DYNAMIC RISK ASSESSMENT AND FAULT DETECTION

IN PROCESS FACILITIES

by

© Omid Zadakbar

A Thesis submitted to the

School of Graduate Studies

In partial fulfillment of the requirements for the degree of

Doctoral of Philosophy

Faculty of Engineering & Applied Science

Memorial University of Newfoundland

June 2014

St. John's Newfoundland

Dedicated to:

My mother, Nahid.

Her support, encouragement, and endless love have sustained me throughout my life.

ABSTRACT

A new multivariate risk-based fault detection and diagnosis technique targeting the safety issues of a process system is being proposed. In contrast to typical fault detection methods which only aim to detect operational faults that affect the control objectives of the process, this method targets the safety of the process. Typical fault detection and diagnosis methods are inadequate as none of the methods considers the consequences of the fault on process safety, integrity and the environment. However the proposed method provides a dynamic process risk indication based on the probability of occurrence of a fault and its consequences. In this method, the consequence is expressed in economic value that demonstrates the potential economic impact of the fault on the process, equipment, workers and the environment. Through this approach, warning system and risk management strategies may be activated when the risk of operation exceeds the acceptable threshold. This is an important concept because it can direct the attention and effort of operators to the faults which poses the most operational or safety risk. Both model based and history based fault detection and diagnosis techniques have been extended to a risk-based fault detection and diagnosis framework. Application of this new risk-based approach provides early warnings and early activation of safety systems prior to the fault impacting the system. This multivariate technique provides much early warning compared to the univariate methods. It has more power in discerning between operational changes and abnormal conditions which have potential to cause accidents. The main benefits of this approach are improved safety, minimum interruption of operation, better alarm management or early warning system and higher availability of process. The novelties and contributions of this work are development of multivariate dynamic risk assessment methodology using history based and model

based methods for linear and nonlinear models combined with a newly developed economic consequence analysis methodology. This methodology makes the severity of the faults more sensible by quantifying consequences in economic terms. This new economic consequence methodology helps to integrates real time process state to accident scenarios via loss functions.

The proposed framework when implemented on a process could serve as a real-time process risk monitor. This would help to take preventive actions in order to minimize process risks.

ACKNOWLEDGEMENT

I would like to express my sincere and profound sense of gratitude and respect to my supervisors Dr. Faisal Khan and Dr. Syed Imtiaz for their expert guidance and untiring support throughout my research at the Memorial University of Newfoundland. Without their encouragement and advice, this thesis might not have been completed successfully. I am especially thankful to them for their friendship and sharing their vast experiences and knowledge over past four years. I am also thankful to my supervisory committee members, and my proposal examiners for their critical comments and constructive suggestions.

I am grateful to my family for their unconditional love and affection during all these years.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Overview

The increasing complexity of process plants have made necessary the use of advanced automation and control techniques which aim to improve safety, productivity and quality during operation. However, even with such advanced techniques, process units are frequently subject to faults and disturbances caused by equipment failures, changes in process conditions or human errors. Therefore, there is a need for additional early detection and preventive measures.

This has given rise to a plethora of monitoring and supervisory functions which are studied under abnormal event management (AEM). Other synonyms for AEM are abnormal condition management, fault management or the term abnormal situation management (ASM). AEM involves well-timed detection of an abnormal event, diagnosing its causal origins and then taking appropriate supervisory decisions and actions to bring the process back to a normal, safe operating state. In this respect, AEM has become crucial in safety related process operation.

Fault detection and diagnosis is the central component of AEM. A detection and diagnosis tool aids the operations in the treatment of process malfunctions, improving process safety, reducing the number of shutdowns, downtime and thereby increased productivity.

1.2 Problem statement

Fault detection techniques are typically aimed at detecting operational faults that affect the control objectives of the process. However, in the context of process safety these methods are inadequate as none of the methods take into account the potential impact of the fault on the process and the environment. In order to address this issue, risk-based fault detection method was proposed by Bao *et al.* (2011) [1]. They used univariate charting method to calculate the probability of fault. However, use of univariate method potentially limits the effectiveness of the method due to inherent limitations of univariate fault detection and diagnosis approach. This research attempts to overcome the following limitations of the past work:

- Lack of early warnings prior to the fault impacting the system
- Lack of multivariate approach to take into account the probability of occurrence of the fault as well as potential impact of the fault on the process and the environment
- Lack of practical approach to prioritize warnings and recovery options of a process
- Lack of direct approach to link process deviations to economic loss

1.3 Objectives and scope of the proposed research

1.3.1 Objectives

The main objective of this research is to develop a risk-based fault detection and diagnosis technique targeting the safety issues of a process system. This main objective is achieved through the following sub-objectives:

i) to detect a fault early and to reduce the number of the false alarms

ii) to incorporate consequence information with multivariate fault detection and diagnosis techniques and improve its effectiveness in process monitoring

iii) to develop economic risk models that are linked with process variables and can be dynamically updated

1.3.2 Scope

To achieve the above objectives, this work covers integration of quantitative multivariate model based (Kalman filter and particle filter) and history based (principal component analysis) fault detection methods with dynamic risk assessment.

In addition, a comprehensive economic consequence analysis is proposed for better understanding and quantification of the risk.

Modeling and simulation of some industrial case studies to validate the developed integrated risk-based models are also studied. The scope of the fault detection is limited to qualitative fault detection and diagnosis methods.

1.4 Organization of Proposal

This thesis is written in manuscript format. Outline of each chapter is discussed below

Chapter 2 briefly presents the literature review pertaining to this work. The importance of the risk-based approach as well as the multivariate approach is discussed in this chapter.

Chapter 3 reports the integration of a history based fault detection and diagnosis technique (Principal component analysis) to the risk-based fault detection and diagnosis framework. This chapter is an original research paper published in the *Journal of Industrial & Engineering Chemistry Research, Volume 52, Issue 2, January 2013, Pages: 809-816* [2].

Chapter 4 reports the integration of a model based fault detection and diagnosis technique (Kalman filter) to the risk-based fault detection and diagnosis framework. This chapter is an original research paper published in the *Journal of Process Safety Progress, Volume 32, Issue 4, December 2013, Pages 365–375* [3].

Chapter 5 introduces a new detailed economic consequence methodology. This chapter is an original research paper accepted for publication in the *Journal of Risk Analysis, 2013* [4].

Chapter 6 reports the integration of a nonlinear and non-Gaussian model based fault detection and diagnosis technique (Particle filter) to the risk-based fault detection and diagnosis framework enhanced with the detailed economic consequence analysis. This chapter is an original research paper submitted for publication to the *Canadian Journal of Chemical Engineering, 2013* [5].

Finally, Chapter 7 presents the summary of this work and the main conclusions drawn through this work. Recommendations for future work are presented towards the end of this chapter.

References

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

[2] Zadakbar O., Khan F., Imtiaz S., Dynamic Risk Assessment and Fault Detection Using Principal Component Analysis, Industrial & Engineering Chemistry Research, **2013**, 52, 2, 809–816.

[3] Zadakbar O., Khan F., Imtiaz S., Dynamic risk assessment and fault detection using a multivariate technique, Process Safety Progress, **2013**, 32, 4, 365–375.

[4] Zadakbar O., Khan F., Imtiaz S., Development of Economic Consequence Methodology for Process Risk Analysis, Journal of Risk Analysis, In Press, **2013**.

[5] Zadakbar O., Khan F., Imtiaz S., Dynamic risk assessment using non-linear non-Gaussian fault detection and detailed consequence analysis, Canadian Journal of Chemical Engineering, Submitted, **2014**.

Chapter 2: Fault Detection and Diagnosis

2.1 Overview

The term fault has many connotations in the literature. Typically, any departure of the process variables from an acceptable range is considered as a fault [1]. Simply, a fault is an unpermitted deviation of at least one characteristic property or variable of the process from acceptable/usual/standard behaviour. Fault detection is also defined as a determination of faults present in the system and the time of detection.

Mehra and Peschon (1971) have suggested several approaches to fault detection based on innovations [2]. They presented the special case of linear dynamic systems with Gaussian random inputs and it was shown how the statistical properties of the innovation process can be used for fault detection and diagnosis. Jones (1973) has approached the fault detection problem by using a statistical reference model in invariant continuous systems [3]. Willsky and Jones (1974) have used the maximum likelihood principle to detect jumps in signals [4]. They presented an adaptive filtering system for state estimation and the detection of the jumps using Kalman-Bucy filtering and generalized likelihood ratio techniques. Alburquerque and Biegler (1996) have applied a method based on statistical analysis of data provided by sensors [5-7].

Since the early application of multivariate statistical method based fault detection and diagnosis technique by Kresta *et al.* (1991), data based fault detection and diagnosis has been used extensively in process industries [8]. In such industries, multivariate data-based approaches, for example, principal components analysis (PCA) and partial least squares (PLS), have had the most success in fault detection [6, 9]. Nonlinear versions, for example, nonlinear PLS and kernel PCA, were also applied [10]. Typically, any breakdown in correlation between the process variables is flagged as a fault. In order to detect the change in correlation a test is performed on the residuals. The residuals are processed to generate a signal that indicates the presence of faults in the process. Statistical tests are commonly used to process the residuals [11]. Different statistical tests such as Chi-squared (Ψ^2) test, sequential probability likelihood ratio test (SPRT) and generalized likelihood ratio test (GLR) to be performed on the filtered residual have been developed [12-14]. A literature review showed that these conventional methods like the ordinary Chi-squared (Ψ^2) test and GLR test have limitations such as too many false alarm problems [15].

MacGregor and Kourti, (1996); Wise and Gallagher (1996); Raich and Çinar (1996); Negiz and Çinar, (1997) developed Several history based methods that manipulate the measured data to reduce their dimension and extract information from the data using principle component analysis (PCA) or partial leastsquares (PLS) techniques[8, 16-18]. Aradhye *et al*. (2003) grouped data based on process structure or process distinct timescales as in multi-block or multi-scale PCA [19]. Romagnoli and Palazoglu (2006), Yoon and MacGregor (2001) represented history based methods to extract information about the fault by comparing the location and/or direction of the system with past faulty behavior particularly trajectory in the state space for linear process systems [20-21].

Different model based fault detection methods have been developed in the last 30 years, [9, 14, 22-26]. An extensive comparison of the various methods can be found in [22, 27-31].

Willsky (1976) devised an adaptive filtering system for state estimation and the detection of the jumps using Kalman-Bucy filtering and generalized likelihood ratio techniques [14]. Willsky (1974) described three methods for using parity relations to generate residuals for fault detection [14]. He characterized the notion of analytical redundancy in terms of a generalized parity space.

Isermann (1984) proposed a nonlinear adaptive state observer and a special correlation technique was developed, for the early detection and localization of small leaks in pipelines, based on pressure and flow measurements at the pipeline inlet and outlet [24]. Isermann (1992) also used state space approaches for fault detection in control loops [25]. A new fault detection scheme consists of continuous time parity equations for fast detection has been deduced and described by Iserman (1996) [32].

Li *et al.* (2005) proposed an online and real time sensor and actuator fault detection scheme which takes care of the process disturbances for a multivariate dynamic system, by extending the well-known Chow–Willsky approach [33-34]. Li *et al.* (2008) proposed a Kalman filter-based methodology for unified detection and isolation of sensor, actuator, and process faults [35].

Fault detection methods can be classified based on different criteria depending on whether an explicit model is required or not. Fault detection and diagnosis methods can be classified into model based and history based. The model based methods require some a priori process knowledge. On the other hand, data based methods rely on data history solely. Both model based and history based methods can be classified as quantitative or qualitative [36]. The quantitative model based methods rely on the mathematical relations that exist between variables. However, the qualitative model based methods are based on the cause-effect reasoning about system behavior. The fault-tree is the most popular among these methods. It uses backward chaining until a primary event is found that presents a possible root cause for observed process deviation [37]. Signed Digraphs [38] are another representation of the causal information in which the process variables are represented as graph nodes and causal relations by directed arcs.

Quantitative history based methods are divided to statistical and non-statistical techniques. However, qualitative methods are either rule-based methods or require trend analysis. This classification is shown in figures 2.1 and 2.2. An extensive description of all methods shown in figure 2.1 can be found in [29-31].

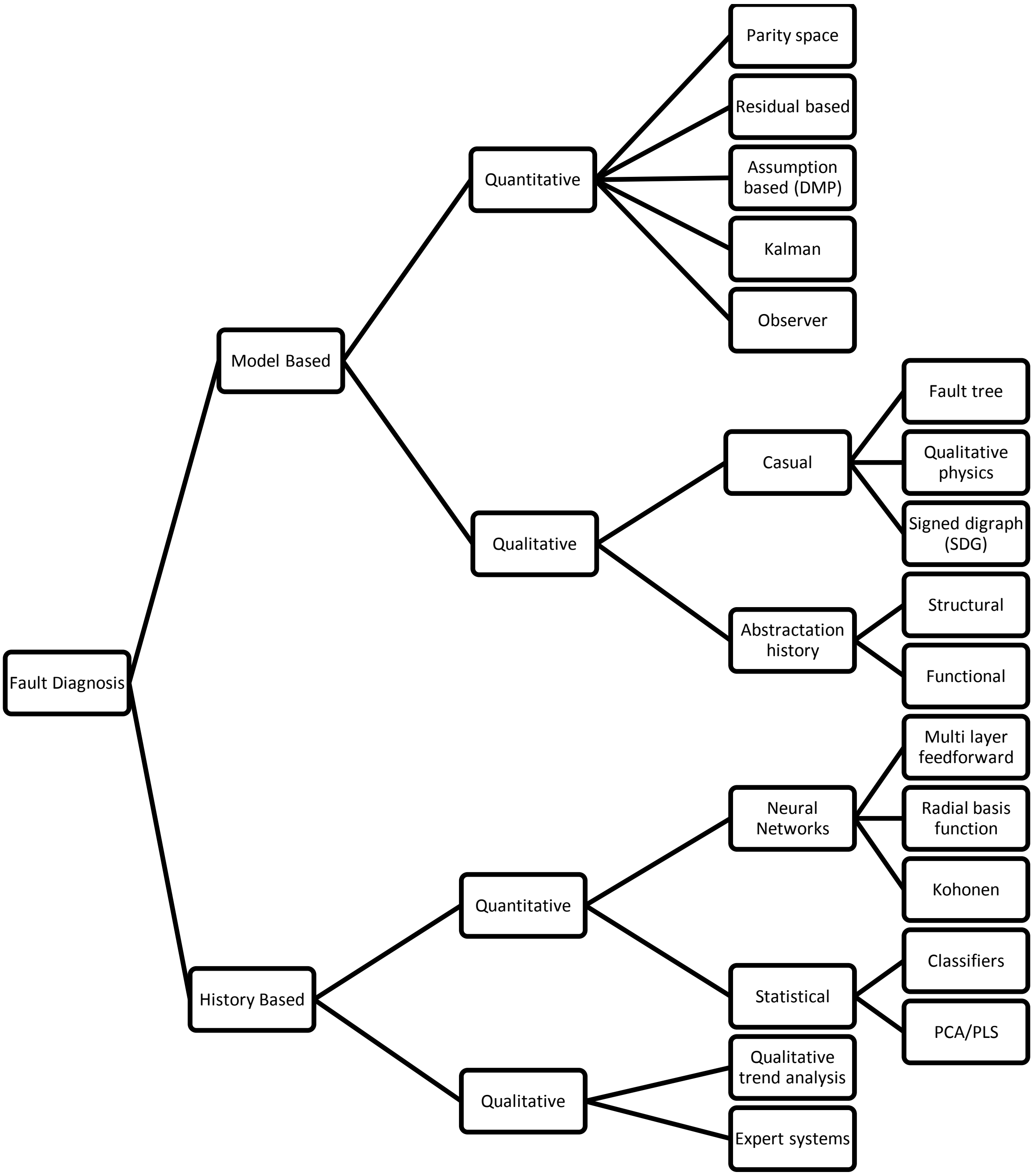

Figure 2.1. The classification of fault detection and diagnosis methods [29-31]

Another way of classifying fault detection methods is based on the number of signals used to detect the fault [39]. Fault detection can done using single signal or multiple signals and models. The signal threshold based approaches for detecting faults consists of checking the measurable variables of a system in regard to a threshold. The signal thresholds can be set into respectively fixed, adaptive, and trend change.

