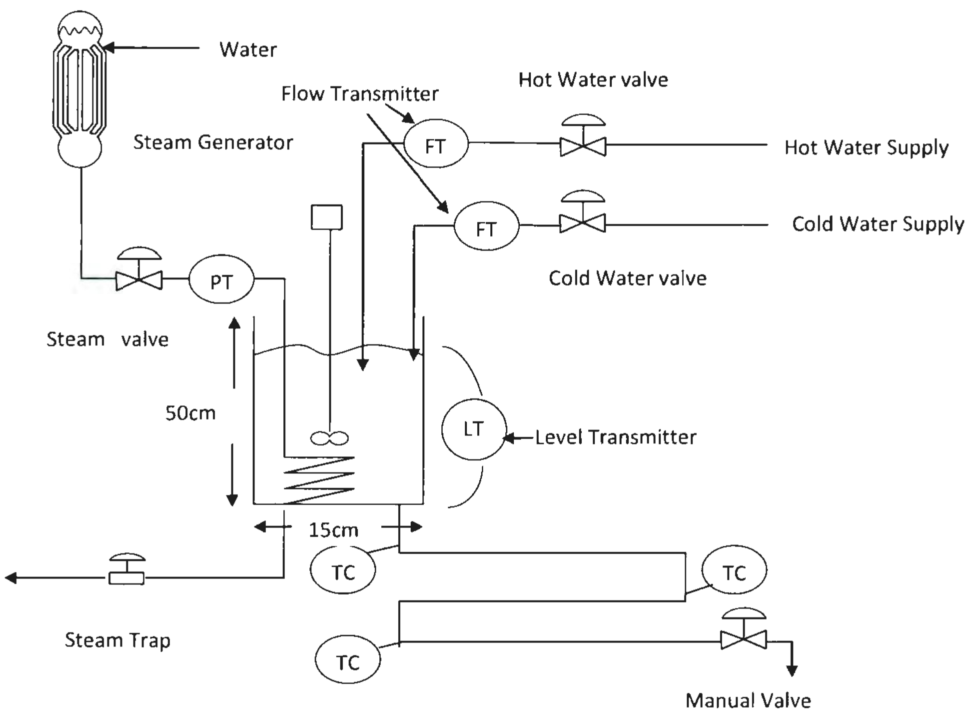


Figure 4.6: Schematic Diagram of the CSTH plant

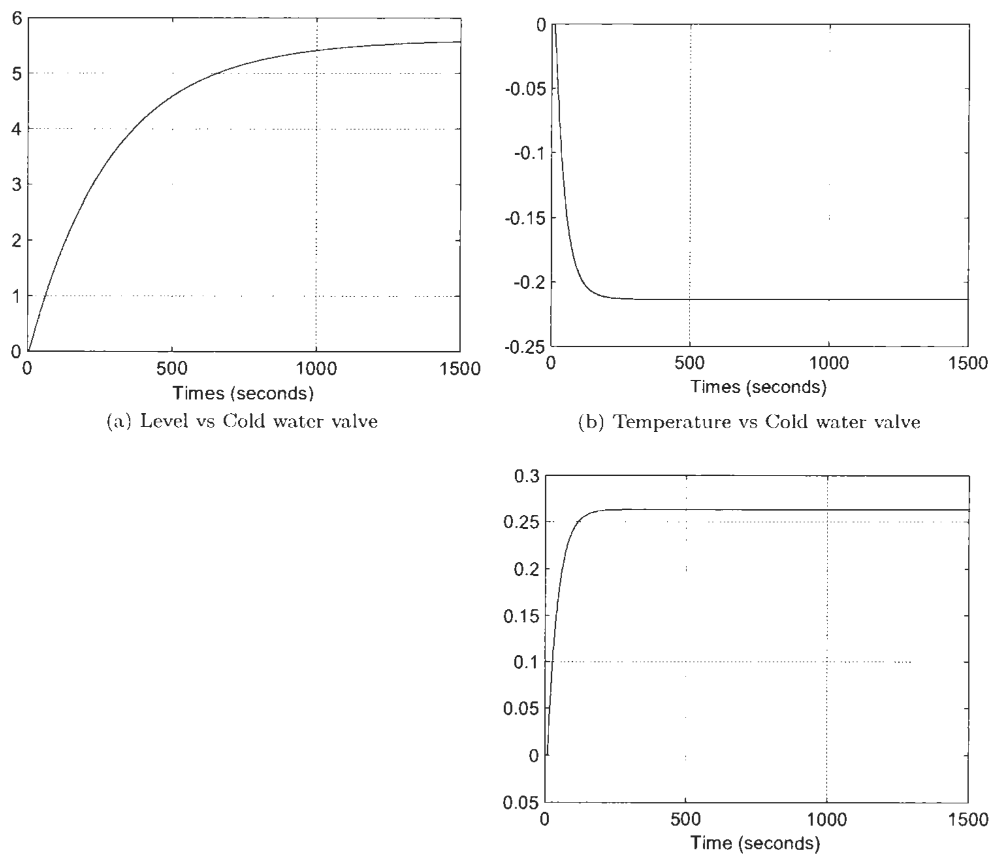
Table 4.2: Operating points of CSTH for predictive monitoring

Variable	Op Pt
Level/cm	20.50
Temperature/Deg C	42.50
CW valve/percent	17.64
Steam valve/percent	9.77
HW valve/percent	7.14

Model Predictive controller is implemented using those step response models. Design of the controller will be discussed in details elsewhere. This study will be limited in detection of an abnormal situation and early generation of alarm. For this study hot water valve position is used as disturbance as it is not being connected with any of the controller outputs. Initially hot water valve is set at standard operating point. For generating faulty condition different step change in the hot water valve is performed as disturbance input. Disturbance models is also generated using a step change in hot water valc position. Unit step response models for process inputs and outputs are shown in Figure 4.7 and unit step response for the disturbance input and process output is shown in Figure 4.8. Step responses shown here is in mA unit. Using the calibration curve in the cited literature, it can easily be found the model gain when the variables are express in their conventional unit (e.g. cm, deg. c). Water level of the tank and water temperature in the tank are two measured outputs of the system. The DMC manipulates the steam valve and cold water valve to control the water level and water temperature of the tank. Design parameters of the DMC controllers are provided in 4.3.

Table 4.3: Design parameters of the DMC controller

Variable	Value
Prediction horizon(p)	15
Control horizon (m)	5
Weighting factor (Q)	1

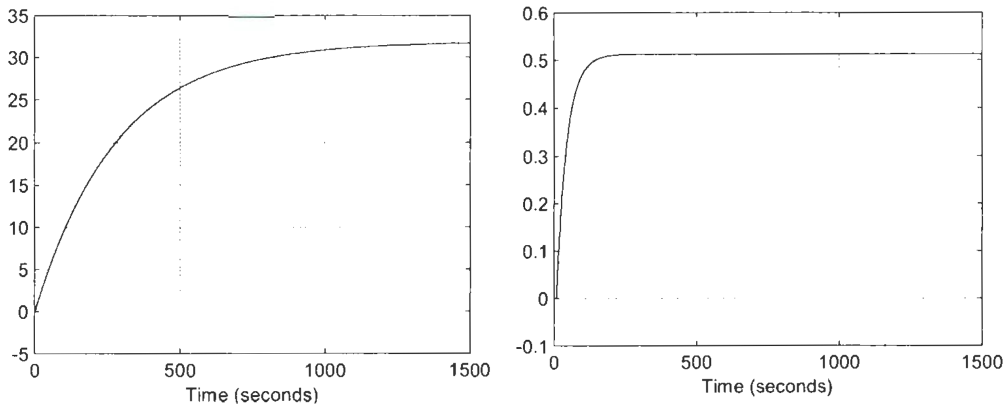


(a) Level vs Cold water valve

(b) Temperature vs Cold water valve

(c) Temperature vs Steam valve

Figure 4.7: Step response models between the Process outputs and inputs



(a) Level vs Hot water valve

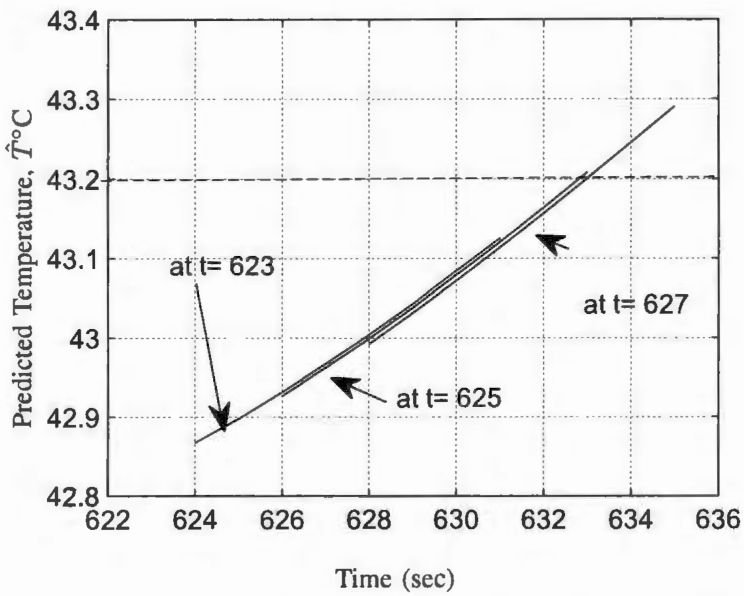
(b) Temperature vs Hot water valve

Figure 4.8: Step response models between the Process outputs and disturbance input

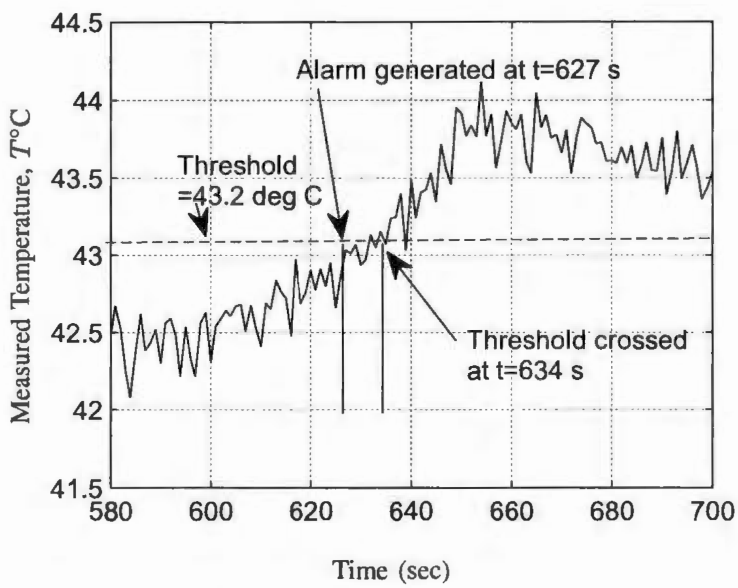
4.2.1 Dynamic Alarm Generation

Several disturbance scenarios were simulated by making changes to the hot water valve position. The monitoring system was used to monitor the process for these abnormal conditions. Here, we report two such scenarios.

At time t=600 s, the hot water valve is opened from 7 % to 8 %. Both output variables start to increase from their set-point with the introduction of the disturbance. Due to process time delay, the disturbance starts to affect the process measurements at t=608 s. However, level and temperature predictions are continuously monitored over the monitoring horizon, which, in this case are the next eight samples at every time instant. The prediction model used for the alarm generation contains the process model as well as the disturbance model relating hot water flow to level and temperature. The predictions are also corrected for bias at every second based on actual measurements. Since the monitoring scheme contains the disturbance model, as soon as the disturbance entered the system the prediction showed the effect of the disturbance on the output variables. In this case, we considered the alarm threshold



(a) Predicted output over horizon at different time



(b) Measured Output with threshold

Figure 4.9: Predictions at different time and process measurement in the dynamic state

to be at 43.2°C. The predictions are shown at different time instants in Figure 4.9a. At 625 sec the prediction showed that the output will exceed the threshold at 633 sec. However, in order to have more confidence, the alarm was issued at 627 sec when three predicted values exceeded the threshold. In Figure 4.9b the closed-loop process measurement shows that the measured output exceeded the threshold at 634 sec. If an alarm was generated solely based on the process measurement, the earliest an alarm could be issued is at 634 sec. The proposed scheme gave a 7 sec lead time to the operator. In this case, the proposed alarm system issues the alarm 7 seconds earlier than an alarm system based on a measured signal, which gives the operator time to react and take corrective action.

4.2.2 Steady State Alarm Generation

The outputs, level (y_1) and temperature (y_2), and inputs steam valve position (u_1) and cold water valve position (u_2), give rise to four constraints. The output constraints arise from the safe operational consideration of the process system, and the input constraints are due to the limited capacities of the valves. In addition to these constraints there also exists the input-output relationships arising from the steady state process gain. Equation 4.12 gives the input-output relationship for the CSTH system at steady state, which are calculated from the steady state gain of the step response of the transfer functions of the system

$$\Delta y_1^{ss} = 2.766\Delta u_1 \tag{4.12a}$$

$$\Delta y_2^{ss} = -0.293\Delta u_1 + 0.369\Delta u_2 \tag{4.12b}$$

where Δy_1^{ss} is the change in level, Δy_2^{ss} is the change in temperature, Δu_1 is the change in cold water valve position and Δu_2 is the change in steam valve position.