The signal models, such as correlation analysis, Fourier analysis, time series analysis and wavelet analysis, are adopted to extract sample signal features, which including amplitude domain, time domain, frequency domain, power spectrum etc. [40]. A variety of quantitative model based methods as well as quantitative history based methods such as neural network, principal component analysis (PCA) can be classified into the multiple signal models.

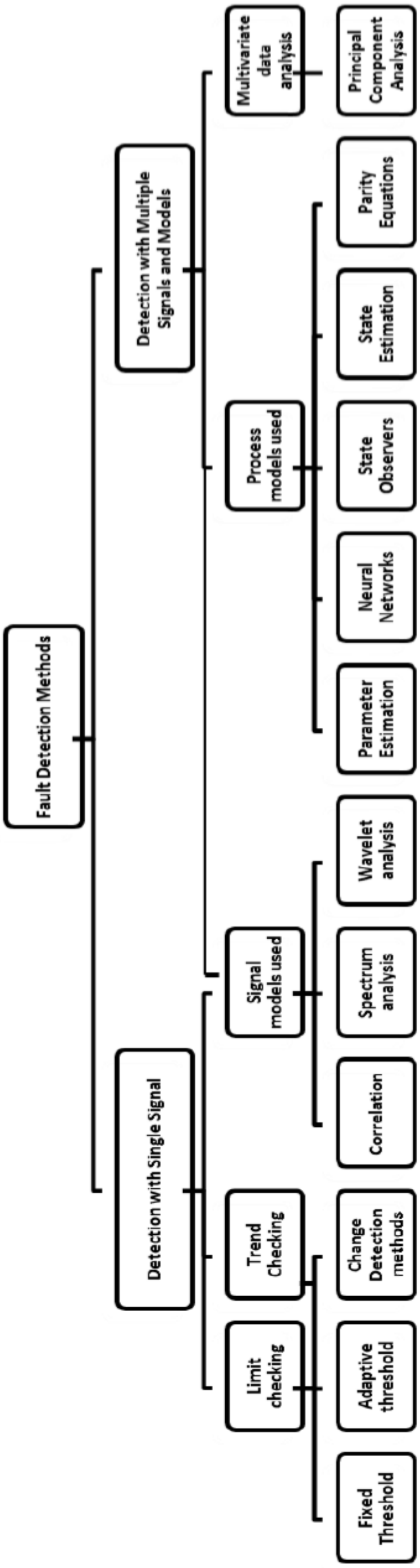

Figure 2.2. The classification of univariate and multivariate fault detection methods [39]

Although all these methods have some advantages in fault detection none of them considers the consequence associated with the fault in detection and invoking warning.

The main aim of this work is to develop a new fault detection method relevant to process safety that incorporates consequence information in fault detection to allow operators to prioritize their action. The proposed risk-based fault detection approach falls under the multivariate category. Methodologies have been developed to implement the strategy in combination with model based as well as history based fault detection methods.

2.2 Why Risk-based Fault Detection

In general fault detection techniques are aimed to detect operational faults that affect the control objectives of the process. However, in the context of process safety these methods are inadequate as none of the methods take into account the potential impact of the fault on the process and the environment. In order to address this issue, risk-based fault detection method was earlier proposed by Bao *et al.* (2011) [41]. Instead of generating an alarm based on residuals or signals crossing the threshold, the risk based fault detection method issues an alarm only when the risk of a process exceeds the acceptable threshold [41]. The risk of a process is defined as a combination of probability of the fault and severity of the fault. This is an important concept as it eliminates faults which are not operational and process safety concerns and also gives a dynamic indication of the operational risk. Bao *et al.* (2011) used a univariate charting method to calculate the probability of fault [41]. This potentially limits the effectiveness of the method due to inherent limitations of univariate fault detection and diagnosis approach. A multivariate risk-based fault detection and diagnosis technique is proposed here. This technique can be implemented with both model-based and history-based approaches.

2.3 Univariate vs Multivariate

Historically, univariate methods, such as limit or trend checking of measurable output variables, are used for fault detection. However, the applicability of univariate method is limited as it is unable to distinguish between normal operational changes and real abnormal faulty conditions. Therefore univariate methods are prone to false alarms [40-41]. More success in fault detection has been observed when multivariable fault detection techniques are applied, as these methods take advantage of the dependence among the process variables (i.e., between input and output), and flags a fault when any deviation from this correlation structure is detected [29-31]. These methods have more power in discerning between operational changes and abnormal conditions. Multivariable methods can be model based, where *a priori* process model is required or it can be purely data based where correlation between the variables are captured from the process data history [42].

2.4 Economic Consequence Analysis

Over the years quantitative risk assessment (QRA) has emerged as an acceptable framework to numerically evaluate the overall risk of a process [43]. QRA was originally used in aerospace, electronics, and nuclear power industries; it is now being employed in the process industries [43–45]. QRA involves four major steps: hazard identification, consequence analysis, frequency assessment, and risk quantification [46]. Consequence analysis is the heart of the risk assessment. Consequence analysis is generally defined as assessment of likely consequences if an accident is to occur. The complexity of a QRA depends on the scenario and the availability of the data and consequence information [45]. Consequence assessment involves a wide variety of mathematical models. For example, source models to predict the rate release of hazardous materials; fire and explosion models; impact intensity models; and toxic gas models [47].

Since 1970, several methodologies have been proposed for quantitative and qualitative risk assessment in process industries [43]. Most of these methodologies are focused on prediction or assessment of failure probability [43-47], and there was not much emphasis given to consequence analysis in evaluating the dynamic risk of a process integrating between process deviation and economic losses.

Taguchi (1986) proposed a quadratic form of loss function to quantify losses associated with deviations of a product characteristic from the target value [48]. The Taguchi loss function has been widely used to determine engineering tolerance. However due to the unbounded and the symmetric characteristic of the Taguchi loss function, it is unrealistic to use this function in many manufacturing processes [49]. The boundlessness of the original Taguchi loss function has been improved by truncating the function at the points where it intersects the maximum loss [50]. On the other hand, asymmetric forms have been proposed [50-52]. Spiring (1993) proposed the reflected or the inverted normal loss function (INLF) in response to criticisms of the quadratic loss functions [49].

A revised inverted normal loss function (RINLF) was also proposed to quantify losses only out of the acceptable region between upper and lower thresholds [53]. In addition to these two major loss functions, a general loss function based on the inversions of gamma loss function has been proposed [54]. Bartholomew *et al* (2004) presented a complimentary general class of inverted probability loss functions (IPLF) based on the inversion of common probability density functions including the inverted normal loss function (INLF), inverted gamma loss function (IGLF), inverted beta loss function (IBLF) and their associated properties for the uniform, normal, gamma and beta distributions [55]. Although they recommended the normal distribution for the INLF, the gamma distribution for the IGLF and the beta distribution for the IBLF, they also suggested some other distributions, such as uniform distribution, to use when the process measurements do not exactly follow the corresponding conjugate distributions [56]. These functions have widely been used in many other engineering areas, particularly in quality control to quantify the economic losses associated with deviations from the target value of a product characteristic [54]. However, these loss functions have not been used in process safety analysis. The present work aims to enhance existing knowledge of consequence analysis through developing a methodology that helps to quantify consequences in economic terms using loss functions.

References

[1] Park S, Himmelblau DM, Fault Detection and Diagnosis via Parameter Estimation in Lumped Dynamic Systems, Ind. Eng. Chem. Process Des. Dev.22:482-487, 1983.

[2] Mehra, RK and Peschon, J, ―An introduction approach to fault detection and diagnosis in dynamic systems", Automatica, 7: 637–640, 1971.

[3] Jones H. L., ―Failure detection in linear systems", PhD thesis, MIT, **1973**.

[4] Willsky A. S., Jones H L, A generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes, Proceedings of the Conference on Decision and Control, Phoenix, Arizona, **1974**.

[5] Albuquerque J. S., Biegler L. T., Data reconciliation and gross error detection for dynamic systems, AIChE Journal, **1996**, 42: 2841–2856,.

[6] Liebmann, M. J., Edgar T. F., Lasdon L. S., Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques, Computers & Chemical Engineering, **1992**, 16, 963,.

[7] Sistu P. B., Gopinath R. S., Bequette B. W., Computation issues in nonlinear predictive control, Computers & Chemical Engineering, **1993**, 17: 361–366.

[8] Kresta J., MacGregor J. F., Marlin T. E., Multivariate Statistical Monitoring of Process Operating Performance, **1991**, The Canadian Journal of Chemical Engineering, 69, 35-47.

[9] Gertler J., ―Fault-detection and diagnosis in engineering systems", New York: Marcel Dekker, **1998**.

[10] Rengaswamy R., Venkatasubramanian V., A syntactic pattern-recognition approach for process monitoring and fault diagnosis. Engineering Applications of Artificial Intelligenc, **1995**, 8, 1,35-51.

[11] Orchard M. C., Cipriano A. M., Cipriano A. Z., Viale M., Vigliocco A., A model based fault detection and diagnosis system for rolling mill equipment, Proceedings of the European Control Conference, Porto, Portugal, **2001**.

[12] Mariabadi A., Mort N., Schmid F., A fault tolerant train navigation system using multisensory multifilter integration technique, FUSION 98 International Conference **1998**.

[13] Kerr T. H., Decentralized filtering and redundancy management for multisensory navigation, IEEE Transaction on Aerospace and Electronic Systems, **1987**, 23, 1, 83-119,.

[14] Willsky A. S., A survey of design methods for failure detection systems, Automatica, **1976**, 12, 601–611.

[15] Asokan A., Sivakumar D., Model Based Fault Detection and Diagnosis Using Structured Residual Approach in a Multi-Input Multi-Output System, Serbian Journal Of Electrical Engineering, **2007**, 4, 2, 133-145.

[16] Wise B. M. , Gallagher N.B., The process chemometrics approach to monitoring and fault detection, Journal of Process Control, **1996**, 6, 329–348.

[17] Raich A., Çinar A., Statistical process monitoring and disturbance diagnosis in multivariable continuous processes, AIChE Journal, **1996**, 42, 995–1009.

[18] Negiz A., Çinar A., Statistical monitoring of multivariable dynamic processes with state-space models, AIChE Journal,**1997**, 43, 2002–2020.

[19] Aradhye H.B., Bakshi B.R., Strauss R.A., Davis J.F., Multiscale SPC using wavelets: theoretical analysis and properties, AIChE Journal, **2003**, 49, 939–958.

[20] Romagnoli J., Palazoglu A., Introduction to Process Control. CRC Press, Baco Raton, **2006**.

[21] Yoon S., MacGregor .J, Fault diagnosis with multivariate statistical models part I: using steady state fault signatures. Journal of Process Control, **2001**, 11, 387–400.

[22] Himmelblau D. M., Fault-detection and diagnosis in chemical and petrochemical processes, Amsterdam: Elsevier, **1978**.

[23] Frank P. M., Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy, Automatica, **1990**, 26, 459–474.

[24] Isermann R., Process fault-detection based on modelling and estimation methods – a survey, Automatica, **1984**, 20, 387–404.

[25] Isermann R., Deibert R., Examples for fault detection in closed loops, Annual Review in Automatic Programming, **1992**, 17, 235-240.

[26] Isermann R., Supervision, fault-detection and fault-diagnosis methods: An introduction, Control Engineering Practice, **1997**, 5, 639–652.

[27] Mylaraswamy D., DKit: a blackboard-based distributed multi expert environment for abnormal situation management, PhD thesis, Purdue University, **1996.**

[28] Rengaswamy R., Venkatasubramanian V., A syntactic pattern-recognition approach for process monitoring and fault diagnosis. Engineering Applications of Artificial Intelligenc, **1995**, 8, 1,35-51.

[29] Venkatasubramanian V., Rengaswamy R., Kavuri S.N., Yin K., A review of process fault detection and diagnosis, Part I: Quantitative model-based methods, Computers and Chemical Engineering, **2003**, 27, 293-311.

[30] Venkatasubramanian V., Rengaswamy R., Kavuri S.N., Yin K., A review of process fault detection and diagnosis, Part II: Qualitative models and search strategies, Computers and Chemical Engineering, **2003**, 27, 313-326.

[31] Venkatasubramanian V., Rengaswamy R., Kavuri S.N., Yin K., A review of process fault detection and diagnosis, Part III: Process history based methods, Computers and Chemical Engineering, **2003**, 27, 327-346.

[32] Höfling T., Isermann R., Fault detection based on adaptive parity equations and single-parameter tracking. Control Engineering Practice, **1996**, 4, 10, 1361-1369.

[33] Li W., Shah S. L., Han Z., Fault detection and isolation in the presence of process uncertainties, Control Engineering Practice, **2005**, 13, 5, 587-599.

[34] Chow E. Y., Willsky A. S., Analytical redundancy and the design of robust failure detection systems, IEEE Transaction on Automatic Control, **1984**, 29, 7, 603-614.

[35] Li W., Shah S. L., Xiao D., Kalman filters in non-uniformly sampled multirate systems: For FDI and beyond, Automatica, **2008**, 44, 1, 199-208.