The high and low limit values for the level are defined as 25.2 cm and 15.8 cm, respectively. For temperature, high and low limits are 43.2°C and 39.2°C respectively. For both the cold water valve position and steam valve position, high and low limit values are selected as 19% and 0% respectively. Using these values, four inequality constraints for the system, can be written as in Equations 4.13a to 4.13d.

$$15.8 - y_1^{ss} \leq \Delta y_1^{ss} \leq 25.2 - y_1^{ss} \tag{4.13a}$$

$$39.2 - y_2^{ss} \leq \Delta y_2^{ss} \leq 43.2 - y_2^{ss} \tag{4.13b}$$

$$0 - u_{1,t} \leq \Delta u_1 \leq 19.05 - u_{1,t} \tag{4.13c}$$

$$0 - u_{2,t} \leq \Delta u_2 \leq 19.05 - u_{2,t} \tag{4.13d}$$

For the first disturbance scenario the hot water valve position is changed from 7.1% to 7.6% percent at $t=$ 600 s. This change of hot water valve position causes a rise in both the level and temperature of the water from their nominal values 20.5 cm and 42.5°C, respectively. The steady state value of the process variables are predicted using Equation 3.5. For this scenario, the predicted open-loop steady state value of the process variables are, $y_1^{ss} = 50$ cm, $y_2^{ss} = 42.85$°C. Also, the input values at t=600 s are $u_{1,600} = 17.95\%$ $u_{2,600} = 9.79\%$. Substituting these values in Equation 4.13 we

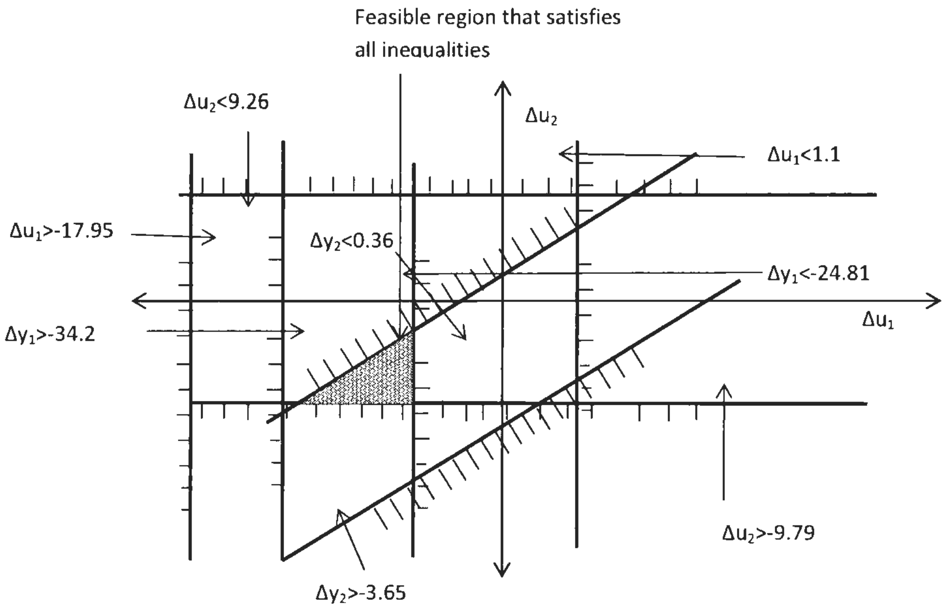


Figure 4.10: Constraints inequalities for the first scenario.

get the following output and input constraints

$$-34.2 \leq \Delta y_1 \leq -24.81 \tag{4.14a}$$

$$-3.65 \leq \Delta y_2 \leq 0.36 \tag{4.14b}$$

$$-17.95 \leq \Delta u_1 \leq 1.1 \tag{4.14c}$$

$$-9.79 \leq \Delta u_2 \leq 9.26 \tag{4.14d}$$

These inequalities, together with the steady state input output relationships described in Equations 4.12a and 4.12b, are used to check for feasibility for Δu_1 and Δu_2. At every instant, the LP algorithm checks whether there is a feasible solution for these constraints. For this disturbance scenario, the LP is able to find a feasible solution, therefore, no alarm is issued. This is also depicted in Figure 4.10 where the feasible region that satisfies all four constraints simultaneously is shown by the hatched area.

Therefore, the process will be at no alarm state at $t = 600$ s despite of the disturbance being present. This is also supported by the actual closed-loop measurements which show that both level and temperature do not exceed the alarm limits for the above disturbance scenario (Figure 4.11).

The second disturbance scenario is similar to scenario 1, except a bigger step size was considered. The hot water valve position is changed from 7.1% percent to 9.5% percent at $t = 800$. The consequence of the introduction of this disturbance is the same as the previous scenario with a greater intensity as the disturbance size is larger. For this scenario, the predicted open-loop steady state values of the process variables are $y_1^{ss} = 50$ cm, $y_2^{ss} = 44.21°$C. and the input values at 800 s are $u_{1,800} = 17.95\%$, $u_{2,800} = 9.79\%$ respectively. Using these values output and input constraints for the process can be written as

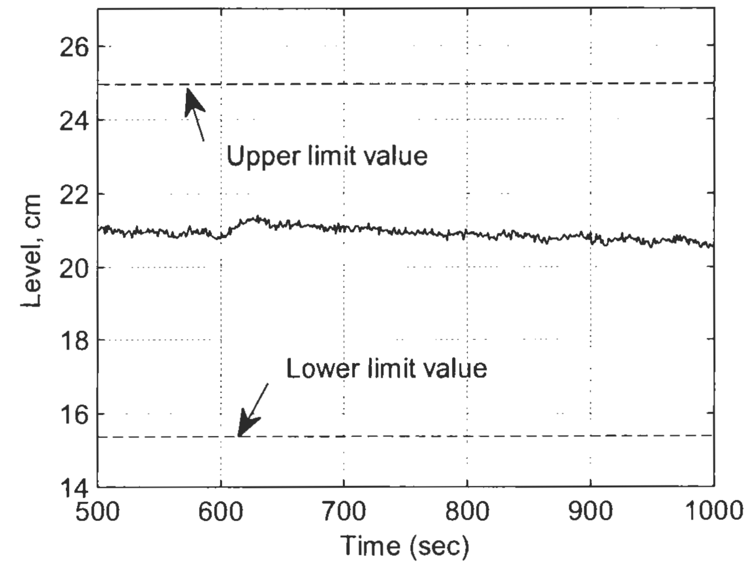
$$-34.2 \leq \Delta y_1 \leq -24.81 \tag{4.15a}$$

$$-5.01 \leq \Delta y_2 \leq -1.01 \tag{4.15b}$$

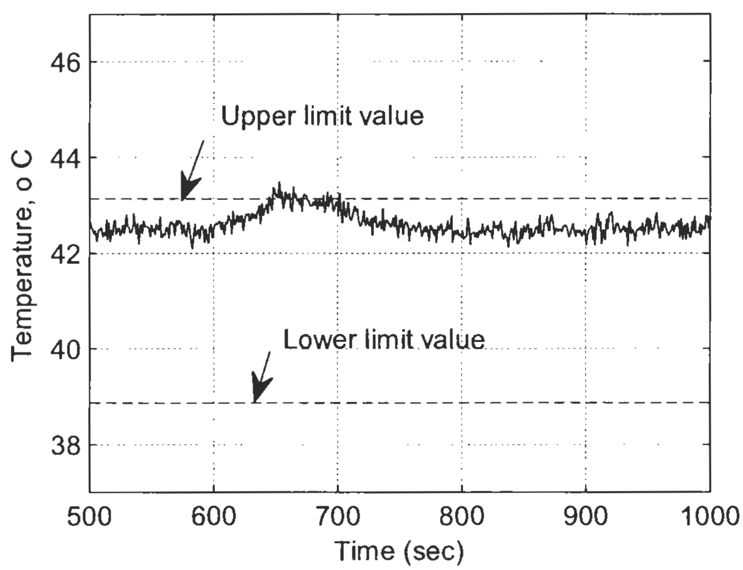
$$-17.95 \leq \Delta u_1 \leq 1.1 \tag{4.15c}$$

$$-9.79 \leq \Delta u_2 \leq 9.26 \tag{4.15d}$$

For the given conditions, a feasible solution does not exist that meets all the constraints. This is also depicted in Figure 4.12, which shows there is no feasible region for the given conditions. Therefore, an alarm is issued at the time the disturbance is measured (at t=800 s). Closed-loop process measurements for this particular scenario are presented in Figure 4.13, which validates that the tank temperature exceeds the alarm limit at 825 sec.



(a) Level measurement with target limit



(b) Temperature measurement with target limit

Figure 4.11: Simulated results of level and temperature measurement with limit value for Scenario 1

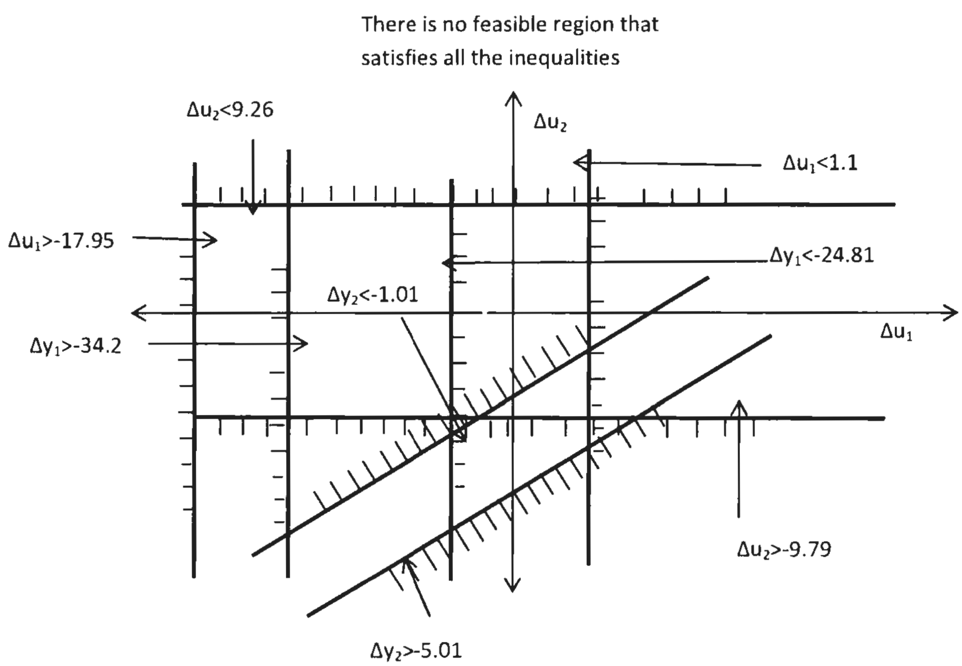
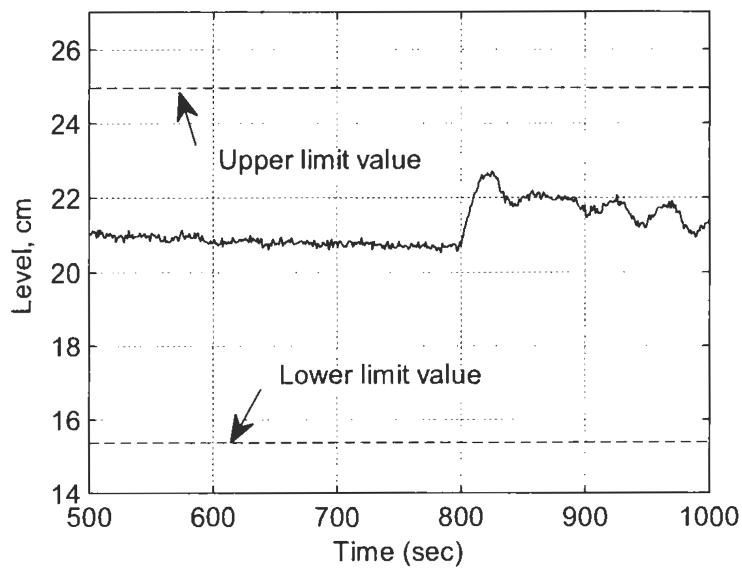
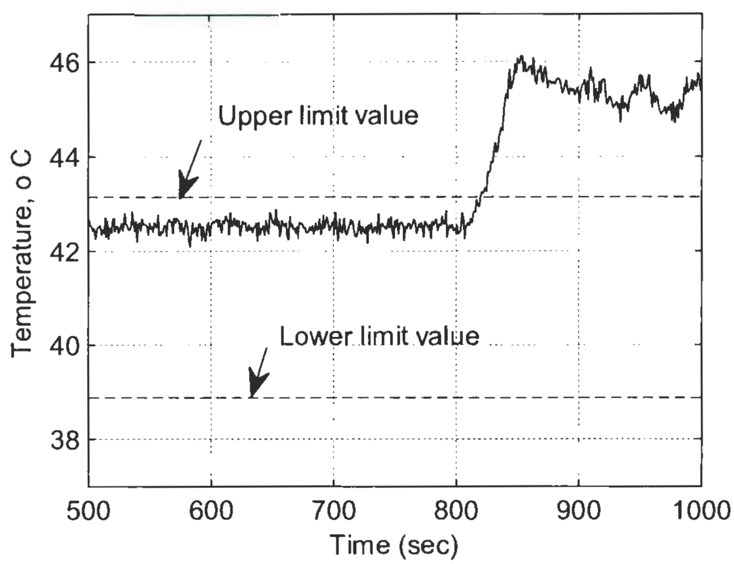


Figure 4.12: Constraints inequalities for the second scenario.



(a) Level measurement with target limit



(b) Temperature measurement with target limit

Figure 4.13: Simulated results of level and temperature measurement with limit value for Scenario 2

Chapter 5

Conclusions of Predictive Early Warning Generation

The aim of the current study was to develop a methodology to generate an early warning to the system. Predictive power of process open-loop model was used to design a predictive alarm system, through which the objective was achieved. In proposed methodology, alarms were generated based on the predictive signals which is capable of provide an alarm earlier to compared to conventional systems that issue alarm based on process measurement. The performance of the methodology was validated using extensive simulation study. The method also proved to be robust and did not generate nuisance alarm when process was at normal operating state. The specific contributions are listed below:
In this part of the study, the following contributions have been achieved.

- A novel model-based predictive alarm generation technique is proposed. The proposed methodology uses an open-loop process and disturbance model to predict the system responses. The open-loop responses are bias-corrected using the available measurements. The alarms are generated based on the updated predictions.

- Two limiting conditions were postulated, and based on these two conditions, alarm generation methodologies were developed. The dynamic alarm generation procedure looks at the immediate effect of the disturbance and the steady state alarm generation procedure monitors the process for long term effect of a disturbance in a process.

- The proposed method provides early alarm compared to any conventional alarm generation method. The nethod is relatively maintenance free, it does not require frequent updating as it uses open-loop predictions.

- The technique is robust as it exploits the fundamental limitations of the controller and actuator for alarm generation. The methodology has been applied to a SISO system and a more complex MIMO system where technique was used to monitor the system for different disturbance scenarios. In both examples, the methods generated alarms in a consistent manner and demonstrated robustness to the false alarms.

5.1 Future Recommendations

- Experimental validation : The effectiveness of the current study can be further demonstrated using an experimental set up (i.e., experimental CSTH set up).

- Use of a quantitative value for alarm generation: In the predictive monitoring section, an alarm is generated based on a heuristic procedure. When three or more values lie above the threshold, an alarm is generated. This can be quantified using the risk value and risk threshold, which provide more scope to prioritize the response of alarms.

- Use of multi-linear model: The current study used linear model for a fixed operating points. However, in practice, with the change of the operating point process dynamics may change drastically. Therefore, effectiveness of the proposed methodology will be compromised in case of a system where operating points change regularly. Multiple models can be used for different operating points to deal with this problem.

Simulation studies performed in this study shows a good prospect of proposed alarm generation method. As our simulation study was based on a lump model with sensor noise, plant model mismatch, these results supports to the fact that, this methodology may be useful for the process industry. The method requires more experimental validation in pilot plant before it can be implemented in industry.

Part II:

A Comparative Study between PID-free MPC and Hybrid Control Structure

Chapter 6

Predictive Control

6.1 Introduction

Model predictive controllers are typically used as a supervisory layer over the base level PID controller, especially in large-scale applications. This structure gained acceptance mainly because it allows the implementation of MPC with minimal changes to the existing control structure. Also, the PID layer can act as a fall back when the MPC is turned off for any reason. However, this structure does not allow the potential benefits of the MPC to be fully harnessed. In practice, it was observed that there are many incentives in breaking the PID loop and directly manipulating the valve output using the MPC. One common example is when trying to use the full valve capacity (e.g., maximize feed, maximizing cooling) it is common practice to break the PID loop and manipulate the valve directly from MPC. Also when multiple feed forward affects a process variable, it is common to replace the PID loop with MPC.

Recently, a software called ADMC from the original inventors of DMC is being marketed that uses the DMC to directly manipulate the actuator. It is claimed that this controller performs better than the hybrid MPC-PID structure. Therefore, an objective investigation of the performance of these competing control structures is necessary. In this study, a simulation-based comparative study is carried out between two control structures: MPC cascaded to PID and MPC directly manipulating the valve output.

6.2 Literature Review

6.2.1 Current State of PID Controller

PID is a widely used control structure in the industry. Desborough and Miller estimated that 98 percent of the controllers in a median chemical plant are PID controllers [Desborough and Miller, 2001]. Though it is widely used for its simplicity of implementation, it has different limitations. The main limitation of the PID is that it has no straightforward tuning method. The impact of this fact is evident from the result reported by Van Overschee and De Moor [Overschee and Moor, 2001]. They summarized that 80 percent of industrial PID controllers are poorly tuned; 30 percent of these PID loops operate in manual mode; and 25 percent of the PID loops in automatic mode operate under default factory settings.

[Na, 2001] proposed a control structure to overcome the drawbacks of the conventional PID controller with fixed tuning parameters. The proposed control structure is presented in Figure 6.1. In this arrangement, PID gains are automatically tuned in order to keep a predefined cost function to a minimum. MPC is applied to minimize the cost function using the second order linear model. The proposed methodology is applied to a linear model for nuclear steam generators. The applied methodology showed an improved performance compared to that of PID in both set point tracking

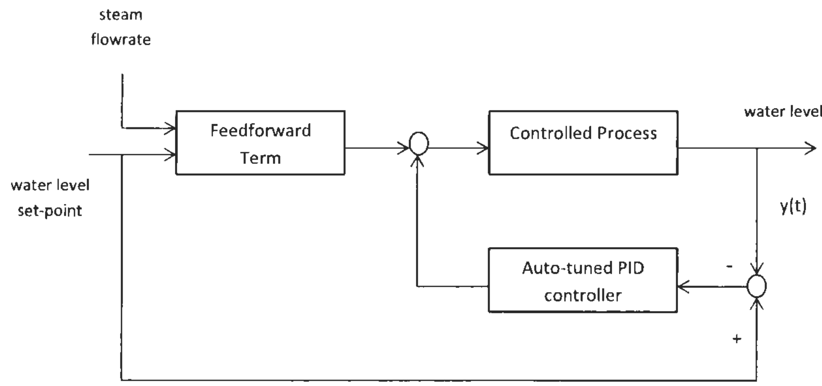


Figure 6.1: Structure of MPC based auto-tuned PID [Na, 2001]

and regulatory control.

A simple but robust technique is described in [Ogunnaike and Mukati, 2006]. In this work, the simplicity of PID and versatility of MPC are combined together. The output is predicted over the horizon using a first order plus dead time (FOPDT) model. Thus, predictive features of MPC are retained. Use of FOPDT ensures that prediction requires as few parameters as the PID controller design required. Though simpler modelling approaches have been considered, there always exists a modelling error. This modelling error is minimized, making a bias correction at each step comparing the predicted output and the actual measurements. Tuning parameters are defined based on the key performance indices such as set point tracking, disturbance rejection, and the robustness and aggressiveness of the controller. The controller showed better performance in set point tracking and disturbance rejection compared to an IMC-tuned PID controller in extensive simulation studies.

[Astrom and Hagglund, 2001] describes the potential alternatives for PID in industrial settings. The proposed alternatives are a discrete-time linear MISO controller, state feedback and observers (SFO) and model predictive controller (MPC). Fuzzy control is also mentioned as a potential alternative. All alternatives provide an improved performance, especially for systems which are poorly damped. Controllers based on SFO require a greater modelling afford, as such its use is justified only when modelling efforts are moderate. MPC is typically used as a supervisory layer to the base layer PID. The use of MPC provides a drastic improvement of set point tracking. Moreover, computational complexity is minimized in this case, as MPC executes at a slower rate, regulating the slower dynamics of the system. The PID layer acts with the fast interactions.

[Pannocchia et al., 2005] proposed an offset-free constrained linear quadratic (CLQ) controller as a potential candidate to replace PID. CLQ consists of three main modules based on a state-space model of the system: a state and disturbance estimator, a constrained target calculation module, and a constrained dynamic optimizer. Each module is designed to minimize the computational load and, as such, the controller implementation load is comparable to a PID controller. The CLQ controller outperformed the PID controller in all the simulated cases reported in the paper. The controller was limited to SISO systems, however, it may be expendable for MIMO systems.

[Han, 2009] described active disturbance rejection control (ADRC) as an improved control scheme to replace PID. ADRC is error driven similar to PID, using a state observer to utilize the power of non-linear feedback. The major limitations of PID pointed, are error computation, noise degradation in derivative control, oversimplification of control law and complications from the integral control. ADRC is aimed at overcoming these PID limitations. First, a simple differential equation is used to gen-

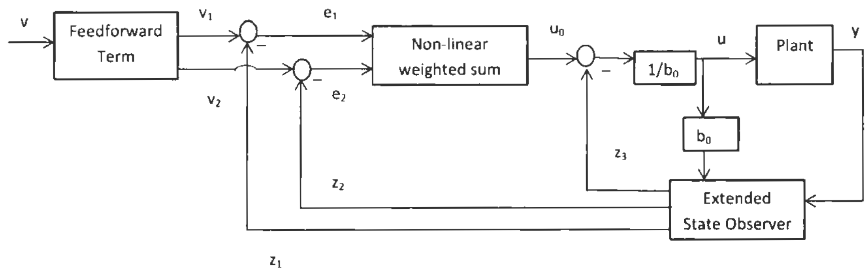


Figure 6.2: ADRC control structure [Han, 2009]

erate a transient trajectory. A differentiator is capable of noise-tolerant tracking. A non-linear control law is used instead of a simple weighted error. An ADRC structure is shown in the Figure 6.2.

Though various controllers have been proposed as an alternative to PID controllers, MPC has tprobably the most potential to replace a portion of the PID controllers in process industry. In the following sections, a historical account of MPC is given and some of the articles that compared MPC with PID, are reviewed.

6.2.2 Historical Review of MPC

MPC has been widely used as an advanced control strategy in the process industry over the last 30 years. The theory of MPC, however, existed long before it was implemented in the process industry. The modern control concept was first developed through the work of Kalman in the early 1960s. In his work, an objective function was minimized that penalized expected values of squared input and state deviation from origin. The solution to this problem was known as a linear quadratic Gaussian (LQG) controller. However, LQG had little impact on control technology development in the process industry. The significant reasons for this failure are cited in [Richalet et al., 1976] and [Garcia et al., 1989].

The failure of LQG led to the development of a more general model based control methodology where dynamic optimization problems were solved on-line at each control execution. Process inputs were computed optimizing future process behaviour over a time interval which is later defined as 'prediction horizon'. Explicit plan models are used to describe the plant dynamics. Parallel to this development, new identification methodologies also emerged which helped to propel the technology. The model based control, together with the industrial process modelling, was referred as MPC technology. MPC technology is first described in [Richalet et al., 1976] and later summarized in [Richalet et al., 1978]. They summarized their approach as model predictive heuristic control and provided solution software using this approach.

Curtler and Ramaker first described the dynamic matrix control (DMC) algorithm in [Cutler and Ramaker, 1979] and [Cutler and Ramaker, 1980]. In a companion work [Prett and Gillette, 1980], Prett and Gillete described an application of DMC for an industrial case. with a modified algorithm capable of handling the non-linearity.

During the 1980s both industrial and academic interest in MPC started to grow. The idea of cost function and optimization is employed with the algorithm. Thus, the MPC algorithm is further modified based on the cost function and optimization of the process. According to the different types of cost functions MPC can be classified into different types; DMC with least squares satisfaction of input constraints [Prett and Garcia, 1988], DMC with constrained linear programming optimization [Morshedi et al., 1985], and Quadratic programming solution of DMC [Garcia and Morshedi, 1986].

The understanding of MPC has now reached a new level and a specific framework has been built for both theoretical and practical purposes.

Some excellent reviews on MPC technology can be found in [Muske and Rawlings, 1993],[Bemporad and Morari, 1999] and [Qin and Badgwell, 2003] . In these papers, a generalized algorithm of the model predictive controller is provided. Moreover, a historical review of the MPC technology and the algorithms of the MPC at different stages are also presented. These papers also discuss the strengths, limitations and evolution of the MPC algorithms in detail.

6.2.3 Comparative Study between MPC and PID

A comparative study between standard PID and predictive controller is presented for a heat exchanger in [Bonivento et al., 2001]. In this work, the modelling of the heat exchange is performed using the dynamic properties of the heat exchanger. The step response model for the input-output is identified. Then, using the identified model, PID controller and Generalized Predictive Control (GPC) are designed. GPC provides better results compared to standard PID for both set-point tracking and disturbance rejection.

Another work on MPC implementation in heat exchanger is presented in [Krishna Vinaya et al., 2012]. For heat exchanger control, PID is a widely used control technique. The motivation of the work was to optimize and conserve energy. A heat exchanger is highly non-linear process. In this work, the system is divided into different zones and for each zone a model is developed. Using these models, a model predictive controller was designed and used to control the temperature of a fluid stream. A PID controller was also designed for the process using the Ziegler Nichols method. Comparative studies on the two controllers' performance show that MPC provides better results based on the rise time, overshoot and settling time.

A comparative study of PID controllers, MPC controllers and model free adaptive controllers (MFA) is performed in [Lukacova and Borzikova, 2010]. In this work, MFA is designed using an artificial neural network,and MPC is designed based on conventional dynamic matrix controller (DMC). The results show that PID is the fastest of the three controllers but it has overshoot and steady state error. Both MFA and MPC are steady state error-free. MFA tracks the set point faster than MPC, but MFA has overshoot. However, advanced control strategies provide superior performance compared to PID.

The above literature survey shows, even though there were several studies to evaluate the performance of MPC against PID controller, there was no effort to compare hybrid MPC-PID structure with PID-free structure. This study is aimed to fill in this gap.

Chapter 7

Theory of Dynamic Matrix Control

7.1 Dynamic Matrix Control

In the present work, dynamic matrix control (DMC)is used as a representative MPC algorithm. DMC has been a widely used algorithm in the industry since its introduction back in 1980s. The theory of DMC is available in several books and literatures such as [Ogunnaike and Ray, 1994] and [Seborg et al., 1989]. Discussions in this chapter will be limited to the representation of the central idea of DMC algorithm. Initially, the algorithm is presented for a simple SISO system. Later, the algorithm is extended to MIMO system.

The DMC algorithm is executed in two stages: prediction and control. In the prediction stage, the process variable is predicted using the receding horizon algorithm. In the control stage, an objective function is defined and minimized to get the control actions.

7.1.1 Prediction

A step response model of the plant is used for prediction in this formulation, while the disturbance is considered to be constant along the horizon. A step response model for a SISO system can be written as in Equation 7.1

$$y_t = \sum_{i=1}^{\infty} a_i * \Delta u_{t-i} \tag{7.1}$$

where, y_t is the model output, a_i is the i-th coefficient of the step response, and Δu are the past input changes step. Using the time-shifting property and taking the constant disturbance into account, a future predicted value can be written as in Equation 7.2

$$\hat{y}_{t+k} = \sum_{i=1}^{\infty} a_i * \Delta u_{t+k-i} + \nu_{t+k} \tag{7.2}$$

where, $\hat{y}_{t+k}$ is the predicted output at time $t+k$, ν_{t+k} is the disturbance at time $t+k$. As the disturbance is assumed to be constant over the horizon, it is given by Equation 7.3

$$\nu_{t+k} = \nu_t = y_m(t) - \hat{y}_t \tag{7.3}$$

where $y_m(t)$ is the measured output at time t. The value of the ν_{t+k} from 7.3 can be replaced in Equation 7.2 and can be rewritten in the following form

$$\hat{y}_{t+k} = \sum_{i=1}^{k} a_i * \Delta u_{t+k-i} + \sum_{i=k+1}^{\infty} a_i * \Delta u_{t+k-i} + y_m(t) - \sum_{i=1}^{\infty} a_i * \Delta u_{t-i} \tag{7.4}$$

Now, the last three terms of Equation 7.4 actually express the output of the system if no control action is taken from time t to $t+k$, and is termed as the free response of the system, y^*_{t+k}. The free response of the system thus can be expressed mathematically as follows

$$y^*_{t+k} = y_m(t) + \sum_{i=k+1}^{\infty} (a_{k+i} - a_i) * \Delta u_{t-i} \tag{7.5}$$

Now, if the process is asymptotically stable, the step response tend to be a constant value after N sampling time. Therefore, finite step response of N samples can be used instead of infinite step response model as, $a_{k+i} - a_i \simeq 0$ for $i > N$. Using this finite step response model, free response of the system can be represented as,

$$y^*_{t+k} = y_m(t) + \sum_{i=k+1}^{N} (a_{k+i} - a_i) * \Delta u_{t-i} \tag{7.6}$$

Using the free response of the system, Equation 7.7 can be rewritten in the following form

$$\hat{y}_{t+k} = \sum_{i=1}^{k} a_i * \Delta u_{t+k-i} + y^*_{t+k} \tag{7.7}$$

Equation 7.4 will be used to predict along the prediction horizon *(k=1, 2, ..., p)* with m control actions.