[36] Kaszkurewicz E., Bhaya A., Ebecken N. E., A Fault Detection and Diagnosis Module for Oil Production Plants in Offshore Platforms, Expert Systems With Applications, **1997**, 12, 2, 189-194.

[37] Lapp S. A., Powers G. A., Computer-aided synthesis of fault-trees, IEEE Transactions on Reliability, **1977**, 37, 2-13.

[38] Iri M., Aoki K., O'Shima E., Matsuyama H., An algorithm for diagnosis of system failures in chemical processes, Computers & Chemical Engineering, **1979**, 3, 489-493.

[39] Isermann R., Fault-diagnosis systems: an introduction from fault detection to fault tolerance, Springer, **2006**.

[40] Shui A., Chen W., Zhang P., Hu S., Huang X., Review of fault diagnosis in control systems, Annual International Conference on Chinese Control and Decision, 5360-5369, **2009**.

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

[42] Isermann R., Model-Based Fault-Detection and Diagnosis – Status and Applications, Annual Reviews in Control, 2005, 29, 71–85.

[43] Khan F., Abbasi S. A., Techniques and methodologies for risk analysis in chemical process industries, Journal of Loss Prevention in the Process Industries, **1998**; 11, 261–277.

[44] Lees F. P., Loss Prevention in the Process Industries, Butterworth, New York, **1996**.

[45] CCPS/AIChE, Guidelines for Chemical Process Quantitative Risk Analysis, Wiley, New York, **2000**.

[46] Pula R., Khan F., Veitch B., Amyotte P., A Grid Based Approach for Fire and Explosion Consequence Analysis, IChemE Transactions, Part B, Process Safety and Environmental Protection, **2006**, 84(B2), 79–91.

[47] Crowl D. A., Louvar J. F., Chemical Process Safety: Fundamentals with Applications, Prentice Hall PTR, Upper Saddle River, New Jersey, **2002**.

[48] Taguchi G., Introduction to Quality Engineering: Designing Quality into Products and Processes. Kraus, White Plains, NY, **1986**.

[49] Spiring F. A., The Reflected Normal Loss Function, Canadian Journal of Statistics, **1993,** 21, 321-330.

[50] T.P. Ryan, Statistical Methods for Quality Improvement, Wiley, New York, NY, **1989**.

[51] Phadke M.S., Quality Engineering Using Robust Design, Prentice Hall, New York, NY, **1989**.

[52] Baker T.B. Engineering Quality by Design: Interpreting the Taguchi Approach, Marcel Dekker, New York, NY, **1990**.

[53] Sun F., Laramee J., Ramberg J., On Spiring's Inverted Normal Loss Function, Canadian Journal of Statistics, **1996**, 24, 241-249.

[54] Pan J., Pan J., A Comparative Study of Various Loss Functions in the Economic Tolerance Design, IEEE International Conference on Management of Innovation and Technology, **2006**.

[55] Leung, B. P., Spiring F. A., Some properties of the family of inverted probability loss functions. Quality Technology and Quantitative Management, **2004**, 1, 1, 125-47.

[56] Pan J.N., Wang J.H., A study of loss functions for product interference analysis", Industrial Engineering Research, **20002**, 1, 80-100.

Chapter 3: Dynamic Risk Assessment and Fault Detection Using Principal Component Analysis†

Preface

A version of this chapter has been published by the *Journal of Industrial & Engineering Chemistry Research*, in 2013. Zadakbar was the main lead on the work. The co-authors, Dr. Khan and Dr. Imtiaz supervised the work and helped to develop the methodology.

Abstract

A methodology to calculate process risk in combination with data based fault detection method is proposed in this chapter. The proposed approach aims to identify and screen the faults which are not safety concern and also to dynamically update process risk at each sampling instant. The approach is built upon principal components analysis (PCA) combined with a quantitative operational risk assessment model. Through this approach a warning system is activated only when the risk of operation exceeds the acceptable threshold. Combining PCA with the risk assessment model makes this approach more robust against false alarms. Application of this new risk-based approach provides early warnings and early activation of safety systems prior to the fault impacting the system. This method has more power in discerning between operational deviations and

† *O. Zadakbar, F. Khan, S. Imtiaz, Industrial & Engineering Chemistry Research, 2013, 52 (2), pp 809–816, DOI: 10.1021/ie202880w, Publication Date (Web): November 26, 2012, Copyright © 2012 American Chemical Society*

abnormal conditions which potentially may cause an unwanted situation (for example: an accident).

KEYWORDS

Dynamic Risk Assessment, Fault Detection, PCA

3.1 Introduction

The need for fault detection and diagnosis tools to monitor complex chemical processes is well established. Both model based fault detection and historical data based methods have been used for process fault detection and diagnosis. In chemical processes data based methods are often the only choice due to the unavailability of models of complex processes. Also, in many cases the cost associated with building a first principles model cannot be justified by the marginal gain in model based fault detection and diagnosis compared to using the multivariate statistical methods. Classification of fault detection methods and an extensive comparison of the various methods can be found in [1-4]. Since the early application of multivariate statistical method based fault detection and diagnosis technique by [5], data based fault detection and diagnosis has been used extensively in process industries. In such industries, multivariate data based approaches, for example, PCA and partial least squares (PLS), have had the most success in fault detection [6-8]. Nonlinear versions, for example, nonlinear PLS and kernel PCA, were also applied [4]. Typically, any breakdown in correlation between the process variables is flagged as a fault. In order to detect the change in correlation a test is performed on the residuals.

In general, fault detection techniques are aimed to detect operational faults that affect the control objectives of the process. The residuals are processed to generate a signal that indicates the presence of faults in the process. Different statistical tests such as Chi-squared (Ψ^2) test [9], sequential probability likelihood ratio test (SPRT) [10], and generalized likelihood ratio test (GLR) [11, 12] have been developed to test the residuals. Although all these methods have some advantages in fault detection none of these methods take into account the potential impact of the fault on the safety of process, personnel and the environment, as such warnings can be generated for trivial changes in the process which do not have a big impact on process safety or operation. Thus these methods often generate spurious alarms which can lead to alarm flooding in the control room.

To address this issue risk-based fault detection method was earlier proposed by Bao *et al.* (2011) [13]. Instead of generating warning based on residuals or signals crossing the threshold, the risk-based fault detection method issues warning when the risk of a process exceeds the acceptable threshold. The operational risk is defined as a combination of probability of a fault and severity of the fault. This is an important concept as it eliminates faults which are not safety concerns and also gives a real time indication of the operational risk. Bao *et al.* (2011) [13] combined a univariate charting method with risk calculation. The probability of fault was calculated based on the deviation of signal from threshold. Univariate methods do not take into account the changes in the input. Therefore they have less power in discerning between process fault and operational changes. This potentially limits the effectiveness of the risk-based method based on univariate fault detection and diagnosis approach.

Zadakbar *et al.* (2011) [14] extended the method to a multivariate framework. Residuals generated from a Kalman filter were used to calculate the probability of fault. However, as Kalman filter requires a model its application is limited due to lack of availability of a model for complex processes. Therefore, there is a need to develop a model-free risk-based fault detection and diagnosis method.

In this Chapter a multivariate risk-based fault detection and diagnosis technique using historical process data is proposed. We extended principal components analysis in risk-based framework to detect process faults which are safety concern to the process.

3.1.1 Organization of the Chapter

The chapter is organized as follows: in Section 3.2, the risk-based fault detection and diagnosis methodology is introduced. This methodology is based on combining the PCA and risk assessments. In Section 3.3, the application of the proposed methodology is demonstrated on two process systems. The first case study, presented in Subsection 3.3.1, is a simulated distillation column while the second case study, presented in Subsection 3.3.2, is a real life industrial case presenting a dissolution tank. Finally, conclusion is presented in Section 3.4.

3.2 Risk-based Fault Detection using PCA

The overall methodology for process risk calculation is shown in Figure 3.1. In the proposed methodology we calculate the process risk at any instant using the output from PCA. At any instant (starting time: θ=0) the measurements from sensors are first projected on the principal components (PC) subspace using the previously calculated loadings. This gives scores in the PCs directions. Typically these calculated scores will contain some noise. In order to suppress the effect of noise the scores are filtered using an exponential filter.

The next step is to calculate the probability of fault and the associated severity. Probability of fault is calculated using the filtered score. Severity of fault is calculated using the filtered scores and the loadings calculated from the offline PCA. Subsequently severity and probability of fault are combined together to get the risk at any sampling instant. The definition of risk and the relevant equations are given in Section 3.2.2.

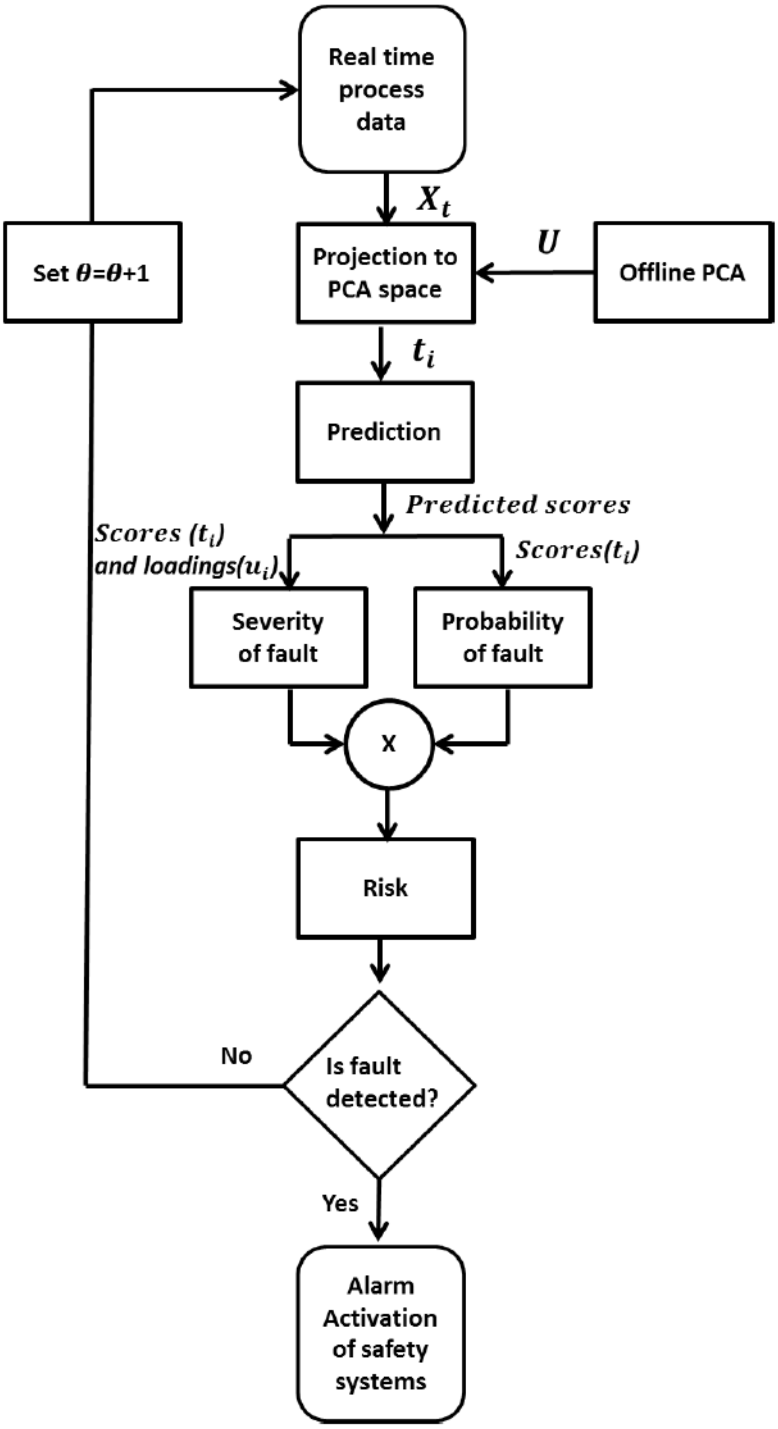

Figure 3.1. Methodology of risk-based fault detection

The risk profile is used for fault detection as well as for taking any supervisory decisions to activate appropriate safety systems in real time. If any fault is detected the next step will be to take steps to mitigate the fault. If fault is not detected then proceed to the next time step.

3.2.1 Principal Component Analysis In Fault Detection

PCA was originally developed as an effective data compression tool for multivariate data analysis. In process industries PCA is widely used as a multivariate statistical process control (SPC) tool for monitoring large number of process variables [5, 15, 16]. In this section we give the relevant theory on PCA. An excellent tutorial on the theory and application of PCA can be found in [17].

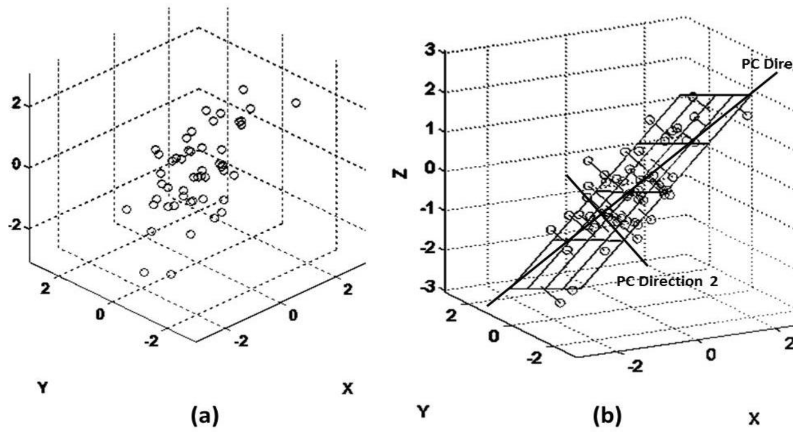

Figure 3.2. Orthogonal transformation and axis rotation by PCA

Consider data matrix $X^{n\times p}$ where n denotes the number of samples and p denotes the number of process variables. Each column in the data matrix is a mean centered and scaled process variable.

There are infinite numbers of linear transformations that can be applied on a data matrix. In PCA the data matrix is transformed to the maximum variability directions of the data matrix. This is depicted in Figure 3.2. The correlated data was originally lying in a three dimensional space (Figure 3.2a). However, most of the variability of the data is only in two directions. Therefore the data is projected in a two dimensional subspace (Figure 3.2b). The coordinates of this new subspace are aligned with the two most variability directions of the data. These two directions are called PC directions. The values relative to this new coordinate system are called scores, T.

To find the appropriate projection matrix singular value decomposition (SVD) is applied to the covariance matrix of $X^{n \times p}$ which decomposes the matrix to a diagonal matrix δ and a set of orthonormal variables collected in matrix $U^{p \times p}$. Mathematically, we can write the covariance matrix $\Sigma = U \delta U^T$.

The columns of U are called the PC loading vectors. The diagonal elements of δ are called eigenvalues of the covariance matrix. The scores, T is given by:

$$T = XU \tag{3.1}$$

Process data contains both systematic process variation and random process noise. Therefore data matrix, X can be written in the following form:

$$X = \hat{X} + E \tag{3.2}$$

Where $\hat{X}$ is the systematic process variation and E is the random process noise.

Accordingly the scores and loadings matrix is partitioned in two parts: $T = [T_r\ T_e]$ and $U = [U_r\ U_e]$ explaining the systematic variation, $\hat{X} = T_r\ U_r^T$ and random variation, $E = T_e\ U_e^T$ in the data, respectively. The number of retained PCs, r should be ideally chosen such that there is no significant process information left in the covariance matrix. In theory, the E matrix should contain only the random error.

Once the number of retained PCs has been decided, the data matrix X can be projected in the PC direction:

$$T_r = XU_r \tag{3.3}$$

Where $T_r^{n\times r} = [t_1, t_2, t_3, \dots, t_r]$ contains the scores of PCs. Each PC or score vector, t_i collected in the matrix T_r and captures as much variation as possible which has not been explained by the former PCs.

Often PCs give an early indication of fault, therefore in process monitoring instead of monitoring individual variables, PCs are monitored. Statistical test is done on each score value and if the value exceeds the threshold it is considered as a fault. However, conventional PCA methodology does not take into account the consequence of the faults and therefore gives equal weightage to all faults. In the proposed methodology we calculate the associated risk for a PC's score exceeding threshold. The scores and the loadings of the PCs are the main inputs for the dynamic risk calculation. In the following section we give the details of dynamic risk calculation methodology for each PC.

3.2.2 Dynamic Risk Calculation for Principal Component

Risk is defined as a measure of likely harm or loss caused by the fault if corrective action is not taken on time. Risk depends on two factors: the probability of occurrence of a fault leading to an unwanted event and severity of the loss caused by the event [18]. Bao *et al.* (2011) [13] proposed the following formula for a univariate deterministic system:

$$Risk = P \times S \tag{3.4}$$

Where P is the probability of occurrence of a fault, and S is the severity of the fault. The probability of the fault for each predicted score point is calculated using Equations (3.5) through (3.7). The probability of an event increases as the process moves away further from the normal operation. This behavior can be captured using a cumulative normal distribution. Therefore the probability of fault is defined as,

$$P = \varphi\left(\frac{x-\mu}{\sigma}\right) \tag{3.5}$$

This also allows us to define the threshold for the normal operating region. We use $\mu - 3\sigma$ and $\mu + 3\sigma$ as the lower and upper threshold for the normal operation. All data points between these two thresholds are considered as normal which is approximately (99.73%) of all the values under normal operating conditions. If any predicted point goes outside this region; it could lead to a fault leading to an unwanted event. For example, when a signal value is at the threshold ($\mu \pm 3\sigma$) the probability of fault is 0.5 as it can either go back to normal or may keep growing and ultimately lead to an event. In the present study we used the same definition with modifications to use the scores instead of original signal. For fault signals when the scores approach the upper threshold, the probability of the fault is calculated by $\varphi\left(\frac{t_{ij}-(\mu_i+3\sigma_i)}{\sigma_i}\right)$. On the other hand, if the scores approach the lower threshold the probability of the fault is calculated by the complement, $1 - \varphi\left(\frac{t_{ij}-(\mu_i-3\sigma_i)}{\sigma_i}\right)$,

$$for\ t_{ij} > \mu_i \longrightarrow\ P = \varphi\left(\frac{t_{ij}-(\mu_i+3\sigma_i)}{\sigma_i}\right) = \int_{-\infty}^{t_{ij}} \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{-(t_{ij}-(\mu_i+3\sigma_i))^2}{2\sigma_i{}^2}} dt \tag{3.6}$$

$$for\ t_{ij} < \mu_i \longrightarrow\ P = 1 - \varphi\left(\frac{t_{ij}-(\mu_i-3\sigma_i)}{\sigma_i}\right) = 1 - \int_{-\infty}^{t_{ij}} \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{-(t_{ij}-(\mu_i-3\sigma_i))^2}{2\sigma_i{}^2}} dt \tag{3.7}$$

Where:

$i = 1, \cdots, n \quad n = number\ of\ PCs$

$j = 1, \cdots, m \quad m = number\ of\ samples$

Figure 3.3 shows a visual depiction of a fault probability. Probability of the fault for a point on the centerline μ, is 0 and for a given point on the thresholds is equal to 0.5.

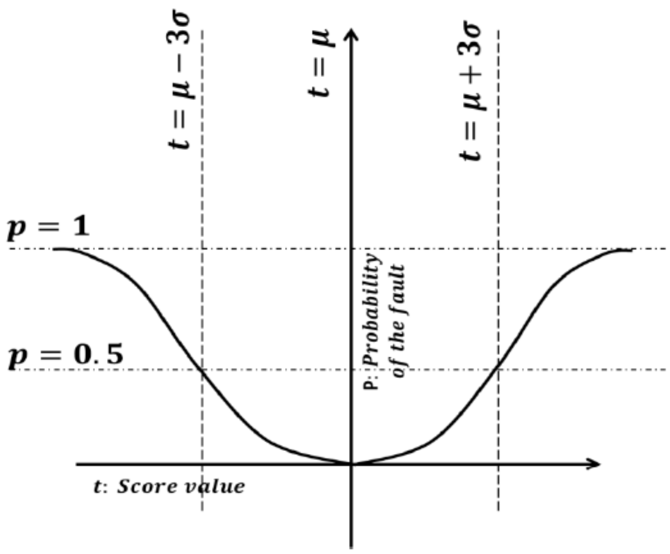

Figure 3.3. Change in probability of fault (φ) with deviation of score from mean

The severity of fault is calculated using Equations 8 through 10. It is a modified version of the equation proposed by Bao *et al.* (2011) [13]. The modifications were required because in this study we use scores of PCs to calculate the risk instead of residuals used in [14].

Each PC is a weighted combination of the original process variables. Violation of threshold by these PCs therefore indicates a fault in the process. The severity of the fault is assessed using Equations (3.7) through (3.9). In this case we calculate the severity based on the scores. Severity has two parts: exponential part and the pre-exponential term. The exponential part gives the magnitude of fault based on the exceedance of the threshold. The pre-exponential term called intensity coefficient gives a relative measure of hazard potential of each PC. Intensity coefficient of a score, a' is calculated as a weighted average of the intensity coefficient of individual variables (a_k) as given in Equation 10. The weights, w_k are the absolute loadings in each PC.

$$for\ t_{ij} > \mu_i \dashrightarrow\ S = a'.(100)^{\frac{t_{ij}-(\mu_i+3\sigma_i)}{t_{ij}-\mu_i}} \tag{3.8}$$

$$for\ t_{ij} < \mu_i \dashrightarrow\ S = a'.(100)^{\frac{t_{ij}-(\mu_i-3\sigma_i)}{\mu_i-t_{ij}}} \tag{3.9}$$

$$a' = \sum_{k=1}^{q} w_k a_k \dashrightarrow\ S = \sum_{k=1}^{q} w_k S_k \tag{3.10}$$

Where:

$i = 1, \cdots, n \qquad n = number\ of\ PCs$

$j = 1, \cdots, m \qquad m = number\ of\ samples$

$k = 1, \cdots, q \qquad q = number\ of\ original\ process\ variables$

Intensity coefficient, a indicates the intensity of the severity of the fault associated with each process variable. For instance, in a simple chemical reactor containing nonhazardous chemicals, the severity of hazard due to uncontrolled temperature (exothermic reaction) is much higher than an uncontrolled concentration of a given component. Thus, a associated with temperature is larger than a associated with concentration. Coefficient a is assigned based on process and operational considerations e.g. process nature, number of people at risk, chemical components, environment and costs.

In order to remove the noise a further moving average filter was applied to the calculated score signal. The filter coefficient of the moving average filter α is a tuning parameter which is further tuned to minimize the false alarms.

$$t_{ij} = \alpha t_{ij} + (1-\alpha)t_{ij-1} \tag{3.11}$$

The next step in the methodology is prediction. Instead of using the scores directly predicted score values are used for risk calculation. This gives some lead time in fault detection. We used a four point backward difference formula for slope calculation. Based on the slope of the filtered scores we predict the scores for four forward points.

$$\left.\frac{dt_{ij}}{d\theta}\right|_{\theta=\theta_j} = \frac{37t_{ij}+36t_{ij-1}-9t_{ij-2}+8t_{ij-3}}{30} \tag{3.12}$$

Where θ is time. The risk of the fault is obtained by multiplying the severity of fault and the probability of fault. If the risk exceeds the threshold an alarm is issued. Subsequently based on assessed risk, supervisory decisions to activate appropriate safety systems are taken.

Multivariate risk-based fault detection using PCA is capable of assessing the contribution of each original variable to the detected fault and finding the variable with the most risk potential using the loadings of a given PC under faulty conditions. Equations 8 and 9 can be expanded to the following expression to determine the contribution of each original variable. S_k is the severity of each original variable in total severity S. Consequently, to diagnose the fault, the associated risk of each original variable, R_k, is calculated using S_k.

$$for\ t_{ij} > \mu_i \dashrightarrow \ S = \sum_{k=1}^{q} w_k a_k .(100)^{\frac{t_{ij}-(\mu_i+3\sigma_i)}{t_{ij}-\mu_i}} = w_1 a_1.(100)^{\frac{t_{ij}-(\mu_i+3\sigma_i)}{t_{ij}-\mu_i}} + w_2 a_2.(100)^{\frac{t_{ij}-(\mu_i+3\sigma_i)}{t_{ij}-\mu_i}} + \cdots w_q a_q.(100)^{\frac{t_{ij}-(\mu_i+3\sigma_i)}{t_{ij}-\mu_i}} = \sum_{k=1}^{q} S_k \quad (3.13)$$

$$for\ t_{ij} < \mu_i \dashrightarrow \ S = \sum_{k=1}^{q} w_k a_k .(100)^{\frac{t_{ij}-(\mu_i-3\sigma_i)}{\mu_i-t_{ij}}} = w_1 a_1.(w_1 a_1.100)^{\frac{t_{ij}-(\mu_i-3\sigma_i)}{\mu_i-t_{ij}}} + w_1 a_1.(100)^{\frac{t_{ij}-(\mu_i-3\sigma_i)}{\mu_i-t_{ij}}} + \cdots w_q a_q.(100)^{\frac{t_{ij}-(\mu_i-3\sigma_i)}{\mu_i-t_{ij}}} = \sum_{k=1}^{q} S_k \quad (3.14)$$

$$R_k = S_k \times P \quad (3.15)$$

3.3 Case Studies

The application of the proposed methodology is demonstrated on two process systems. In the first case study the methodology is applied to a simulated 40 stage distillation column. The second case study is an industrial application of a dissolution tank where solid crystals are dissolved with water.

3.3.1 Distillation Column

The system considered in this study is a binary distillation column with 40 stages that separates a binary mixture with relative volatility of 1.5 into products of 96% purity (Figure 3.4) [19]. The assumptions considered to model the distillation unit are: binary mixture, equilibrium on all stages, constant pressure and relative volatility, total condenser, no vapour holdup and linearized liquid dynamics.

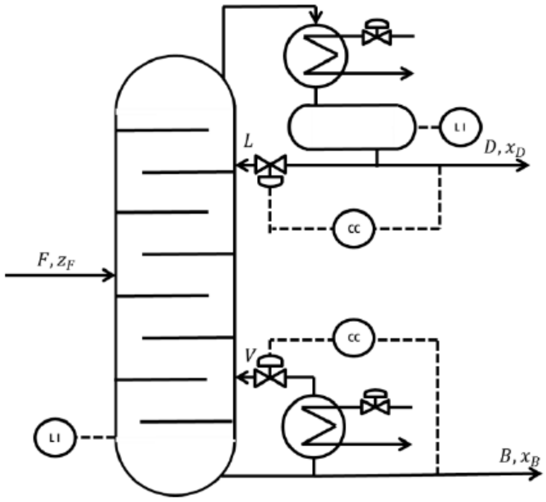

Figure 3.4. Schematic diagram of binary distillation column

The linearized dynamic model has 82 states, 6 inputs and 4 output variables. The first 41 states are compositions of light component in reboiler, condenser and 39 trays in between. In the next 41 states are the holdups in the 39 trays, the reboiler and the condenser. Inputs are reflux flow rate (L), boilup flow rate (V), top or distillate product flow (D), bottom product flow (B), feed rate (F), and feed composition (z_F). There are four sensors to measure top composition (x_D), bottom composition (x_B), condenser holdup (M_D) and reboiler holdup (M_B).

In this study a gradual decrease of reboiler heat flow is considered as a fault. This fault may cause increase in bottom flow rate with the time which would affect other process states, e.g. top and bottom concentrations. It is considered that the reboiler fault occurred at θ =2000s which also shows up at the bottom flowrate at the same time (Figure 3.5).

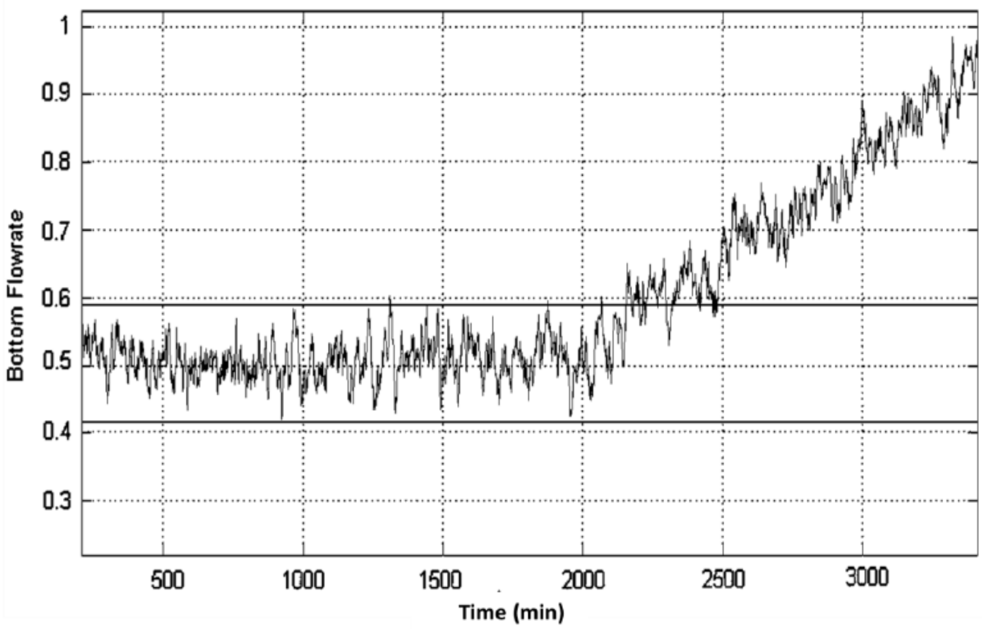

Figure 3.5. Bottom flow rate (B) of distillation unit in faulty conditions

In this case we monitor the following process variables: feed composition (z_F), feed flowrate (F), distillate composition (x_D), distillate flowrate (D), bottom composition (x_B), and bottom flowrate (B). These variables are collected in data matrix, X.

$$X = \begin{bmatrix} z_F \\ F \\ x_D \\ D \\ x_B \\ B \end{bmatrix} \tag{3.16}$$

PCA was applied on these variables. Based on the results of PCA in Table 3.1 the first three PCs are retained which capture about 90% of the total variation.