$\hat{y}_{t+1} = y^*_{t+1} + a_1 * \Delta u_t$

$\hat{y}_{t+2} = y^*_{t+2} + a_2 * \Delta u_t + a_1 * \Delta u_{t+1}$

....

....

$\hat{y}_{t+p} = y^*_{t+p} + \sum_{i=1}^{m}(a_i * \Delta u_{t+p-i})$

These calculated predicted values can be expressed in the following matrix form

$$\boldsymbol{\hat{y}} = \boldsymbol{y^*} + \boldsymbol{A} * \boldsymbol{\Delta u} \tag{7.8}$$

where, $\boldsymbol{\hat{y}}$ is a p dimensional vector containing the predicted output over prediction horizon, $\boldsymbol{y^*}$ is also a p dimensional vector which contains the free response of the system over the horizon, $\boldsymbol{\Delta u}$ is an m dimensional vector of control increments. $\boldsymbol{A}$ is the dynamic matrix of the system, which is defined in Equation 7.9

$$\boldsymbol{A} = \begin{bmatrix} a_1 & 0 & 0 & \dots & \dots & 0 \\ a_2 & a_1 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \\ a_m & a_{m-1} & a_{m-2} & \dots & \dots & a_1 \\ \dots & \dots & \dots & \dots & \dots & \\ a_p & a_{p-1} & a_{p-2} & \dots & \dots & a_{p-m+1} \end{bmatrix} \tag{7.9}$$

Equation 7.8 expresses the relation between the predicted future output with control increment. As such, this can be used to calculate the action necessary to achieve a specific system behaviour [Ogunnaike and Ray, 1994] [Seborg et al., 1989].

7.1.2 Control Algorithm

The objective of the DMC controller is to drive the output close to desired trajectory. An objective function is defined based on the deviation of desired trajectory and predicted output and it is minimized by calculating a set of control actions.

Suppose, a p dimensional vector $\boldsymbol{r}$ is known which contains the desired set-points over the prediction horizon p. The objective function $J(\Delta u)$ is defined in Equation 7.10 that calculates a set of control actions that minimizes the deviation between $\boldsymbol{r}$ and $\hat{\boldsymbol{y}}$

$$J(\boldsymbol{\Delta u}) = (\boldsymbol{r} - \hat{\boldsymbol{y}})^T * \boldsymbol{Q} * (\boldsymbol{r} - \hat{\boldsymbol{y}}) \tag{7.10}$$

where $\boldsymbol{Q}$ is the weighting matrix that defines the aggressiveness of the controller. However, the objective function defined in Equation 7.10 is an unconstrained formulation and may produce undesirable consequences. This is why the control action is also penalized along with the deviation of prediction and set-point. Thus, a more robust objective function can be defined as in Equation 7.11

$$J(\boldsymbol{\Delta u}) = (\boldsymbol{r} - \boldsymbol{\hat{y}})^T * \boldsymbol{Q} * (\boldsymbol{r} - \boldsymbol{\hat{y}}) + \boldsymbol{\Delta u}^T * \boldsymbol{R} * \boldsymbol{\Delta u} \tag{7.11}$$

where, $\boldsymbol{R}$ is the weighting matrix to penalize the control action. Control actions can be calculated by minimizing the objective function described in 7.11. The value of $\boldsymbol{\hat{y}}$ can be replaced from Equation 7.8. Control actions are calculated analytically by taking the first derivative of the objective function with respect to $\boldsymbol{\Delta u}$ and equating it to zero which gives the following explicit expression for $\boldsymbol{\Delta u}$.

$$\boldsymbol{\Delta u} = (\boldsymbol{A}^T \boldsymbol{Q} \boldsymbol{A} + \boldsymbol{R})^{-1} \boldsymbol{A}^T \boldsymbol{Q}^T * (\boldsymbol{r} - \boldsymbol{y}^*) \tag{7.12}$$

More compactly controller can be expressed as, $\boldsymbol{\Delta u} = \boldsymbol{K_c} * \boldsymbol{e}$, where, $\boldsymbol{K_c} = (\boldsymbol{A}^T \boldsymbol{Q} \boldsymbol{A} + \boldsymbol{R})^{-1} \boldsymbol{A}^T \boldsymbol{Q}^T$, and $\boldsymbol{e} = (\boldsymbol{r} - \boldsymbol{y}^*)$.

Thus, using DMC a set of control actions are calculated that drive the output close to the desired set-point over the predicted horizon. However, the total m number of control actions are calculated, but only the first control action is implemented, as at the next control interval, the calculation is repeated to get a new set of control actions.

7.1.3 Extension to Multi-variable Case

The scheme discussed in the previous subsections can be easily extended for a MIMO system. Basic equations will remain the same, except for the fact that the size of the vector and the matrices would be increased and partitioned. Based upon the linearity of the model, the superposition principle can be used to evaluate the predicted outputs.

For a multivariable system with s output and h input variables, the predicted output vector $\hat{\boldsymbol{y}}_{mm}$, free response vector $\boldsymbol{y}^*_{mm}$, set-point trajectory vector $\boldsymbol{r}_{mm}$ and array of future control signal $\boldsymbol{\Delta u}_{mm}$ can be written as [Ogunnaike and Ray, 1994],

$$\hat{\boldsymbol{y}}_{mm} = \begin{bmatrix} \hat{y}_{1,t+1} \\ \cdots \\ \hat{y}_{1,t+p} \\ \cdots \\ \hat{y}_{s,t+1} \\ \cdots \\ \hat{y}_{s,t+p} \end{bmatrix}, \boldsymbol{y}^*_{mm} = \begin{bmatrix} y^*_{1,t+1} \\ \cdots \\ y^*_{1,t+p} \\ \cdots \\ y^*_{s,t+1} \\ \cdots \\ y^*_{s,t+p} \end{bmatrix}, \boldsymbol{r}_{mm} = \begin{bmatrix} r_{1,t+1} \\ \cdots \\ r_{1,t+p} \\ \cdots \\ r_{s,t+1} \\ \cdots \\ r*_{s,t+p} \end{bmatrix}, \boldsymbol{\Delta u}_{mm} = \begin{bmatrix} \Delta u_{1,t} \\ \cdots \\ \Delta u_{1,t+m-1} \\ \cdots \\ \Delta u_{h,t} \\ \cdots \\ \Delta u_{h,t+m-1} \end{bmatrix}.$$

Dynamic matrix for the multivariable system is redefined as

$$\boldsymbol{A}_{mn} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \cdots & \cdots & \boldsymbol{A}_{1h} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \cdots & \cdots & \boldsymbol{A}_{2h} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ \cdots & \cdots & \cdots & \boldsymbol{A}_{ij} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \cdots & \cdots & \boldsymbol{A}_{sh} \end{bmatrix}.$$

Here, the overall dynamic matrix is constructed using submatrices which contain the step weights that relate the individual input-output pair. The sub-matrix relating the i-th output to j-th input can be defined as the same way a dynamic matrix is defined previously and is given below.

$$
\boldsymbol{A_{ij}} = \begin{bmatrix}
a_{ij,1} & 0 & 0 & \dots & \dots & 0 \\
aij,2 & a_{ij,1} & 0 & \dots & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots & \\
a_{ij,m} & a_{ij,m-1} & a_{ij,m-2} & \dots & \dots & a_{ij,1} \\
\dots & \dots & \dots & \dots & \dots & \\
a_{ij,p} & a_{ij,p-1} & a_{ij,p-2} & \dots & \dots & a_{ij,p-m+1}
\end{bmatrix}
$$

With these definitions, the DMC controller is implemented for a MIMO system.

Chapter 8

Simulation Results

8.1 Plant Description

Controllers based on the algorithm described in chapter 7, are designed for a continuous stirred tank heater (CSTH) mentioned in earlier chapter. As stated previously, though the plant is a simulated model, it is very real life like as dynamic equations along with experimental data, were used to build the simulink model. The available simulink model is considered as a plant for this study as it is assumed that the dynamic behaviour of the real plant will be similar to this simulated model.

In this set up, water is heated using steam and hot water. Cold water enters into the tank continuously from supply. Steam is supplied from a steam generator whereas hot water is supplied from building utilities. Control valves manipulate the flow of steam, cold water and hot water. The water level of the tank and the temperature of the water are the two controlled variables. These variables are controlled by manipulating the valve positions of the control valves. Standard operating points used to develop simulink model are stated in Table 8.1.

Table 8.1: Operating points of CSTH for different control structures

Variable	Op Pt
Level/cm	20.50
Temperature/Deg C	42.50
CW valve/percent	42.67
Steam valve/percent	40.81
HW valve/percent	0

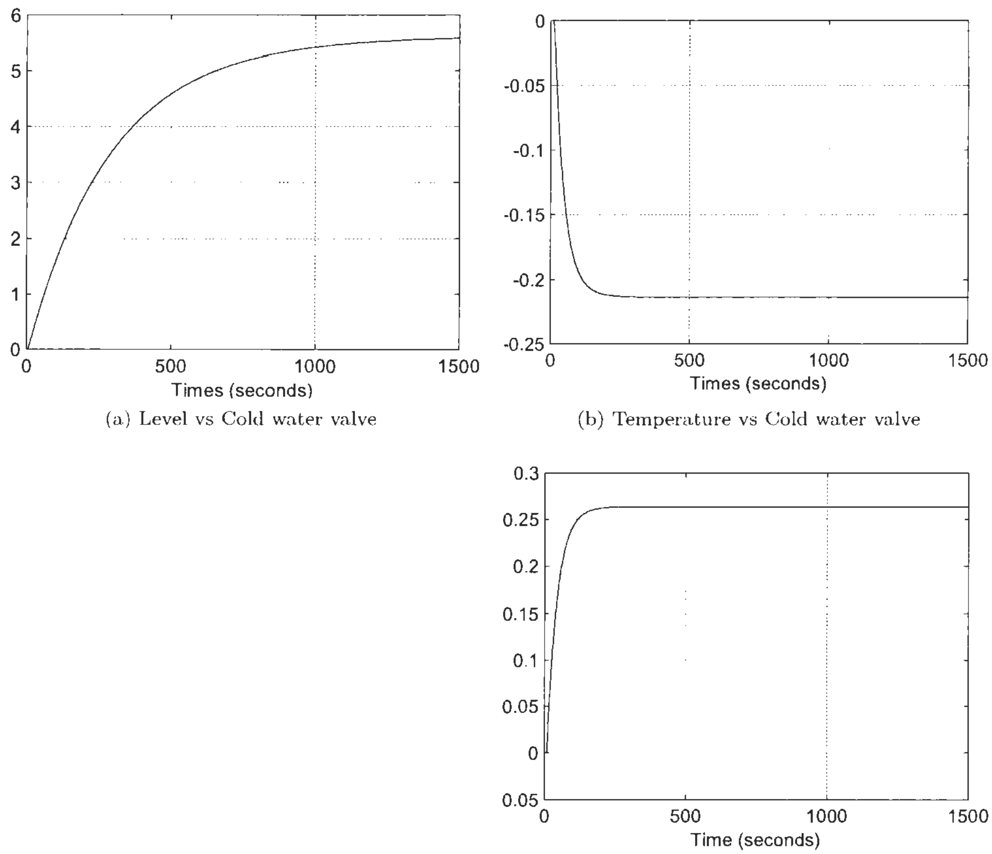
Plant model at the operating point are provided in Figure 8.1. These plant models are extensively used while designing the MPC block which uses DMC algorithm to design a controller. The details algorithm of DMC is discussed in the previous chapters. In the next section different control structures will be discussed.

8.2 Control Structures

In this work, the performance of three different control structures are compared. These are: a two-layer cascaded PID structure; a hybrid structure with PID in the base layer and the set-points of the PID manipulated by DMC; a PID-free structure where the control valve is directly manipulated by DMC.

8.2.1 Two Layer Cascaded PID Structure

The cascaded PID structure is presented in Figure 8.2 using the four measured variables and two manipulated variables. Cold water flow and steam flow are the two measured variables used as the feedback to the base layer PID. The outputs of the base layer PIDs are used to manipulate the position of the control valves of cold water and steam. Set-points of the base layer PID controllers are manipulated by supervisory layer PID. Measured variables, tank level and temperature, are used as feedback signals to the supervisory layer PID, which compares the measured values with their



(a) Level vs Cold water valve

(b) Temperature vs Cold water valve

(c) Temperature vs Steam valve

Figure 8.1: Step response models between the Process outputs and inputs

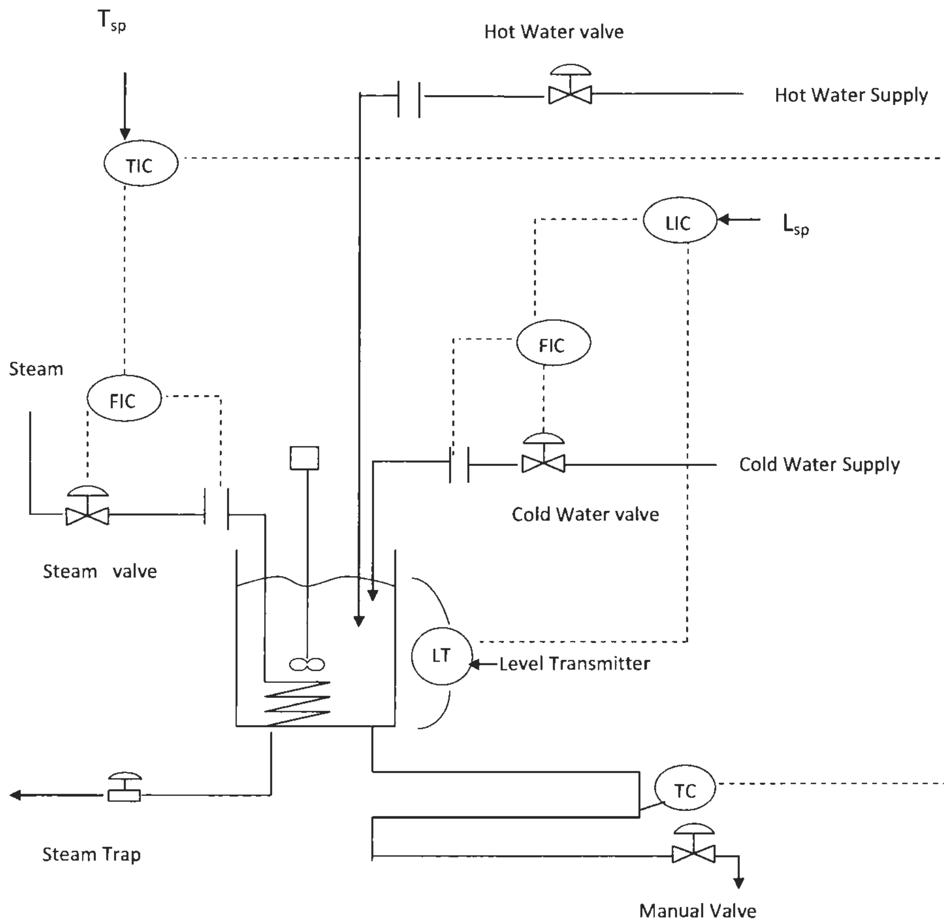


Figure 8.2: Two layer cascaded PID structure

corresponding desired values and provides control actions accordingly.