Table 3.1. Principal component analysis for the first case study

PC Number	% Variance captured this PC	% Variance captured total
1	48.99	48.99
2	23.7	72.69
3	16.7	89.39
4	8.16	97.56
5	2.05	99.6
6	0.4	100

Among the retained PCs, fault was detected in the scores of PC1 as can be seen in Figure 3.6. The scores of PC1 were filtered and based on slope the future scores were calculated. Subsequently these predicted values were used for risk calculation. Risk calculation has two parts: probability of fault and severity calculation. Probability of fault is calculated based on Equations 3.5 and 3.6.

In order to calculate the severity of a score, we first assign the severity of each process variables considered in the analysis. We give a relative number to the severity coefficient for each variable based on process knowledge. Among these variables concentration is a quality variable and the consequence associated with concentration is poor product quality; thus it has the lowest potential to cause hazards. Therefore we assigned concentrations a severity value of 1 and relative to that the severity of other variables were assigned. Relative to the concentration, the feed flowrate has twice hazard potential and the top and the bottom flowrate have three times hazard potential. These values are somewhat subjective at this stage. A detailed cost analysis is necessary to assign more accurate values which will be outlined in a forthcoming article.

$$a = \begin{bmatrix} a_1 = a_{x_F} \\ a_2 = a_F \\ a_3 = a_{x_D} \\ a_4 = a_D \\ a_5 = a_{x_B} \\ a_6 = a_B \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \\ 1 \\ 3 \end{bmatrix} \quad (3.17)$$

Loadings of each PC calculated by PCA are used as weights in Equation 3.10. For the first PC loadings are:

$$w = \begin{bmatrix} w_1 = w_{x_F} \\ w_2 = w_F \\ w_3 = w_{x_D} \\ w_4 = w_D \\ w_5 = w_{x_B} \\ w_6 = w_B \end{bmatrix} = \begin{bmatrix} 0.27 \\ 0.27 \\ 0.04 \\ 0.03 \\ 0.12 \\ 0.27 \end{bmatrix} \quad (3.18)$$

Substituting these values we get the severity intensity coefficient for PC1.

$$a' = \sum_{i=1}^{6} |w_i a_i| = 1.87 \quad (3.19)$$

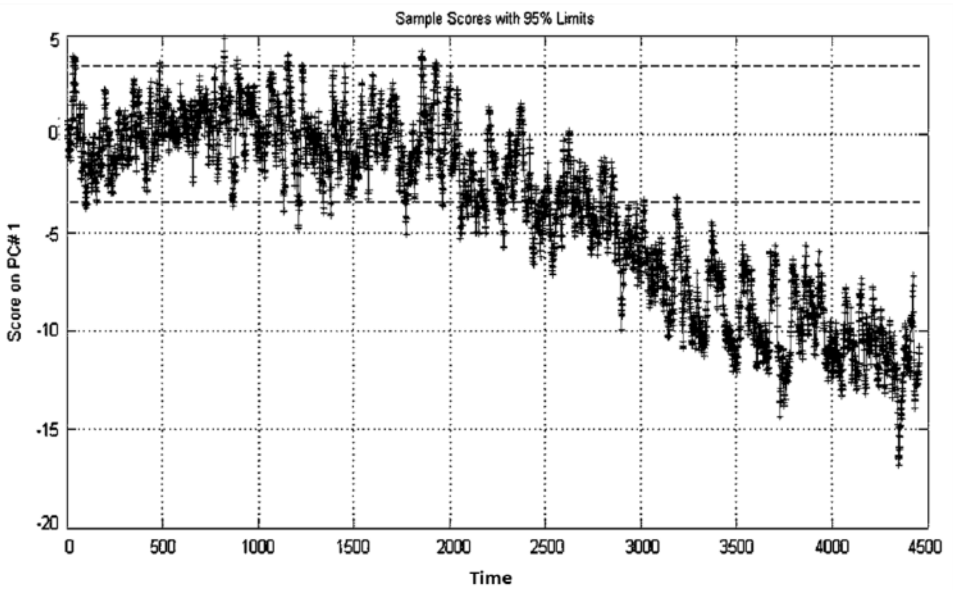

Figure 3.6. First PC's score for the distillation unit data

The risk of the fault is obtained by combining the effect of severity with the probability of fault. Figure 3.7 shows the risk profile for the first PC.

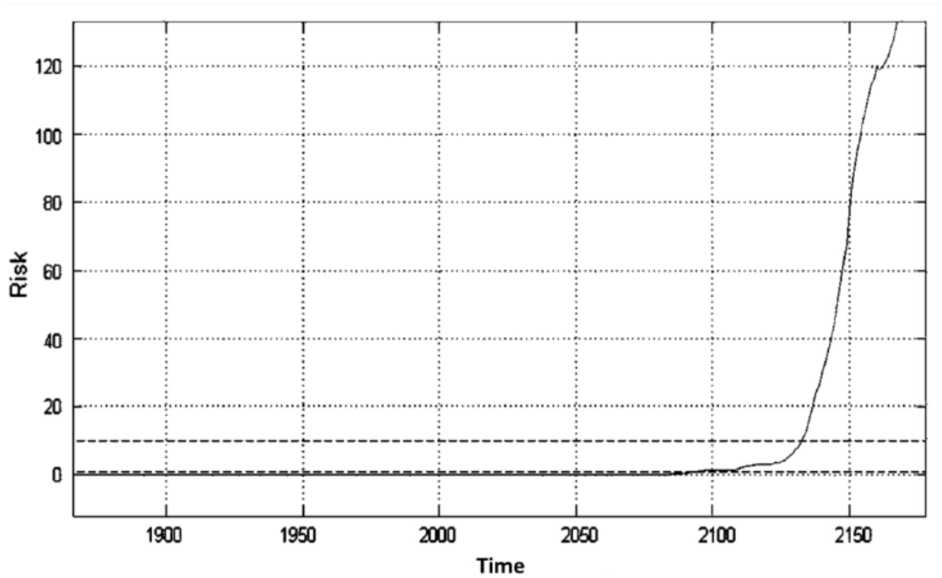

Figure 3.7. Risk profile of the first PC of the distillation unit data

The threshold for the risk signal is based on the acceptable risk criteria of the specific process system. In Figure 3.7 few guiding principles of acceptable risk in process operation are shown. The first risk threshold is at 1. It is used for fault detection in parallel with the warning generation for the operators to take priority response if the automatic systems failed. If no corrective action is taken or the corrective action fails to bring down the risk, and risk exceeds the second threshold then the automatic safety system is activated (emergency shutdown system). Application of the intensity coefficient enhances the ability of the methodology by incorporating the impact of the fault on operational performance and potential accident.

The application of the risk-based fault detection provides early warnings that can be used to correct process fault before it can lead to a catastrophic event. Based on the analysis, an alarm is activated at θ=2090s when the risk of operation exceeds the acceptable threshold (first level) instead of generating an alarm at θ=2400s based on original signal, i.e. bottom flowrate (B), crossing its threshold.

3.3.2 Dissolution Tank

The risk-based methodology is applied to a data set collected from a pure terepthalic acid (PTA) crystal dissolution tank operation. A simplified process diagram for the system is shown in Figure 3.8. Solid crystals are dissolved in a tank with water. Water is pumped into the tank under flow control. A rotary feeder is used to feed the dissolution tank from a hopper. The feed rate of solid crystals to the tank is controlled by the speed of the rotary feeder. The water level in tank and the density of the liquid going out of the tank are measured variables.

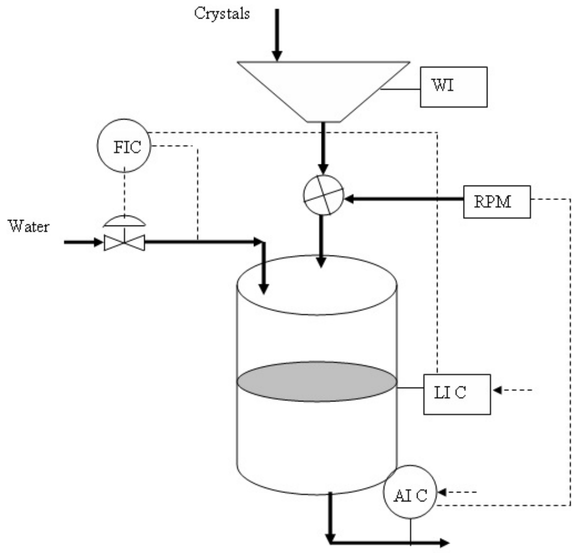

Figure 3.8. Schematic diagram of PTA industrial dissolution tank

The process is known to be impacted by severe disturbances which if undetected can lead to catastrophic faults. Occasionally, because of the variation in moisture content the solid gets lumped in the rotary feeder and does not dispense from the feeder uniformly. After a while when the solid lump gets too big it falls into the tank creating a big disturbance in the density which causes problems in the downstream process. One specific example is shown in Figure 3.9.

In this case because of the actuator problem in the rotary drum, excess solid crystals dropped into the tank at θ =10365 min. Figure 3.10 shows the effect of the excess solid crystals on the density. It creates a large spike in the density at θ =10519 min when operator is able to detect the fault and subsequently takes corrective action.

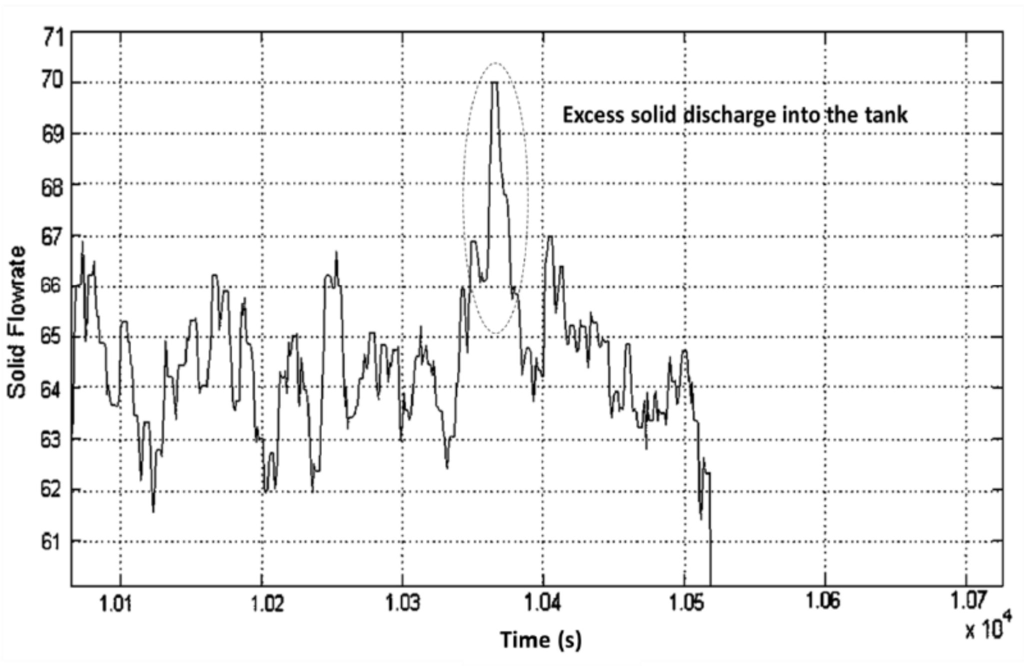

Figure 3.9. Solid flowrate to dissolution tank

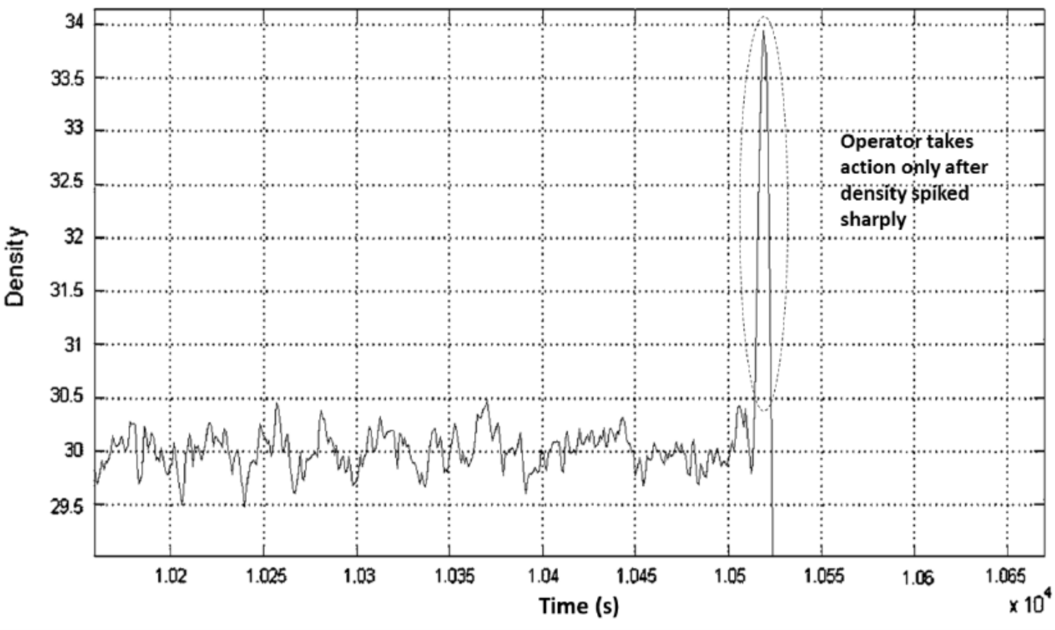

Figure 3.10. Density of liquid in dissolution tank

We applied risk-based methodology to the system for early detection of the fault. In this case we monitor the following process variables: speed of solid rotary feeder (rpm), inlet water flowrate (w_f), inlet solid flowrate (s_f) and density of outlet (d). These variables are collected in data matrix, X.

$$X = \begin{bmatrix} rpm \\ w_f \\ s_f \\ d \end{bmatrix} \tag{3.20}$$

In order to take into account the dynamic nature of the data the variables were adjusted for the time delay. Therefore, the data matrix, X contains the time delay adjusted variables. PCA has been applied to the data matrix. Based on Table 3.2 the first three PCs which collectively capture 80% of the total variance are retained.