Details design methodology of the design of two layer PID is discussed in [Thornhill et al., 2008]. Moreover, an electronic model of the plant controlled with the two layer PID is available online.

8.2.2 Hybrid Structure with Base Layer PID Manipulated by DMC

The hybrid control structure is shown in Figure 8.3. In this structure, the supervisory layer is a DMC controller. This structure is practised widely and gained acceptance mainly because it allows the implementation of MPC with minimal changes to the existing control structure, and also because the PID layer can act as a fall back when MPC is turned off for any reason. In this structure, the plant, together with the PID controller, constitutes the system for the MPC that controls the tank level and temperature by manipulating the set-points of the base layer PID flow controllers.
MPC block is shown in the block diagram which is a centralized controller, that uses the plant model and DMC algorithm to design a controller. In the hybrid case, as plant together with PID is used as the process, model would not be same as shown in the Figure 8.1. An identified FIR filter is used as the process model to design the DMC. FIR model is provided in Equations 8.1 to 8.4. Design parameters of DMC are provided in Table 8.2.

$$G_{11}(z) = \frac{-0.0002347z^{-2}}{1 - 1.935z^{-1} + 0.9352z^{-2}} \tag{8.1}$$

$$G_{12}(z) = \frac{0.0003297z^{-2}}{1 - 1.935z^{-1} + 0.9352z^{-2}} \tag{8.2}$$

$$G_{21}(z) = \frac{-0.0643962z^{-10}}{1 - 0.9089z^{-1} + 0.002072z^{-2}} \tag{8.3}$$

$$G_{21}(z) = \frac{-0.014916z^{-10}}{1 - 0.9089z^{-1} + 0.002072z^{-2}} \tag{8.4}$$

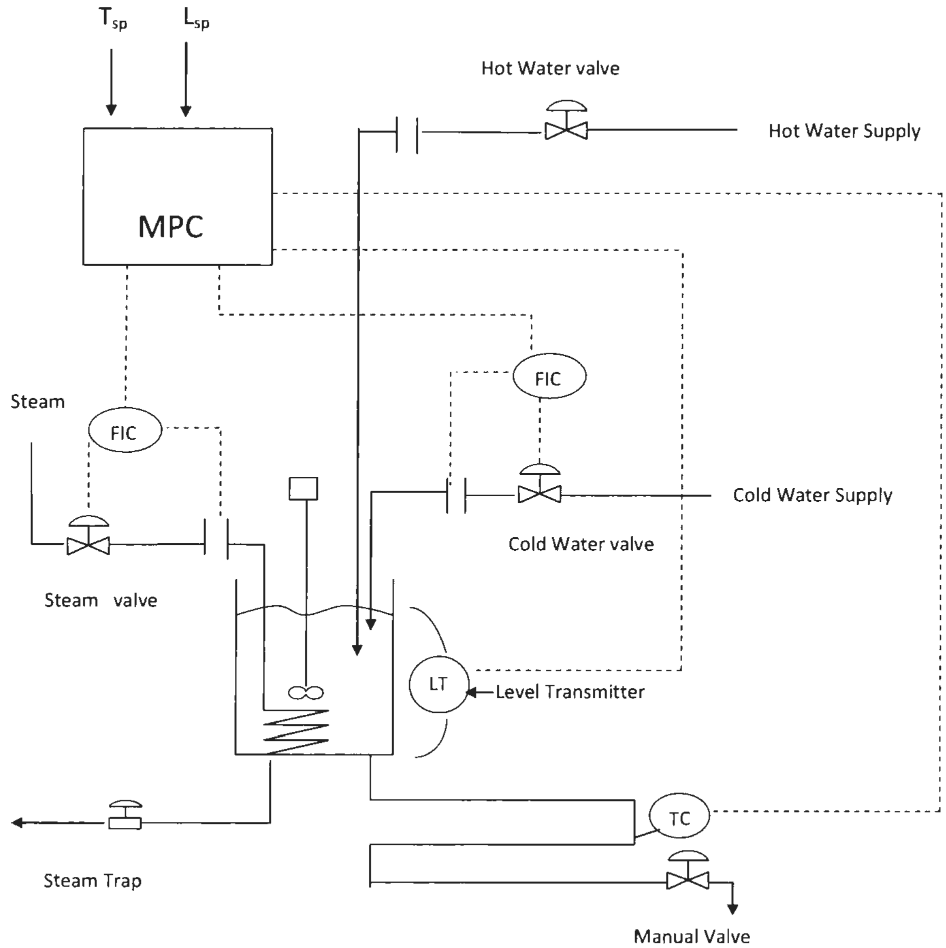


Figure 8.3: Hybrid control structure

Table 8.2: Design parameters of the DMC controller

Variable	Value
Prediction horizon(p)	15
Control horizon (m)	5
Weighting factor (Q)	1

Table 8.3: Design parameters of the DMC controller

Variable	Value
Prediction horizon(p)	15
Control horizon (m)	5
Weighting factor (Q)	1

8.2.3 PID Free MPC Structure

A PID-free control structure is presented in Figure 8.4. In this control structure there is no PID controller. A DMC controls the tank level and temperature by manipulating the cold water valve and the steam valve positions directly. So, in this case open loop model of the process plant provided in Figure 8.1, can be used directly to design MPC. Different design parameters to design DMC for this structure is provided in Table 8.3.

8.3 Performance Comparison of Different Types of Structures

The performances of the three different control structures are evaluated based upon set point tracking and regulatory control. Set point tracking performance describes how well a controller can react to the change of the desired set point of a process variable, whereas regulatory control assesses the ability of the controller to nullify the effect of any disturbance that appears in the system. Apart from these two properties, another desired property of a good controller is minimal fluctuations in the actuator. This will also be evaluated in this study.

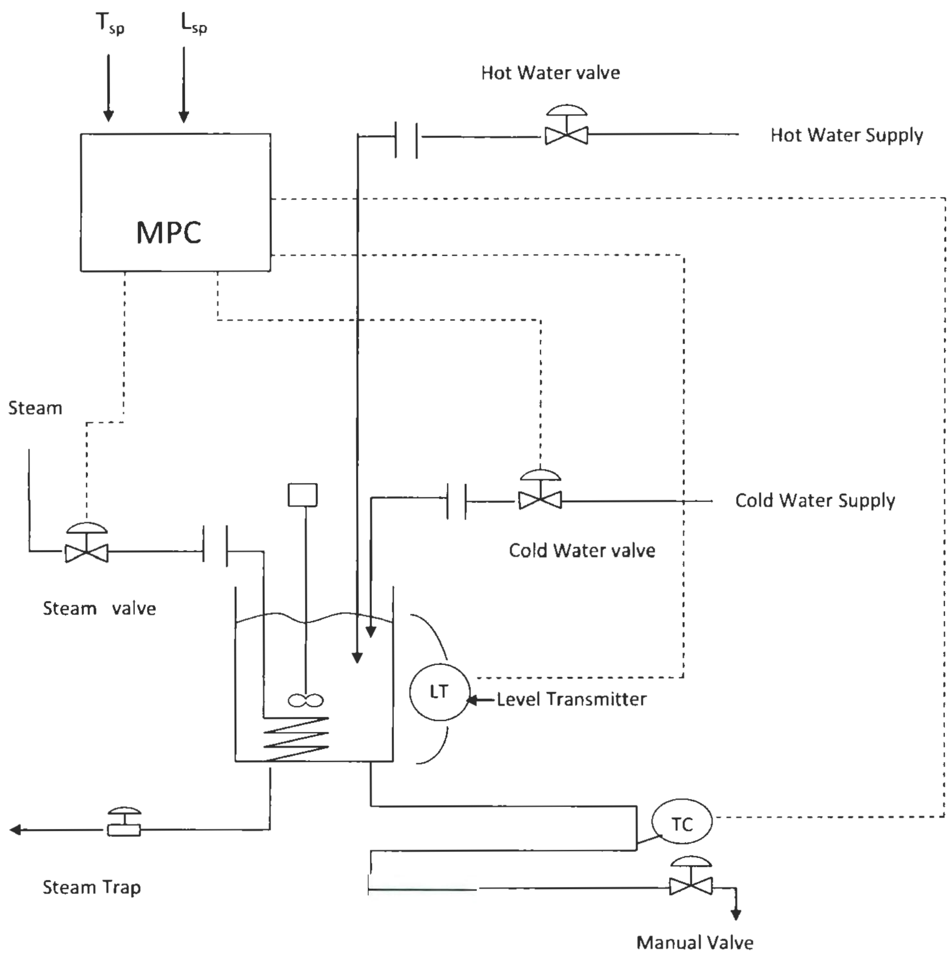
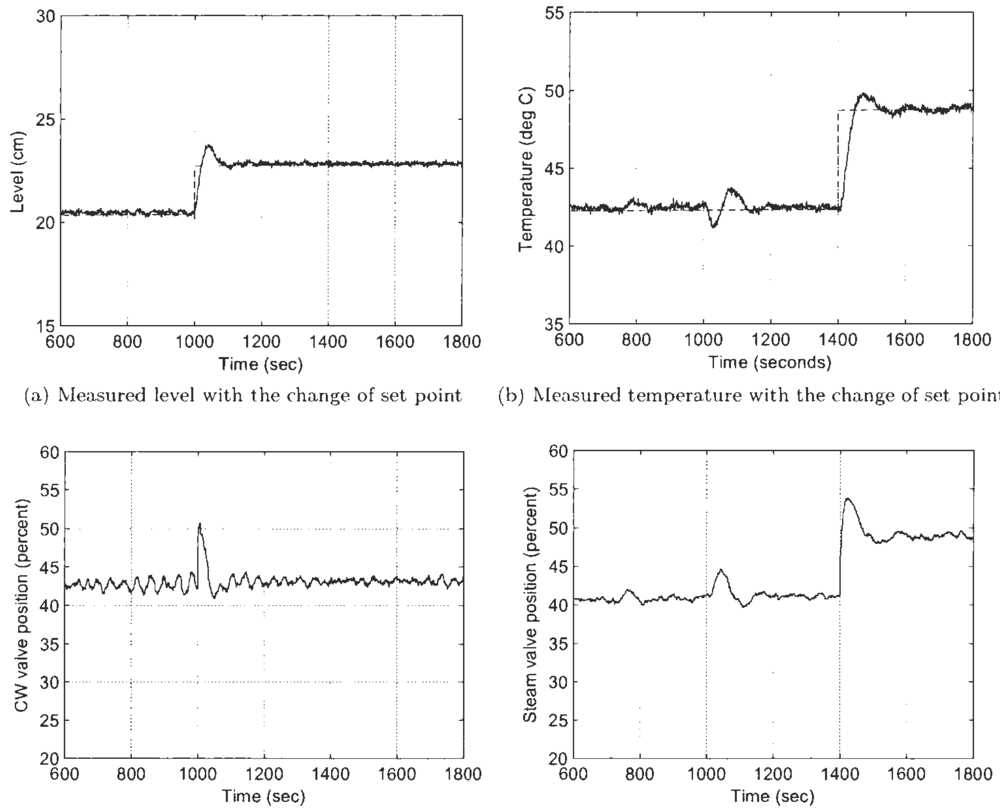


Figure 8.4: PID free control structure structure

8.3.1 Set-point Tracking

For assessing the controllers' response to a change in set point, the same scenario was set for the three different structures. The set points of both level and temperature are changed and the change of measured variables and actuator due to control action are observed. Measured outputs and manipulated variables for the three controllers are shown in Figures 8.5 to 8.7. The set point of the level is changed from 20.5 to 22.85 cm and the temperature set point is changed from 42.5 to 48.73°C. From the figures, it is evident that a PID-free MPC structure can react to a change of set point quicker than the other two structures; however, it demands more movements in the actuators. Considering valve movement, a hybrid structure proved to be better. However, it is much slower to react to the set point change. Both cascaded PID and hybrid structures have some overshoot which is much lower in the case of PID-free MPC. Execution frequency is another concern while designing DMC. In hybrid structure, DMC execution frequency is 15s, while for PID free structure execution frequency is 1s in order to reject any local disturbances. Hence, a PID-free structure has significantly more computational load compared to the hybrid PID.

In order to quantify the control performance of the three structures, they are compared using the integrated squared error (ISE) values for set-point tracking. ISE value is an integrated value of the deviation between the desired set-point and measured output over a certain period. In this case, an integral interval is considered to be the time that is required to achieve a steady state value after a set-point is changed. The ISE values for level and temperature are shown in Figures 8.8 and 8.9. From the figures, it can also be seen that the PID free structure shows superior performance compared to other structures. The hybrid control structure gives a larger ISE value due to steady state error. To sum up, having a large computational load PID-free structure



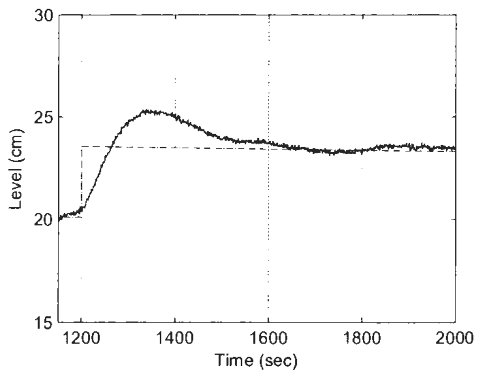
(a) Measured level with the change of set point

(b) Measured temperature with the change of set point

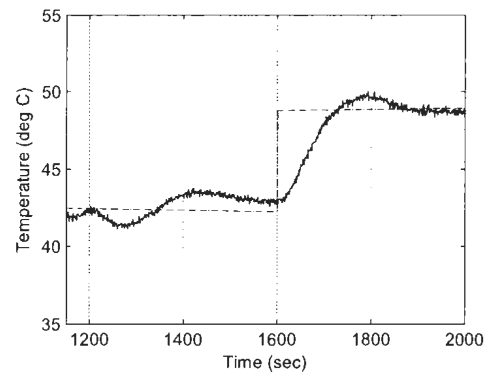
(c) Cold water valve position

(d) Steam valve position

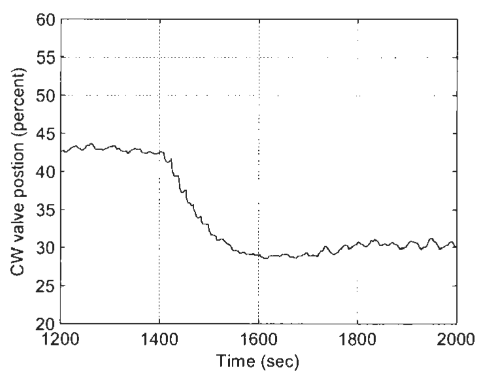
Figure 8.5: Measured output and actuator variable in cascaded PID structure for set point change



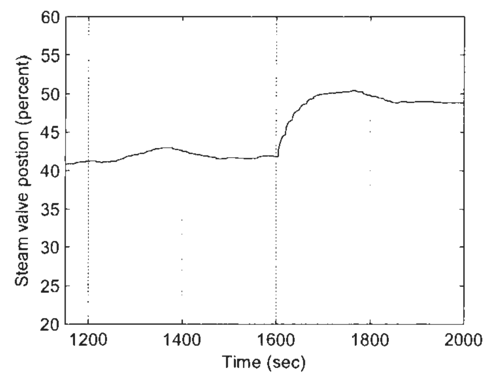
(a) Measured level with the change of set point



(b) Measured temperature with the change of set point

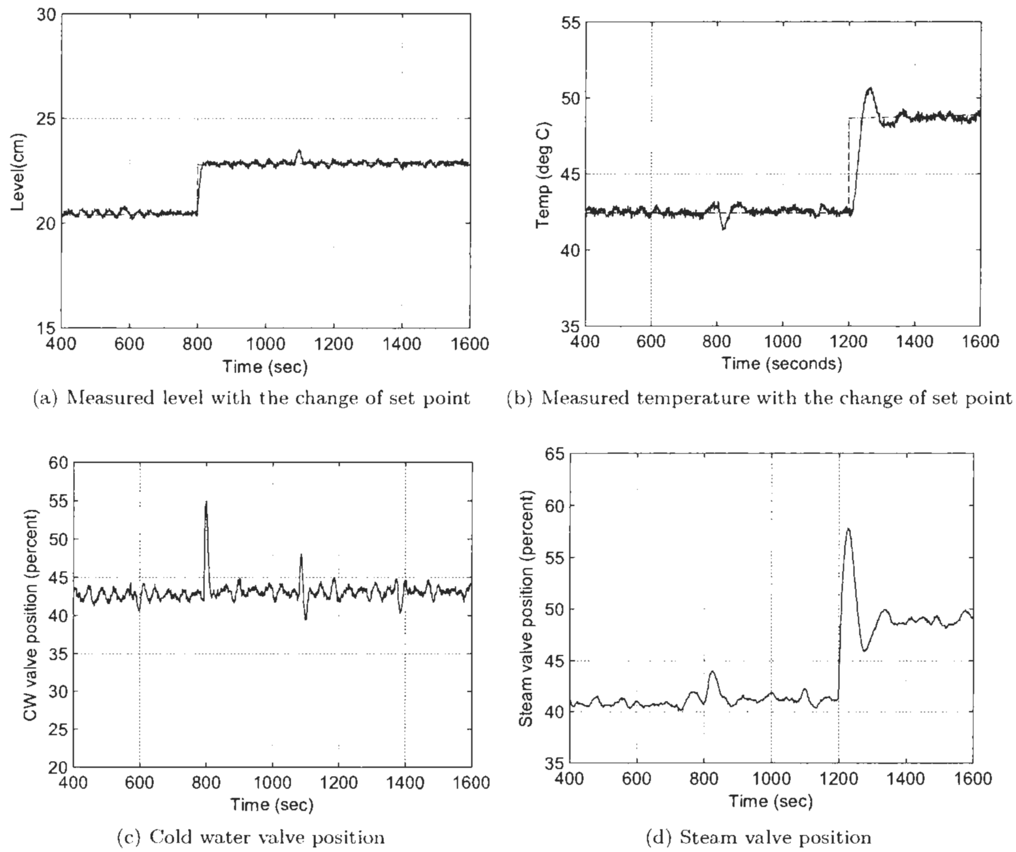


(c) Cold water valve position



(d) Steam valve position

Figure 8.6: Measured output and actuator variable in hybrid structure MPC for set point change



(a) Measured level with the change of set point

(b) Measured temperature with the change of set point

(c) Cold water valve position

(d) Steam valve position

Figure 8.7: Measured output and actuator variable in PID-free MPC structure for set point change

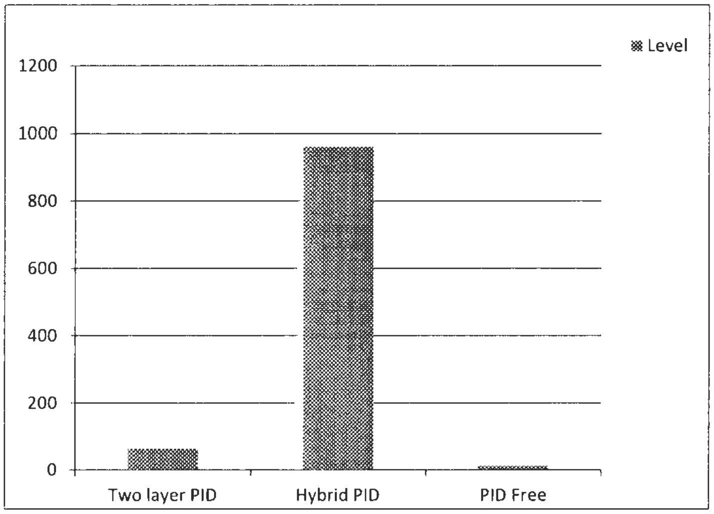


Figure 8.8: Comparison of the ISE value of different control structures for level control

Table 8.4: Settling time of level for different structures

Structure	Settling time(level)
Two layer cascaded PID	100 sec
Hybrid structure	420 sec
PID free structure	50 sec

is bit difficult to implement but it clearly outperforms the other structures in terms of control performances.

Comparison of the controller performance can also be made based on the settling time. Settling time of both level and temperature for each structure are shown in Table 8.4 and 8.5. Values of the settling time also suggest that, PID free structure clearly outperforms the other two structures in terms of stability having significantly lower settling time.

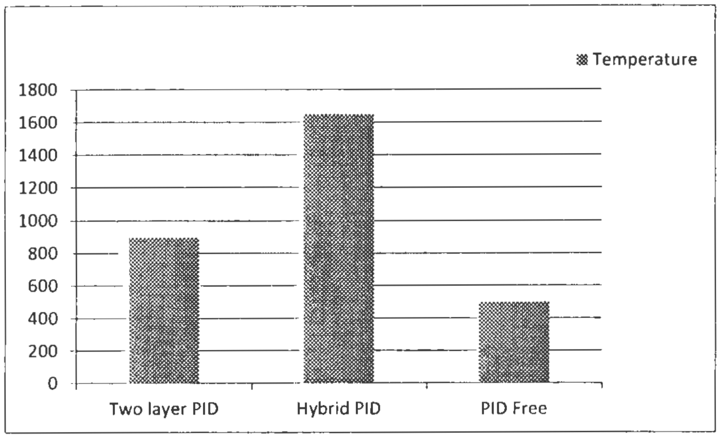


Figure 8.9: Comparison of the ISE value of different control structures for temperature control

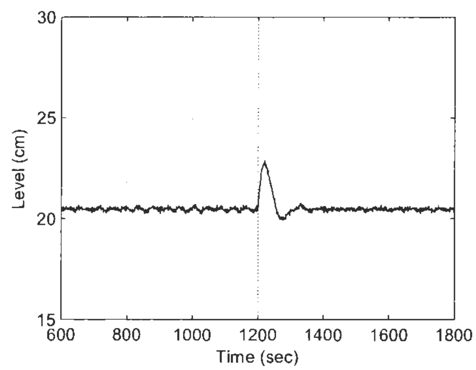
Table 8.5: Settling time of temperature for different structures

Structure	Settling time(temperature)
Two layer cascaded PID	200 sec
Hybrid structure	220 sec
PID free structure	170 sec

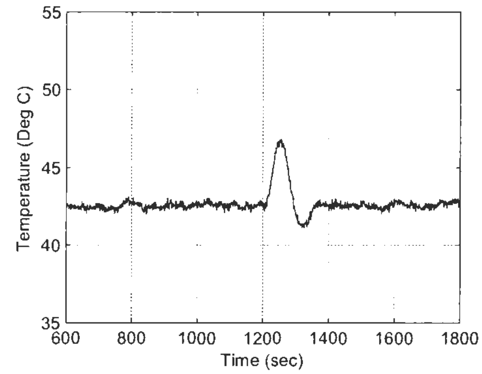
8.3.2 Regulatory Control

Regulatory control assesses a controller's ability to nullify a disturbance when it enters in to the system. In this study, hot water is considered to be the disturbance. Thus, a change in hot water valve position means that a disturbance has appeared in the system. For the nominal operation condition, the hot water valve is kept fully closed. In order to observe the regulatory control action of the controller, the hot water valve position is changed from 0 percent to 4.76 percent. Thus, hot water acts as a disturbance to the system and causes a rise of both measured variables, level and temperature, from their defined set point. The controllers took action to bring back the measured variable to the initial set point.

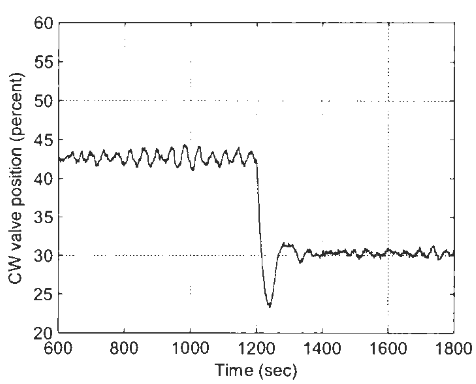
Figures 8.10 through 8.12 show the measured outputs and actuator movements after a disturbance is introduced into the system. From these results it is clear that all the controllers are capable of bringing the process to its initial state. A cascaded PID controller gives the fastest disturbance rejection with an undershoot and it has significant large swing in the actuator, which is not desirable. Both hybrid and PID-free structures reject disturbance without any undershoot. In the case of the actuator movement hybrid structure has less variation. However, the hybrid structure is significantly slower than the PID-free structure in disturbance rejection and allows a bigger rise of the measured output compared to the PID-free structure. The performance of hybrid structure may be improved by increasing the execution frequency of the supervisory DMC.



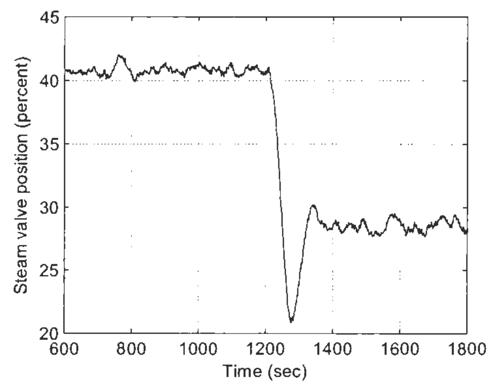
(a) Measured level with disturbance rejected



(b) Measured temperature with disturbance rejected

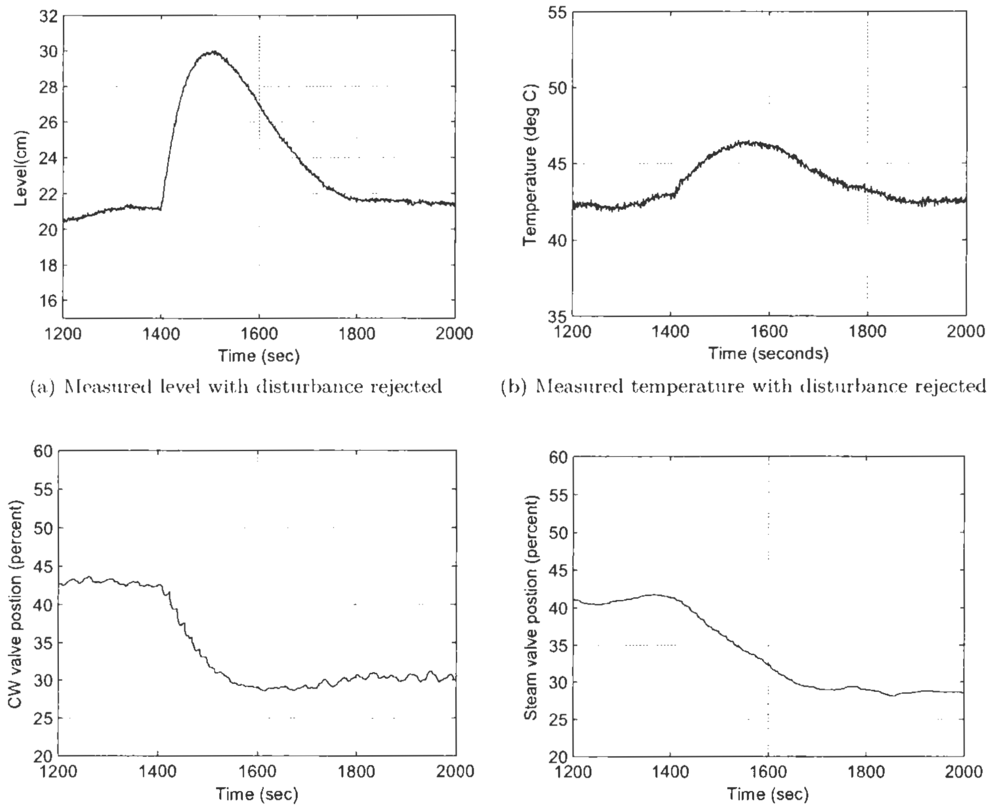


(c) Cold water valve position



(d) Steam valve position

Figure 8.10: Regulatory control of level and temperature using cascaded PID controller