Table 3.2. Principal component analysis for the second case study

PC Number	% Variance captured this PC	% Variance captured total
1	31.35	31.35
2	25.16	56.51
3	23.73	80.23
4	19.77	100

Subsequently the loadings calculated from the training data set were used to project the data onto a PC subspace. It was observed that PC2 was able to detect the fault at the earliest. The scores of PC2 were used to calculate the probability of fault.

Similar to the previous case study we assign severity to different variables using a relative scaling. Severity of water flowrate was taken as a reference point. It is considered that the severity of hazard associated with the speed of solid rotary feeder and solid flowrate are twice the severity of hazard associated with the water flowrate. Finally, any fault in density has a direct impact on the downstream process equipment. Therefore, the severity of hazard associated with the density is assigned three times the severity of hazard associated with the water flowrate.

$$a = \begin{bmatrix} a_{rpm} \\ a_{Wf} \\ a_{Sf} \\ a_d \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 3 \end{bmatrix} \quad (3.21)$$

$$w = \begin{bmatrix} w_{rpm} \\ w_{Wf} \\ w_{Sf} \\ w_d \end{bmatrix} = \begin{bmatrix} 0.007 \\ 0.072 \\ 0.676 \\ 0.245 \end{bmatrix} \quad (3.22)$$

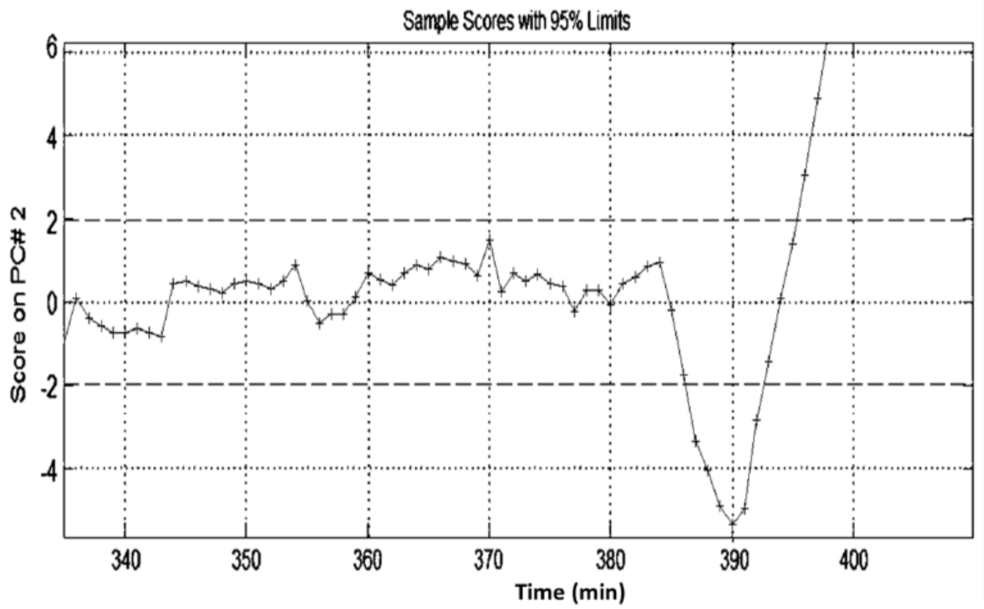

Figure 3.11. Score of the second PC for the dissolution tank data

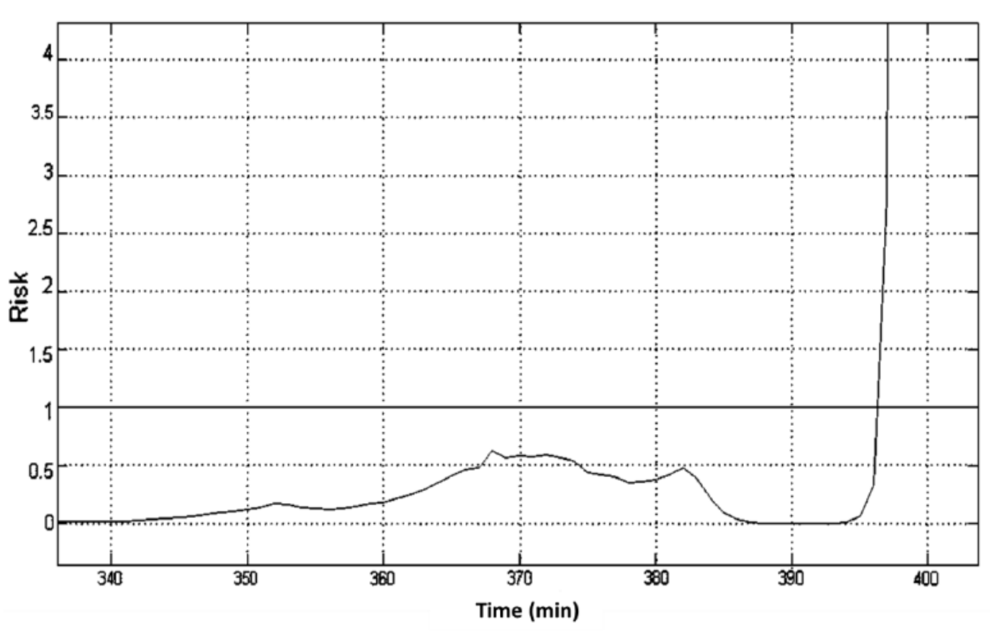

Figure 3.12. Risk profile of the second PC of the dissolution tank data

Risk profile of the fault is obtained by combining the effect of severity with the probability of the fault. Figure 3.12 shows the risk profile of the second PC. The first risk threshold at 1 is used for fault detection in parallel with the warning to the operators for priority response. The second threshold placed at 10 will activate safety systems (emergency shutdown systems).

Application of the multivariate risk-based fault detection provides early warnings and early activation of safety systems prior to the fault impacting the system in comparison with univariate methods. Table 3.3 shows time of fault detection using different signals. In this case, multivariate risk-based approach detects the fault 36 minute earlier than a univariate approach.

Table 3.3. Time of fault detection using different signals

	Monitored Variable	Time (min.)	Improvement
Actual time of Fault	Solid flowrate	10365	n/a
Time of Operator Intervention	Density	10519	Base line
	Scores of 2nd PC	10486	33 min earlier
	Risk of 2nd PC	10483	36 min earlier

Figure 3.13 and 3.14 show the contribution and risk associated with each process variables. This can be used for root cause analysis.

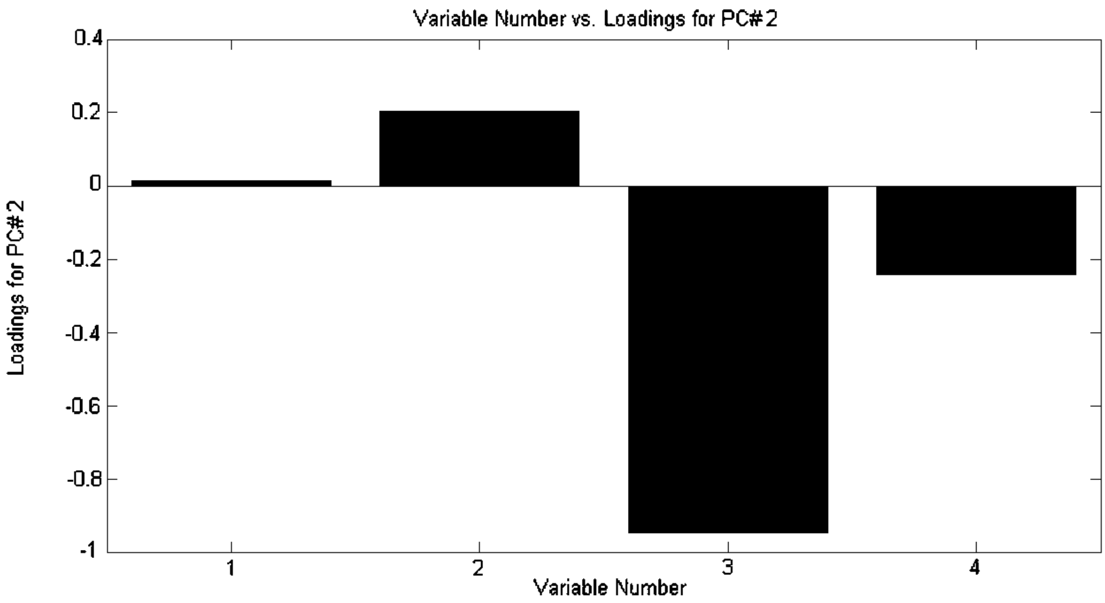

Figure 3.13. Contribution of rotational speed of rotary feeder (variable #1); inlet water flowrate, (variable #2); inlet solid flowrate (variable #3); and density of outlet (variable #4) to the detected fault in the dissolution tank data

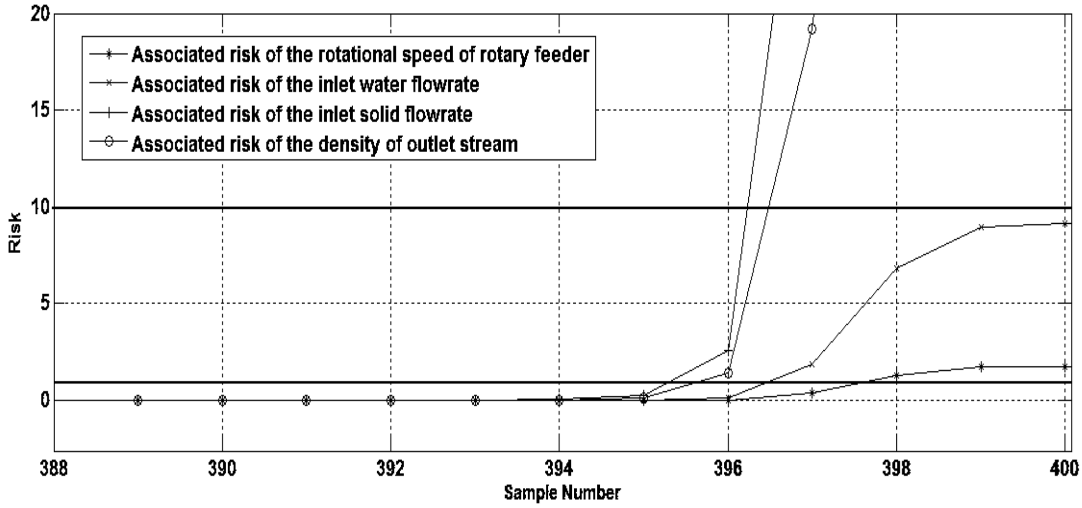

Figure 3.14. The associated risk of each original variable, R_k in the dissolution tank data

In this case because of the actuator problem in the rotary drum excess solid crystals dropped into the tank. As expected, the contribution of the solid flowrate, s_f, as well as its risk is more than other variables.

3.4 Conclusions

PCA based fault detection and diagnosis technique has been extended to a risk-based fault detection and diagnosis framework targeting the safety issues of a process system. In this method a warning system is activated when the risk of operation exceeds the acceptable threshold. Combining PCA with the risk calculation procedure makes this method robust to false alarms. This method has more power in discerning between operational changes and abnormal conditions which can cause accidents. Severity of the fault associated with different process variable is considered to calculate the operational risk. The proposed technique is demonstrated on a simulated system and using real-life industrial data. This technique provides much early warning compared to the univariate methods. Further, due to risk-based approach, warning and recovery options are easy to prioritize.

Acknowledgment

Authors thankfully acknowledge the financial support provided by Vale Inc., Research & Development Corporation (RDC) and Atlantic Canada Opportunities Agency (ACOA) under the Vale Research Chair Program.

References

[1] Himmelblau D. M., Fault-detection and diagnosis in chemical and petrochemical processes; Elsevier: Amsterdam, **1978.**

[2] Mylaraswamy D., DKit: A Blackboard-Based Distributed Multi Expert Environment for Abnormal Situation Management; PhD Thesis, Purdue University, **1996.**

[3] Dash S., Venkatasubramanian, V., Challenges in the Industrial Applications of Fault Diagnostic Systems, Computers & Chemical Engineering, **2003**, 24, 785-791.

[4] Rengaswamy R., Venkatasubramanian V., A Syntactic Pattern-Recognition Approach for Process Monitoring and Fault Diagnosis; Engineering Applications of Artificial Intelligence, **1995**, 8, 1, 35-51.

[5] Kresta J., MacGregor J. F., Marlin T. E., Multivariate Statistical Monitoring of Process Operating Performance, The Canadian Journal of Chemical Engineering, **1991**, 69, 35-47.

[6] Bakshi B. R., Multiscale PCA with Application to Multivariate Statistical Process Monitoring, AIChE Journal, **1998**, 44, 7, 1596–1610.

[7] Liebmann M. J., Edgar T. F., Lasdon L. S., Efficient Data Reconciliation and Estimation for Dynamic Processes Using Nonlinear Programming Techniques, Computers & Chemical Engineering, **1992**, 16, 963.

[8] Gertler J., Fault-detection and diagnosis in engineering systems; Marcel Dekker: New York, 1998.

[9] Chowdhury F., Ordinary and Neural Chi-Squared Tests for Fault Detection in Multioutput Stochastic Systems, Control Systems Technology, IEEE Transactions on, **2000.** 8, 2, 372-379.

[10] Basseville M., Nikiforov I. V., Detection of abrupt changes: theory and application Prentice-Hall, Inc. Upper Saddle River: NJ, USA, **1993**.

[11] Kiasi F., Fault Detection and Isolation of a Benchmark Wind Turbine using the Likelihood Ratio Test, Proc. IFAC World Congress, 18^{th}, Milan, Italy, **2011**.

[12] Willsky A. S., Jones, H. L., A generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes, Decision and Control, IEEE Conference on, Phoenix, Arizona, **1974.**

[13] Bao H., Khan F., Iqbal T., Chang Y., Risk-Based Fault Diagnosis and Safety Management for Process Systems, Process Safety Progress, **2011**, 30, 1, 6–17.

[14] Zadakbar O., Imtiaz S. A., Khan F., Dynamic Risk Assessment and Fault Detection Using Multivariate Technique, AIChE Journal, submitted, **2011.**

[15] Qin S. J., Statistical Process Monitoring: Basics and Beyond, Journal of Chemometrics, **2003**, 17, 8-9, 480–502.

[16] Li W. H., Qin S. J., Consistent Dynamic PCA Based on Errors-in-Variables Subspace Identification, Journal of Process Control, **2001**, 11, 6, 661–678.