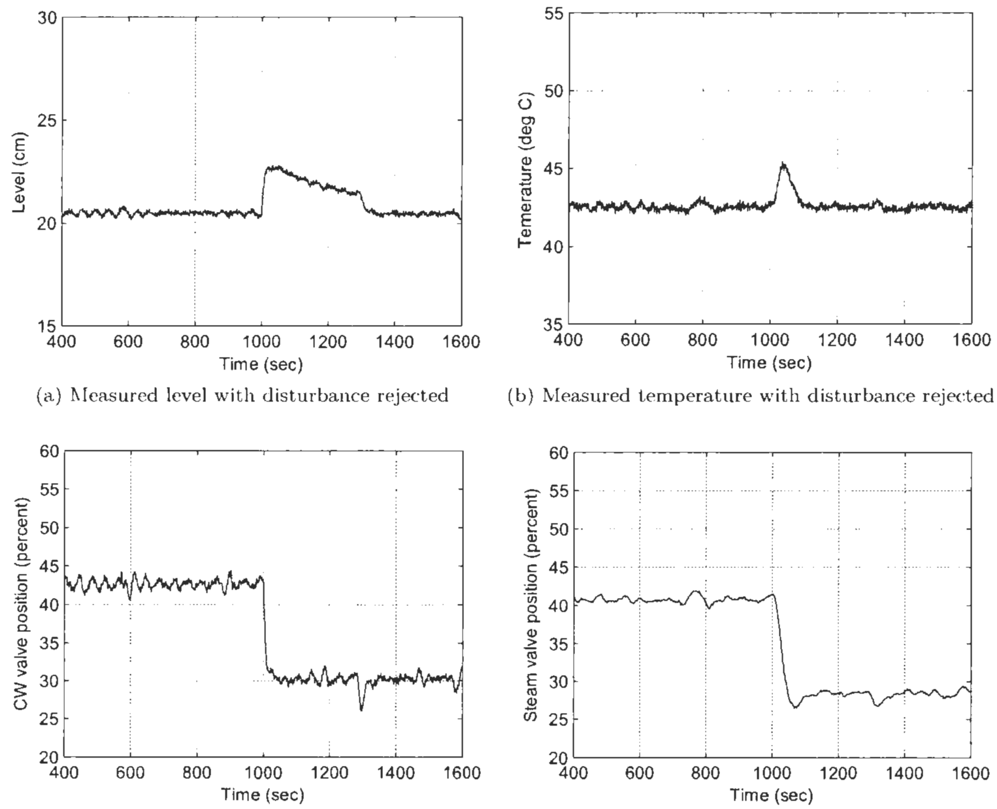
(a) Measured level with disturbance rejected

(b) Measured temperature with disturbance rejected

(c) Cold water valve position

(d) Steam valve position

Figure 8.11: Regulatory control of level and temperature using hybrid DMC-PID controller



(a) Measured level with disturbance rejected

(b) Measured temperature with disturbance rejected

(c) Cold water valve position

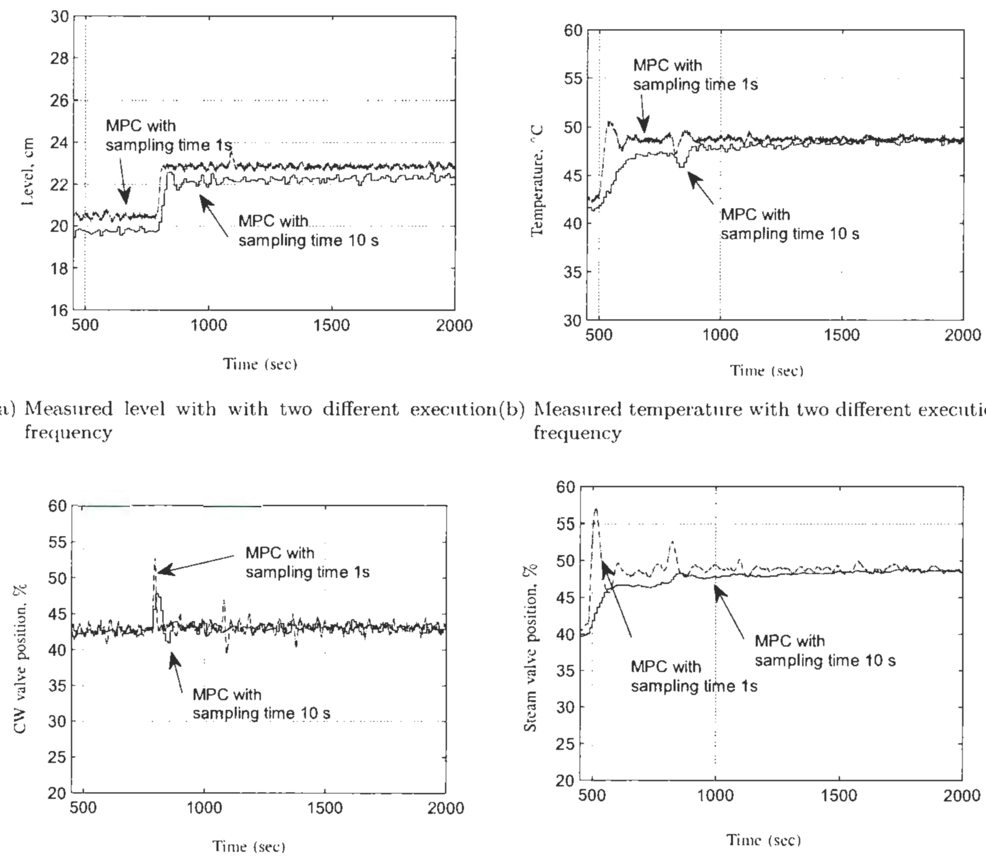
(d) Steam valve position

Figure 8.12: Regulatory control of level and temperature using PID-free DMC controller

8.4 Effect of Execution Frequencies in PID-free MPC Performance

Performance comparison in the previous section convincingly demonstrates that PID-free MPC structure has a superior performance over PID or hybrid structure. However, the main concern for the PID-free MPC is that it has significantly more computational load, as DMC has to provide a control action at every second. Decreasing the execution frequency would help to decrease the computational load. Moreover, PID-free MPC has more fluctuation in the valve position, which can be reduced by increasing sampling time, hence decreasing execution frequency. In this section, PID-free MPC is implemented at two different frequencies and their performances are evaluated. The first one is the controller described in the previous section with a sampling time of 1s, while for the other, a sampling time of 10s is chosen. Set point tracking performances of the PID-free MPC at these two execution frequencies are observed. The set point of level is changed from 20.5 to 22.85 cm at t= 800s and the set point of the temperature is changed from 42.5°C to 48.73°C at t= 500s.

Figure 8.13 shows the measured outputs and actuator movements at different execution frequencies. Comparing the results, we see that, for the lower execution frequency, a steady state error exists between the set-point and the response for a prolonged period. However, the valve movement is significantly reduced for the lower execution frequency. All these phenomena should be taken into account while choosing an execution frequency of a DMC controller.



(a) Measured level with with two different execution frequency

(b) Measured temperature with two different execution frequency

(c) Cold water valve position

(d) Steam valve position

Figure 8.13: Set-point tracking performance comparison of PID-free MPC structure for different execution frequencies

Chapter 9

Conclusions of the Predictive Control

A comprehensive simulation-based study was carried out to compare the performances of the two control structures: 'MPC cascaded to PID' and a 'PID-free MPC', where MPC is directly manipulating the actuators. DMC is used as the representative MPC algorithm. The simulation study was carried out on a CSTH system. The performance of the controllers for set point tracking and disturbance rejection were monitored. ISE is used as the control performance indicator which clearly suggests PID-free MPC structure is the superior one. However, PID-free MPC needs to be executed at a high frequency which increases the computational load.

The findings and contributions are summarized below:

- A hybrid control structure where a DMC cascaded to a PID controller and a PID-free control structure where DMC directly manipulates the actuator are designed for a CSTH system.These two design of control were implemented in Matlab Simulink.

- The performance of the above control structures are evaluated for both set-point change and disturbance rejection. More aggressive control can be achieved by the PID-free DMC structure. Some overshoot in the response is observed for the PID-free DMC structure.

- Quantitatively, the performance of these competing control structures are compared by calculating ISE. This also shows that the PID-free DMC structure outperforms the hybrid DMC-PID structure. In fact, the performance of the hybrid structure is inferior to the cascaded PID structure to some extent.

- The computational load and the movements of an actuator for the PID-free DMC structure is more than that of a hybrid structure. Computational load can be managed by optimizing the execution frequency.

9.1 Future Recommendations

- Experimental validation: The effectiveness of the methodologies are discussed in this study based on the simulation results. This can be further validated using an experimental study.

- The effect of valve non-linearity on the performance of the PID-free MPC structure was not studied. This is an important question which can be studied using an experimental set up.

Simulation results showed that direct use of MPC rather than an MPC in supervisory layer and base layer PID may allow achieving better control performances. This advantage comes with additional computational load. Experimental studies should be carried out to validate the above simulation findings.

Bibliography

[Adnan et al., 2011] Adnan, N. A., Izadi, I., and Chen, T. (2011). On expected detection delays for alarm systems with deadbands and delay-timers. *Journal of Process Control*, 21(9):1318 – 1331.

[Astrom and Hagglund, 2001] Astrom, K. and Hagglund, T. (2001). The future of pid control. *Control Engineering Practice*, 9(11):1163 – 1175.

[Bakshi and Stephanopoulos., 1993] Bakshi, B. R. and Stephanopoulos., G. (1993). Wave-net: a multiresolution, hierarchical neural network with localized learning. *American Institute of Chemical Engineers Journal*, 39 (1):57–81.

[Bao et al., 2011] Bao, H., Khan, F., Iqbal, T., and Chang, Y. (2011). Risk-based fault diagnosis and safety management for process systems. *Process Safety Progress*, 30(1):6–17.

[Basseville and Benveniste, 1986] Basseville, M. and Benveniste, A. (1986). *Detection of abrupt changes in signals and dynamic systems.* Berlin: Springer-Verlag.

[Basseville and Nikiforov, 1993] Basseville, M. and Nikiforov, I. V. (1993). *Detection of abrupt changes-theory and application (Information and System Sciences Series). Prentice Hall.* Prentice Hall.

[Bemporad and Morari, 1999] Bemporad, A. and Morari, M. (1999). Robust model predictive control: A survey. In Garulli, A. and Tesi, A., editors, *Robustness in identification and control*, volume 245 of *Lecture Notes in Control and Information Sciences*, pages 207–226. Springer London.

[Ben-Haim, 1980] Ben-Haim, Y. (1980). An algorithm for failure location in a complex network. *Nuclear Science and Engineering*, 75:191–199.

[Ben-Haim, 1983] Ben-Haim, Y. (1983). Malfunction location in linear stochastic systems-application to nuclear power plants. *Nuclear Science and Engineering*, 85:156–166.

[Bhagwat et al., 2003a] Bhagwat, A., Srinivasan, R., and Krishnaswamy, P. (2003a). Fault detection during process transitions: a model-based approach. *Chemical Engineering Science*, 58(2):309 – 325.

[Bhagwat et al., 2003b] Bhagwat, A., Srinivasan, R., and Krishnaswamy, P. R. (2003b). Multi-linear model-based fault detection during process transitions. *Chemical Engineering Science*, 58(9):1649 – 1670.

[Bonivento et al., 2001] Bonivento, C., Castaldi, P., and Mirotta, D. (2001). Predictive control vs pid control of an industrial heat exchanger. In *Proceedings of the 9th Mediterranean Conference on Control and Automation*.

[Chang and Hwang, 1998] Chang, C. T. and Hwang, J. I. (1998). Simplification techniques for ekf computations in fault diagnosis - suboptimal gains. *Chemical Engineering Science.*, 53 (22):3853–3862.

[Chang et al., 2011] Chang, Y., Khan, F., and Ahmed, S. (2011). A risk-based approach to design warning system for processing facilities. *Process Safety and Environmental Protection*, 89(5):310 – 316.

[Cherry and Qin, 2006] Cherry, G. and Qin, S. (2006). Multiblock principal component analysis based on a combined index for semiconductor fault detection and diagnosis. *Semiconductor Manufacturing, IEEE Transactions on*, 19(2):159 – 172.

[C.Hoskins et al., 1991] C.Hoskins, J., Kaliyur, K. M., and Himmelblau, D. (1991). Fault diagnosis in complex chemical plants using artificial neural networks. *American Institute of Chemical Engineers Journal*, 37 (1):137–141.

[Chow and Willsky, 1984] Chow, E. Y. and Willsky, A. S. (1984). Analytical redundancy and the design of robust failure detection systems. ieee transactions on. *Automatic Control*, 29 (7):603–614.

[Chu et al., 1994] Chu, R., Bullemer, P., Harp, S., Ramanathan, P., and Spoor, D. (1994). Qualitative user aiding for alarm management (qualm): an integrated demonstration of emerging technologies for aiding process control operators. In *Systems, Man, and Cybernetics, 1994. 'Humans, Information and Technology', 1994 IEEE International Conference on*, volume 1, pages 735 –740 vol.1.

[Cutler and Ramaker, 1979] Cutler, C. R. and Ramaker, B. L. (1979). Dynamic matrix control - a computer control algorithm. In *AICHE national meeting, Houston, TX, April.*

[Cutler and Ramaker, 1980] Cutler, C. R. and Ramaker, B. L. (1980). Dynamic matrix control - a computer control algorithm. In *Proceedings of the joint automatic control conference.*

[Desborough and Miller, 2001] Desborough, L. and Miller, R. (2001). Increasing customer value of industrial control performance monitoring - honeywell's experience.

[Dingli et al., 1995] Dingli, Y., Gomm, J. B., and. D. Williams, D. N. S., and Disdell, K. (1995). Fault diagnosis for a gas-fired furnace using bilinear observer method. In *Proceedings of the American control conference , Seattle, Washington.*

[Dong and McAvoy, 1996] Dong, D. and McAvoy, T. J. (1996). Batch tracking via nonlinear principal component analysis. *American Institute of Chemical Engineers Journal*, 42 (8):2199–2208.