[17] Smith L., A Tutorial on Principal Components Analysis.
www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf, 2002

[18] Jackson J. E., A User's Guide to Principal Components, John Wiley & Sons, Inc. **1991.**

[19] Skogestad S., Dynamics and Control of Distillation Columns - A Tutorial Introduction, Trans. IChemE, **1997**, 75A, 539-562.

Chapter 4: Dynamic Risk Assessment and Fault Detection Using a Multivariate Technique†

Preface

A version of this chapter has been published by the *Journal of Process Safety Progress*, in 2013. Zadakbar was the main lead on the work. The co-authors, Dr. Khan and Dr. Imtiaz supervised the work and helped to develop the methodology.

Abstract

In the context of process safety, significant improvements are needed in fault detection methods, especially in the areas of early detection and warning. In this Chapter, a multivariate risk-based fault detection and diagnosis technique is proposed. The key elements of this technique are to eliminate faults that are not serious and to provide a dynamic process risk indication at each sampling instant. A multivariable residual generation process based on the Kalman filter has been combined with a risk assessment procedure. The use of the Kalman filter makes the method more robust to false alarms, which is an important aspect of any fault detection algorithm that targets the safety of a process. In addition, we consider significant differences in the severity of the faults associated with different process variables. We also take into account the varying

† *O. Zadakbar, S. Imtiaz, F. Khan, Dynamic risk assessment and fault detection using a multivariate technique, Process Safety Progress, Volume 32, Issue 4, pages 365–375, December 2013, DOI: 10.1002/prs.11609, Article first published online: 17 JUN 2013, Copyright © 2013 American Institute of Chemical Engineers*

intensity of damage caused by the increasing and decreasing rates of fault and the need to treat those cases differently.

Keywords: Dynamic risk assessment, multivariate fault diagnosis, risk assessment of process system, Kalman filter

4.1 Introduction

Abnormal Event Management (AEM) involves detecting an abnormal event at an appropriate time, diagnosing its cause, and then taking appropriate supervisory control decisions and actions to bring the process back to a normal, safe operating state [1]. Fault detection and diagnosis are key components of abnormal event management. Process fault detection, a critical part of process engineering, is essential to product quality and operational safety.

The term fault has many connotations in the literature. Typically, any departure of process variables from an acceptable range is considered as a fault [2]. Historically, univariate methods, such as limit- or trend-checking of measurable output variables, have been used for fault detection [3-6]. However, the applicability of univariate methods is limited, as they are unable to distinguish between noise and abnormal, faulty conditions; thus, they can provoke false alarms.

More success in fault detection has been observed when multivariable fault detection techniques are applied [4, 5, 7], as these methods take advantage of the dependence among the process variables (i.e., between input and output), and flag a fault when any deviation from the correlation structure is detected. These methods are better able to distinguish between operational changes and abnormal conditions [8]. Multivariable methods are model-based methods, where models are used either explicitly or implicitly to generate residuals: the prediction is subtracted from the observed values. Both data-based models and first- principles models have been used for prediction in connection with fault detection and diagnosis [9].

To make the detection robust, Kalman predictors have been widely used to predict the fault-free variables. Process industries are complex systems, and there is always uncertainty in process models, but the Kalman filter makes the prediction robust as it also optimally updates the prediction with measurements [10]. In general, fault detection techniques seek to detect operational faults that affect the control objectives of the process. However, in the context of process safety, these methods are inadequate, as none of them take into account the potential impact of the fault on the process and the environment. Bao *et al*. (2011) proposed a risk-based fault detection method in order to address this issue. Instead of generating an alarm based on residuals or signals crossing the threshold, the risk-based fault detection method issues an alarm only when the risk of a process exceeds the acceptable threshold [6, 8]. The risk of a process is defined as a combination of probability of fault and severity of fault. This is a crucial concept, as it eliminates faults that are not operational or process safety concerns, and also gives a dynamic indication of the operational risk [8]. Bao *et al.* (2011) used a univariate charting method to calculate the probability of a fault, with limited effectiveness of the method because of inherent limitations of the univariate fault detection and diagnosis approach.

In this chapter, we propose a multivariate risk-based fault detection and diagnosis technique. A multivariable residual generation process based on the Kalman filter has been combined with the risk calculation procedure earlier proposed by Bao *et al.* (2011). The proposed method takes advantage of the correlation between process input and output and is therefore better equipped to detect faults and calculate risk precisely. Also, the use of the Kalman filter makes the method more robust to false alarms; this represents a significant advantage in any fault detection algorithm targeting the safety issues of an operation. The main benefits of this approach are improved safety, minimum interruption of operation, better alarm management or early warning system and higher availability of process.

4.2 Problem Formulation

Consider a linear time invariant system in discrete time,

$$x_k = A_k x_{k-1} + B_k u_{k-1} + W_k w_k \qquad y_k = C_k x_k + V_k v_k \tag{4.1}$$

where $k \in \mathbb{N}$ is the time index and $u_k \in \mathbb{R}^l$ is noise-free input; and $y_k \in \mathbb{R}^m$ is noisy output. $x_k \in \mathbb{R}^n$ is the state; $w_k \in \mathbb{R}^n$ is a stationary Gaussian white noise vector with covariance $Q \in \mathbb{R}^{n\times n}$, e.g., $w_k \sim \aleph(0, Q)$, to represent process disturbances; $v_k \in \mathbb{R}^m$ is a stationary Gaussian white noise vector with covariance $R \in \mathbb{R}^{m\times m}$, e.g., $v_k \sim \aleph(0, R)$. It represents measurement disturbances. $A_k, B_k, C_k, V_k, and\ W_k$ are system matrices with compatible dimensions. The above set of equations can be modified to represent faults arising at different parts of the system. In the case of a process fault, system equations can be represented by

$$x_k = A_k x_{k-1} + B_k u_{k-1} + F_k f_{p_k} + W_k w_k \quad y_k = C_k x_k + V_k v_k \tag{4.2}$$

$f_{p_k} \in \mathbb{R}^n$ is the process fault magnitude vector with zero or non-zero elements.

4.3 Fault Detection and Diagnosis

The fault detection and diagnosis scheme is based on a test of the residuals. The residual of a particular measurement at any instant is given by

$$r_k = \tilde{y}_k = y_k - \hat{y}_{k|k-1} \tag{4.3}$$

where y_k is the measured vector at a particular instant, and $\hat{y}$ contains the corresponding predicted values. Among various methods for residual generation, the Kalman-filter-based residual generation method was employed to calculate the residues [11]. The Kalman filter is an optimal state estimator for a linear system in the presence of Gaussian noise. Unlike model predictions such as those in equation (4.3), Kalman filter predictions were subtracted from the measurements to calculate the residues. The Kalman estimate is robust in the presence of disturbance and uncertainty in the process. Kalman estimations are optimally weighted between the model prediction and the observed value, and they account for measurement noise and disturbances in the process model. This is particularly important in the context of risk-based fault detection (RFDI), since it is aimed at predicting only those faults that constitute operational safety concerns.

Given $A, B, C,$ $V,$ and W and some knowledge about process noise variance Q and disturbance noise variance R, the Kalman filter gives the least squares estimate of x_k

such that $x_k - \hat{x}_{k|j}$ has minimum variance. $\hat{x}_{k|j}$ is the estimation of x_k based on $\{u(1), y(1), \ldots, u(j), y(j)\}$, where $j = k-1$ or $j = k$. In fact, the Kalman filter is capable of predicting one step ahead if $j = k-1$, or filtering if $j = k$ [12]. The following form is considered for the one-step predictor algorithm:

$$\hat{x}_{k+1|k} = A\,\hat{x}_{k|k-1} + Bu_k + K_k\tilde{y}_k \qquad \hat{y}_k = C\hat{x}_{k|k-1} \tag{4.4}$$

Where K_k is the Kalman gain that is computed as below:

$$K_k = P_k^- C^T (C P_k^- C^T + R)^{-1} \qquad P_k^- = E[e_k^- e_k^{-T}] \qquad e_k^- = x_k - \hat{x}_k \tag{4.5}$$

In this work, additive process fault as given in equation (4.2) is considered. In the presence of process fault, the residuals will have the following form:

$$r_k = [C_k(A_k x_{k-1} + B_k u_{k-1} + F_k f_{p_k} + W_k w_k) + V_k v_k]$$

$$- [C_k(A\,\hat{x}_{k|k-1} + Bu_k + K_k\tilde{y}_k)]$$

$$r_k \approx [(C_k F_k f_{p_k} + C_k W_k w_k + C_k V_k v_k) - C_k K_k \tilde{y}_k] \tag{4.6}$$

As is shown in equation (4.6), since the Kalman-predicted values are used for residual calculation, there is an additional term $(-C_k K_k \tilde{y}_k)$ for residual equation. Subtracting a portion of the residuals makes the fault prediction conservative, which makes the method more robust to false alarms. This is an important component of the fault detection algorithm, particularly in this case, when the focus is on safety in the process operation.

In order to give more flexibility to the prediction process, a further filter was applied to the calculated residual signal from Equation (4.6). The filter coefficient of the moving average filter α is a tuning parameter that is further tuned to minimize false alarms.

Finally, the slope of the predicted signal was used to predict the fault at any given instant:

$$r_k = \alpha r_k + (1 - \alpha) r_{k-1} \quad (4.7)$$

4.4 Dynamic Risk Calculation

Risk is defined as a measure of likely harm or economic loss caused by the fault if corrective action is not taken. Risk depends on two factors: the probability of occurrence of a fault leading to an unwanted event and the severity of loss caused by the event [13]. Bao *et al.* (2011) proposed the following formula for a univariate deterministic system:

$$Risk = p \times s \quad (4.8)$$

where p is probability and s is severity of a fault. The probability and severity of the fault for each predicted residual point are calculated using equations (4.9) and (4.10). The probability of an event increases as the process moves further away from the normal operation. In this methodology, $\mu - 3\sigma$ and $\mu + 3\sigma$ are used as the lower and upper thresholds for normal operation. All data points between these two thresholds are considered normal. If any predicted point goes outside this region, it could lead to a fault and, possibly, to an event. When the residuals are at the threshold ($\mu \pm 3\sigma$), the probability of fault is 0.5, as it can either go back to normal or keep growing and ultimately lead to an abnormal event. Based on this intuition, Bao *et al.* (2011) developed a cumulative normal distribution for fault probability, which we also use in the present study. For positive fault signals when the residuals approach the upper threshold, the probability of the fault is calculated by $\varphi\left(\frac{x-(\mu+3\sigma)}{\sigma}\right)$. On the other hand, for negative fault signals, the residuals approach the lower threshold and the probability of the fault is calculated by the complement, $1-\varphi\left(\frac{x-(\mu-3\sigma)}{\sigma}\right)$,

$$for\ r > \mu \longrightarrow\ p = \varphi\left(\frac{r-(\mu+3\sigma)}{\sigma}\right) = \int_{-\infty}^{r} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(r-(\mu+3\sigma))^2}{2\sigma^2}} dr \tag{4.9}$$

$$for\ r < \mu \longrightarrow\ p = 1-\varphi\left(\frac{r-(\mu-3\sigma)}{\sigma}\right) = 1-\int_{-\infty}^{r} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(r-(\mu-3\sigma))^2}{2\sigma^2}} dr \tag{4.10}$$

Figure 4.1 shows a visual depiction of a fault probability. The probability of the fault for a point on the centerline μ is 0 and for a given point on the thresholds is equal to 0.5.

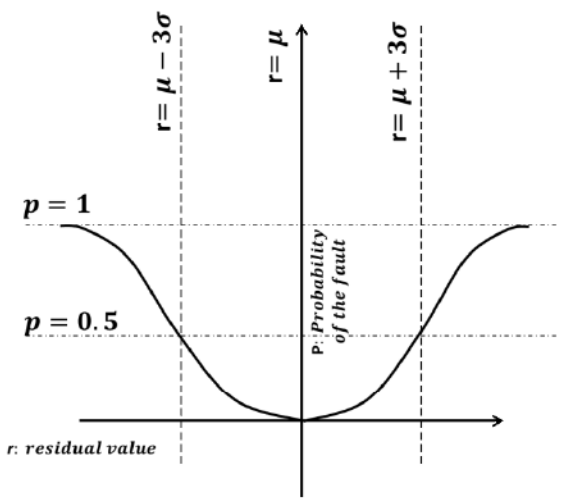

Figure 4.1. Probability of fault (φ)

The severity of fault is calculated using equation (4.11-13). It is a modified version of the original equation proposed by Bao *et al.* (2011), which needed modification because it was developed for univariate methods and did not take into account the various degrees of severity of different types of fault. The modified equation takes into consideration the severity of the fault associated with different process variables. It also accounts for the varying degrees of severity involved in the increasing and decreasing rate of fault. Therefore, these cases should be treated differently. The modified severity equation is given below:

$$for\, r > \mu \dashrightarrow \; s = (a.100)^{\frac{(r-(\mu+3\sigma))}{r-\mu}} \tag{4.11}$$

$$for\, r < \mu \dashrightarrow \; s = (\acute{a}.100)^{\frac{((\mu-3\sigma)-r)}{\mu-r}} \tag{4.12}$$

$$\acute{a} = a.b \tag{4.13}$$

Coefficient a [8] in the above equations is called the *intensity coefficient;* it indicates the intensity of the severity of the fault associated with each process variable. For instance, in a simple chemical reactor containing nonhazardous chemical compounds, the severity of damage caused by uncontrolled temperature is much greater than an uncontrolled concentration of a given component. Thus, a associated with temperature is larger than a associated with concentration, and this can be simply depicted $a_{Temperature} > a_{Concentration}$.

Coefficient b [8] in the above equations is called the *moderation coefficient*. Since the severity of the fault in the case of a decreasing rate may not be equal to the severity of the fault in an increasing rate, coefficient b is used to consider this effect and moderate the severity of the outcome. For example, an abnormally increasing temperature in a given process can have much more damaging effects than an unusually decreasing temperature. On the other hand, a decreasing cooling water flow in a reactor can cause severe damage and needs more immediate attention than an increasing cooling water flow. Coefficient b provides the flexibility to treat increasing and decreasing faults differently. Both coefficients a and b are selected based on process and operational considerations in a given system, such as the nature of a process, number of people at risk, chemical and physical components, environment, and costs.

4.4.1 Risk-based Fault Detection

The methodology for risk-based fault detection is given in Figure 4.2. The first step of the methodology is to model the process. Then faulty conditions are generated in a given process model. Process states may not be measured. Therefore, in the next step the Kalman filter estimates all process states and is also used for residual generation of measured process states. After that, filtered residuals are used to predict the next residual points based on the slope of the three previous successive real-time data points in time series. In the next step, the risk of each predicted point is calculated using probability and the severity of the fault. Finally, the risk profile is used for fault detection, and necessary supervisory decisions are taken to implement the appropriate safety systems.