[Fan et al., 1993] Fan, J. Y., Nikolaou, M., and White, R. E. (1993). An approach to fault diagnosis of chemical processes via neural networks. *American Institute of Chemical Engineers Journal*, 39 (1):83–88.

[Farell and Roat, 1994] Farell, A. E. and Roat, S. D. (1994). Framework for enhancing fault diagnosis capabilities of artificial neural networks. *Computers and Chemical Engineering*, 18 (7):613–635.

[Fathi et al., 1993] Fathi, Z., Ramirez, W. F., and Korbiez, J. (1993). Analytical and knowledge-based redundancy for fault diagnosis in process plants. *AIChE J.*, 39:42–56.

[Fernandez et al., 2005] Fernandez, L., Pogrebnyak, O., and Marquez, C. (2005). Neural network and trend prediction for technological processes monitoring. In Gelbukh, A., de Albornoz Ãlvaro, and Terashima-MarAn, H., editors, *MICAI 2005: Advances in Artificial Intelligence*, volume 3789 of *Lecture Notes in Computer Science*, pages 731–740. Springer Berlin / Heidelberg.

[Frank, 1990] Frank, P. M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy*/a survey and some new results. *Automatica*, 26:459–474.

[Garcia and Morshedi, 1986] Garcia, C. E. and Morshedi, A. (1986). Quadratic programming solution of dynamic matrix control(qdmc). *Chemical Engineering Communications*, 46:73–87.

[Garcia et al., 1989] Garcia, C. E., Prett, D. M., and Morari, M. (1989). Model predictive control: Theory and practice - a survey. *Automatica*, 25 (3):335–348.

[Gertler et al., 1995] Gertler, J., Costin, M., Fang, X., Kowalczuk, Z., Kunwer, M., and Monajemy, R. (1995). Model based diagnosis for automative engines - algorithm development and testing on a production vehicle. *IEEE Transactions on Control Systems Technology*, 3:61–69.

[Gertler et al., 1990] Gertler, J., Fang, X., and Luo, Q. (1990). Detection and diagnosis of plant failures: the orthogonal parity equation approach. *Control and Dynamic Systems*, 37:159–216.

[Han, 2009] Han, J. (2009). From pid to active disturbance rejection control. *Industrial Electronics, IEEE Transactions on*, 56(3):900 –906.

[Holcomb and Morari, 1991] Holcomb, T. and Morari, M. (1991). Local training of radial basis function networks: towards solving the hidden unit problem. In *American control conference*.

[Isermann, 2005] Isermann, R. (2005). Model-based fault-detection and diagnosis - status and applications. *Annual Reviews in Control*, 29(1):71 – 85.

[Izadi et al., 2009a] Izadi, I., Shah, S. L., Shook, D. S., and Chen, T. (2009a). An introduction to alarm analysis and design fault detection, supervision and safety of technical processes. In *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*.

[Izadi et al., 2009b] Izadi, I., Shah, S. L., Shook, D. S., Kondaveeti, S. R., and Chen, T. (2009b). A framework for optimal design of alarm systems. In *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes.*

[J. Gertler and Monajemy., 1995] J. Gertler, J. and Monajemy., R. (1995). Generating directional residuals with dynamic parity relations. *Automatica*, 31:627-635.

[Juricek et al., 2001] Juricek, B. C., Seborg, D. E., and Larimore, W. E. (2001). Predictive monitoring for abnormal situation management. *Journal of Process Control*, 11(2):111 - 128.

[Kavuri and Venkatasubramanian, 1994] Kavuri, S. N. and Venkatasubramanian, V. (1994). Neural network decomposition strategies for large scale fault diagnosis. *International Journal of Control*, 59 (3):767–792.

[Kresta et al., 1991] Kresta, J. V., Macgregor, J. F., and Marlin, T. E. (1991). Multivariate statistical monitoring of process operating performance. *The Canadian Journal of Chemical Engineering*, 69(1):35–47.

[Krishna Vinaya et al., 2012] Krishna Vinaya, V., Ramkumar, K., and Alagesan, V. (2012). Control of heat exchangers using model predictive controller. In *Advances in Engineering, Science and Management (ICAESM), 2012 International Conference on*, pages 242 –246.

[Lee et al., 2006] Lee, J.-M., Qin, S. J., and Lee, I.-B. (2006). Fault detection and diagnosis based on modified independent component analysis. *AIChE Journal*, 52(10):3501-3514.

[Lee et al., 2004] Lee, J.-M., Yoo, C., and Lee, I.-B. (2004). Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chemical Engineering Science*, 59(14):2995 – 3006.

[Leonard and Kramer, 1990] Leonard, J. A. and Kramer, M. A. (1990). Limitations of backpropagation approach to fault diagnosis and improvements with radial basis functions. In *AIChE annual meeting , Chicago*.

[Li et al., 2000] Li, W., Yue, H., Valle-Cervantes, S., and Qin, S. (2000). Recursive pca for adaptive process monitoring. *Journal of Process Control*, 10(5):471 – 486.

[Lukacova and Borzikova, 2010] Lukacova, I. and Borzikova, J. (2010). Comparison of advanced control methods with classical pid control for using in heating process control based on outdoor temperature compensation. *Journal of applied science in the thermodynamics and fluid mechanics*, 4.

[MacGregor et al., 1994] MacGregor, J. F., Jaeckle, C., Kiparissides, C., and Koutondi, M. (1994). Process monitoring and diagnosis by multiblock pls methods. *American Institute of Chemical Engineers Journal*, 40 (5):826–838.

[MacGregor and Kourti, 1995] MacGregor, J. F. and Kourti, T. (1995). Statistical process control of multivariate processes. *Control Engineering Practice*, 3:403–414.

[Massoumnia, 1986] Massoumnia, M. A. (1986). A geometric approach to the synthesis of failure detection filters. *IEEE Transactions on Automatic Control*, 31:839–846.

[Misra et al., 2002] Misra, M., Yue, H., Qin, S., and Ling, C. (2002). Multivariate process monitoring and fault diagnosis by multi-scale pca. *Computers & Chemical Engineering*, 26(9):1281 – 1293.

[Morshedi et al., 1985] Morshedi, A. M., Cutler, C. R., and Skrovaneck, T. A. (1985). Optimal solution of dynamic matrix control with linear programing techniques (ldmc),. In *American Control Conference (199-2080*.

[Muske and Rawlings, 1993] Muske, K. R. and Rawlings, J. B. (1993). Model predictive control with linear models. *AIChE Journal*, 39(2):262–287.

[Na, 2001] Na, M. G. (2001). Auto-tuned pid controller using a model predictive control method for the steam generator water level. *Nuclear Science, IEEE Transactions on*, 48(5):1664 –1671.

[N.Clark, 1979] N.Clark, R. (1979). The dedicated observer approach to instrument fault detection. In *Proceedings of the 15th IEEE-CDC*.

[Nomikos and MacGregor, 1994] Nomikos, P. and MacGregor, J. F. (1994). Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40(8):1361–1375.

[Ogunnaike and Mukati, 2006] Ogunnaike, B. A. and Mukati, K. (2006). An alternative structure for next generation regulatory controllers: Part i: Basic theory for design, development and implementation. *Journal of Process Control*, 16(5):499 – 509.

[Ogunnaike and Ray, 1994] Ogunnaike, B. A. and Ray, W. (1994). *Process Dynamics, Modeling, and Control*. Oxford.

[Overschee and Moor, 2001] Overschee, P. V. and Moor, B. D. (2001). The end of heuristic pid tuning, in preprints of the ifac workshop on digital control: past, present and future of pid control cbs.

[Pannocchia et al., 2005] Pannocchia, G., Laachi, N., and Rawlings, J. B. (2005). A candidate to replace pid control: Siso-constrained lq control. *AIChE Journal*, 51(4):1178–1189.

[Prett and Garcia, 1988] Prett, D. M. and Garcia, C. E. (1988). *Fundamental Process Control*. Butterworths-Heinemann, Boston, MA.

[Prett and Gillette, 1980] Prett, D. M. and Gillette, R. D. (1980). Optimization and constrained multivariable control of a catalytic cracking unit. in. In *Proceedings of the joint automatic control conference*.

[Qin and Badgwell, 2003] Qin, S. and Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733 – 764.

[Qin and McAvoy, 1992] Qin, S. J. and McAvoy, T. J. (1992). Nonlinear pls modeling using neural networks. *Computers and Chemical Engineering*, 16 (4):379–391.

[Raich and Cinar, 1996] Raich, A. and Cinar, A. (1996). Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *American Institute of Chemical Engineers*, 42 (4):995–1009.

[Richalet et al., 1978] Richalet, J., Rault, A., L.Testud, J., and Papon, J. (1978). Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14:413–428.

[Richalet et al., 1976] Richalet, J., Rault, A., Testud, J. L., and Papon, J. (1976). Algorithmic control of industrial processes. In *Proceedings of the 4th IFAC symposium on identification and system parameter estimation*.

[Rothenberg, 2009] Rothenberg, D. H. (2009). *Alarm Management for Process Control*. Momentum Press.

[Ruiz et al., 2002] Ruiz, D., Benqlilou, C., Nougués, J. M., Puigjaner, L., and Ruiz, C. (2002). Proposal to speed up the implementation of an abnormal situation management in the chemical process industry. *Industrial & Engineering Chemistry Research*, 41(4):817–824.

[Seborg et al., 1989] Seborg, D. E., Edgar, T., and Mellichamp, D. (1989). *Process Dynamics and Control*. Wiley.

[Shahriari et al., 2006] Shahriari, M., Shee, A., and Örtengren, R. (2006). The development of critical criteria to improve the alarm system in the process industry. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 16(3):321–337.

[Shewhart, 1931] Shewhart, W. A. (1931). *Economic Control of Quality Manufactured Product*. Oxford, England.

[Thornhill et al., 2008] Thornhill, N. F., Patwardhan, S. C., and Shah, S. L. (2008). A continuous stirred tank heater simulation model with applications. *Journal of Process Control*, 18(3-4):347 – 360.

[Tsai and T.Chang, 1995] Tsai, C. S. and T.Chang, C. (1995). Dynamic process diagnosis via integrated neural networks. *Computers and Chemical Engineering*, 19:747–752.

[Ungar et al., 1990] Ungar, L. H., Powell, B. A., and Kamens, S. N. (1990). Adaptive networks for fault diagnosis and process control. *Computers and Chemical Engineering*, 14 (4-5):561–572.

[Vaidyanathan and Venkatasubramanian, 1992] Vaidyanathan, R. and Venkatasubramanian, V. (1992). Representing and diagnosing dynamic process data using neural networks. *Engineering Applications of Artificial Intelligence*, 5 (1):11–21.

[Varga et al., 2010] Varga, T., Szeifert, F., and Abonyi, J. (2010). Detection of safe operating regions: A novel dynamic process simulator based predictive alarm management approach. *Industrial & Engineering Chemistry Research*, 49(2):658–668.

[Venkatasubramanian, 1985] Venkatasubramanian, V. (1985). Inexact reasoning in expert systems: a stochastic parallel network approach. In *Proceedings of the second conference on artificial intelligence applications*.

[Venkatasubramanian and Chan, 1989] Venkatasubramanian, V. and Chan, K. (1989). A neural network methodology for process fault diagnosis. *American Institute of Chemical Engineers Journal*, 35 (12):1993–2002.

[Venkatasubramanian et al., 2003a] Venkatasubramanian, V., Rengaswamy, R., and Kavuri, S. N. (2003a). A review of process fault detection and diagnosis: Part ii: Qualitative models and search strategies. *Computers & Chemical Engineering*, 27(3):313 – 326.

[Venkatasubramanian et al., 2003b] Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., and Yin, K. (2003b). A review of process fault detection and diagnosis: Part iii: Process history based methods. *Computers & Chemical Engineering*, 27(3):327 – 346.

[Venkatasubramanian et al., 1990] Venkatasubramanian, V., Vaidyanathan, R., and Yamamoto, Y. (1990). Process fault detection and diagnosis using neural networks i: steady state processes. *Computers and Chemical Engineering*, 14 (7):699–712.

[Watanabe et al., 1994a] Watanabe, K., Hirota, S., Iloa, L., and Himmelblau, D. M. (1994a). Diagnosis of multiple simultaneous fault via hierarchical artificial neural networks. *American Institute of Chemical Engineers Journal*, 40 (5):839–848.

[Watanabe et al., 1994b] Watanabe, K., Hirota, S., Iloa, L., and Himmelblau, D. M. (1994b). Diagnosis of multiple simultaneous fault via hierarchical artificial neural networks. american institute of chemical engineers journal. 40 (5):839–848.

[Willsky, 1976] Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12:601–611.

[Willsky and Jones, 1976] Willsky, A. S. and Jones, H. L. (1976). A generlized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control*, 21:108–112.

[Yang et al., 2010] Yang, F., Shah, S., and Xiao, D. (2010). Correlation analysis of alarm data and alarm limit design for industrial processes. In *American Control Conference (ACC), 2010*, pages 5850 –5855.

[Yin et al., 2012] Yin, S., Ding, S. X., Haghani, A., Hao, H., and 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*, 22(9):1567 – 1581.

[Yoon and MacGregor, 2000] Yoon, S. and MacGregor, J. F. (2000). Statistical and causal model-based approaches to fault detection and isolation. *AIChE Journal*, 46(9):1813–1824.

[Yuki, 2002] Yuki, Y. (2002). Alarm system optimization for increasing operations productivity. *ISA Transactions*, 41(3):383 – 387.

[Zadakbar et al., 2012] Zadakbar, O., Imtiaz, S., and Khan, F. (2012). Dynamic risk assessment and fault detection using principal component analysis. *Industrial & Engineering Chemistry Research*, 0(0):null.

[Zadakbar et al., 2011] Zadakbar, O., Imtiaz, S. A., and Khan, F. (2011). Dynamic risk assessment and fault detection using multivariate technique. In *AIChE J submitted.*

[Zamanizadeh et al., 2008] Zamanizadeh, E., Salahshoor, K., and Manjili, Y. (2008). Prediction of abnormal situation in nonlinear systems using ckf. In *Networking, Sensing and Control, 2008. ICNSC 2008. IEEE International Conference on*, pages 681 –686.