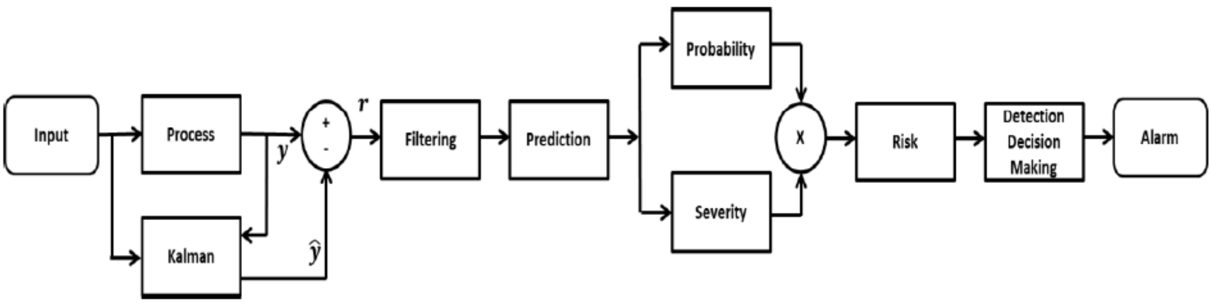

Figure 4.2. Methodology of risk-based fault detection

The criteria for selecting risk thresholds depend on the process system. The risk threshold is defined based on the acceptable risk of each process system. The risk profile is used not only for fault detection but also for making decisions about the activation of a series of safety systems corresponding to the level of the risk, in parallel with a warning for the operators to initiate a priority response if the automatic systems should fail.

4.5 Case Studies

In this section, risk-based fault detection methodology is applied to two simulated case studies. The first case study is a Continuous Stirred Tank Reactor (CSTR) and the second is a 40-stage distillation column. In the CSTR, a fault is introduced in the temperature; in the distillation column, a fault is introduced in the reboiler.

4.5.1 Continuous Stirred Tank Reactor (CSTR)

A simple schematic of the Continuous Stirred Tank Reactor (CSTR) is shown in Figure 4.3. In addition to complete mixing assumption; density of the reactant, the specific heat of the cooling and the heat of the reaction are considered independent of the temperature and there is assumed to be no phase change in the reaction mass.

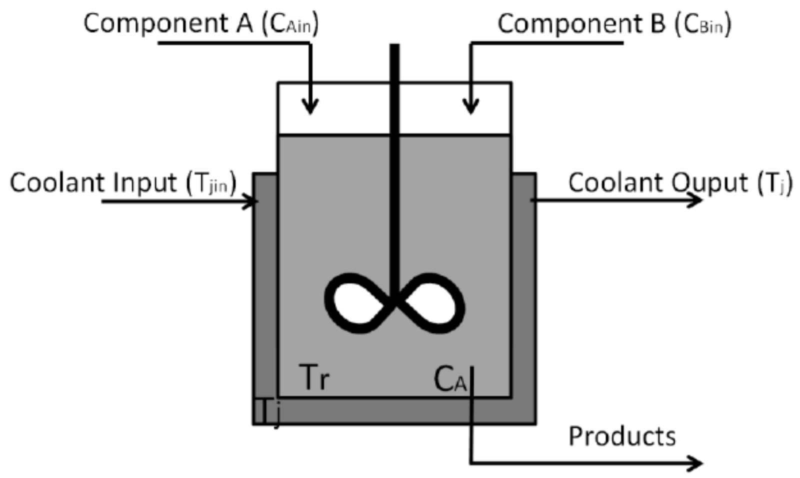

Figure 4.3. Continuous Stirred Tank Reactor

The reaction in the system is the oxidation of sodium thiosulfate by hydrogen peroxide, which is an irreversible exothermic oxido-reduction reaction. This reaction is expressed by the following equation [11]:

$$2Na_2S_2O_3 + 4H_2O_2 \dashrightarrow Na_2S_3O_6 + Na_2SO_4 + 4H_2O$$

$$2A + 4B \dashrightarrow C + D + 4F$$

$$r_A = 4k(T_r)C_AC_B = -(k_0 + \Delta k_0)\exp\left(\frac{-E+\Delta E}{R_gT_r}\right)C_AC_B \tag{4.14}$$

where k_0 is the pre-exponential factor, and C_A and C_B are the concentrations of components A and B respectively. E is activation energy; R_g is the gas constant; Δk_0 and ΔE represent uncertainty and T_r is the reactor temperature.

The system equations describe the reactant concentration, C_A, reactor temperature, T_r, and jacket temperature, T_j [11]:

$$\frac{dC_A}{dt} = \frac{F}{V}(C_{Ain} - C_A) - 4kC_A^2 \tag{4.15}$$

$$\frac{dT_r}{dt} = \frac{F_r}{V_r}(T_{rin} - T_r) + 4\frac{(-\Delta H_r)+\Delta(-\Delta H_r)}{\rho C_p}kC_A^2 - \frac{UA+\Delta UA}{\rho C_p V}(T_r - T_j) \tag{4.16}$$

$$\frac{dT_j}{dt} = \frac{F_w}{V_w}\left(T_{jin} - T_j\right) + \frac{UA+\Delta UA}{\rho_w C_{p_w} V_w}(T_r - T_j) \tag{4.17}$$

To implement the Kalman filter, the nonlinear CSTR model is linearized around its nominal operating conditions. The discrete state space model coefficients are given as [11]:

$$A_k = \begin{bmatrix} -0.08 & -5.89e-4 & -1.26e-5 \\ 156.23 & 1.04 & 0.02 \\ 33.55 & 0.26 & 0.72 \end{bmatrix} \quad B_k = \begin{bmatrix} -1.69e-7 \\ 4.23e-4 \\ 0.02 \end{bmatrix} \quad C_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(4.18)

The concentration of component A, C_A, the temperature in the reactor, T_r, and the temperature in the coolant jacket, T_j, are selected as the elements of process state matrix. It is assumed that there are two sensors to measure concentration of component A, C_A, and the temperature in the coolant jacket, T_j:

$$x(t) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \Delta C_A \\ \Delta T_r \\ \Delta T_j \end{bmatrix} \qquad y(t) = \begin{bmatrix} \Delta C_A \\ \Delta T_j \end{bmatrix} \tag{4.19}$$

The process control variables are feed to the reactor, F, and cooling water flow rate, F_w. Both variables could be manipulated. The feed concentration and cooling water inlet temperature T_{jin} are uncontrolled inputs in the process; thus, they may be considered as disturbances. It is assumed that the cooling water inlet temperature changes as a pulse function. Figure 4.4(a) shows the residuals of jacket temperature T_j under normal conditions, generated by subtracting Kalman estimates from the measurements. The thresholds are placed at $\pm 3\sigma$, which cover 99.7% of the values.

The noisy residual signal is subsequently filtered and then, based on the slope of three previous successive real-time data points in time series, the next five data points of the residual are predicted. Figure 4.4 shows the T_j residuals and the filtered and predicted residuals in normal conditions with $\pm 3\sigma$ limits.

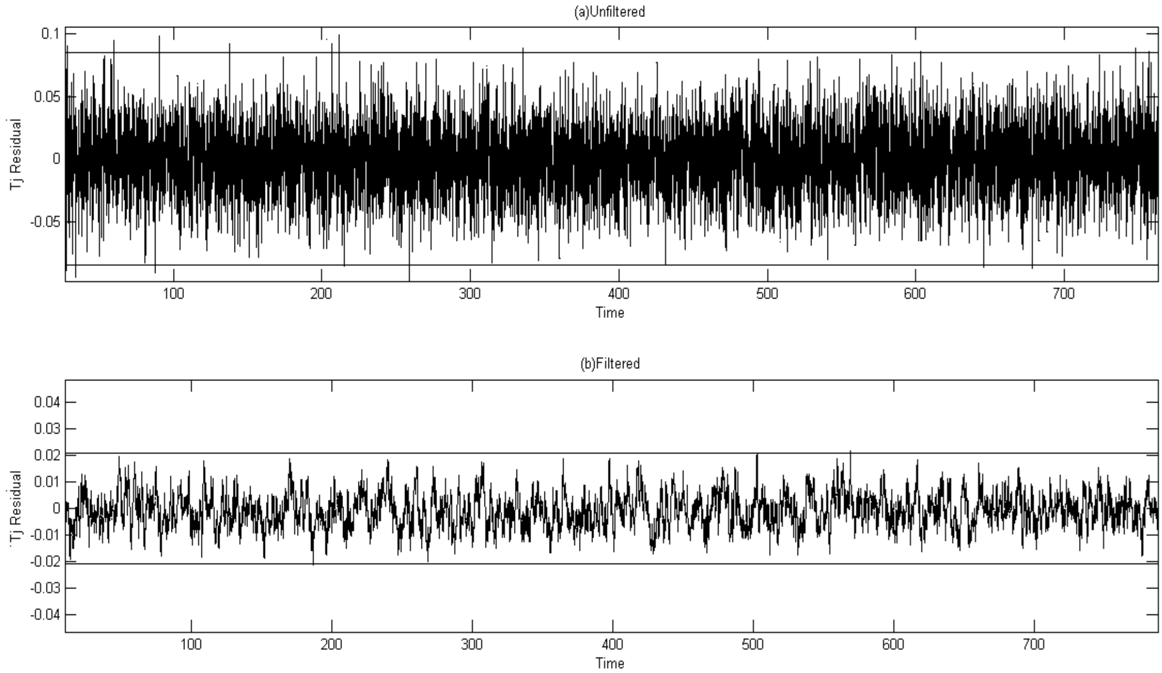

Figure 4.4. T_j residuals and the filtered/predicted residuals in normal conditions (a) unfiltered residuals (b) filtered residuals

A process fault was simulated by adding a ramp-type fault to the reactor temperature T_r. Figure 4.5 shows T_r in normal and faulty conditions where an additive fault occurred at t=400s.

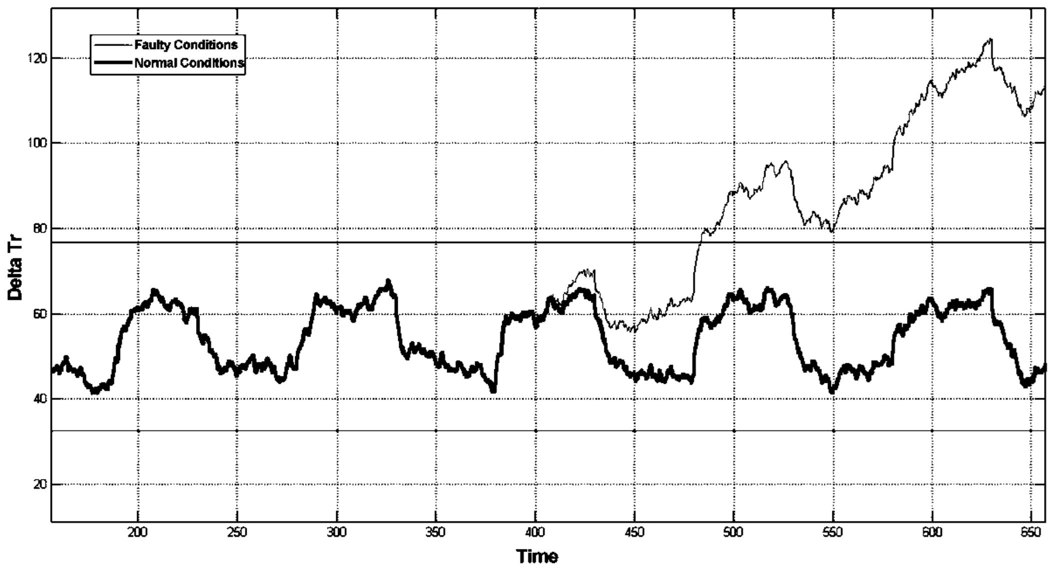

Figure 4.5. ΔT_r in normal (down) and faulty conditions (up)

The reactor temperature T_r is not directly measured, whereas T_j is directly measured. The T_j residuals are directly affected by the T_r fault. Figure 4.6 shows the T_j residuals in faulty conditions.

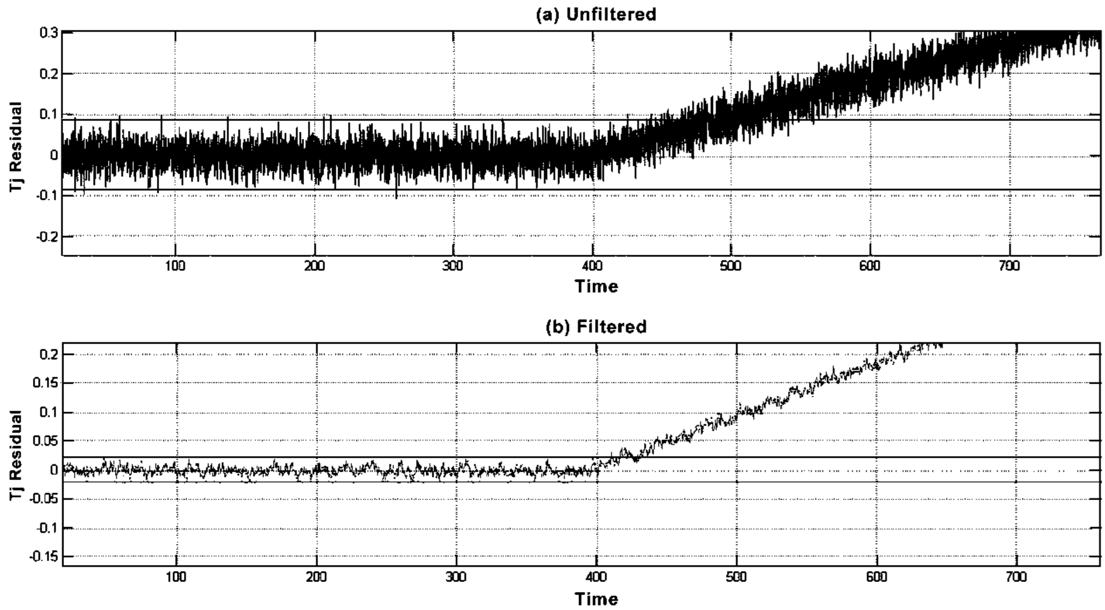

Figure 4.6. T_j residuals in faulty conditions (a) unfiltered residuals (b) filtered residuals

4.5.1 Dynamic Risk Calculation

The probability of the fault for the T_j residuals is illustrated in Figure 4.7. In normal operating conditions, the probability of the fault fluctuates around zero. After t=400s, when the fault has occurred, the probability of the fault starts increasing. When predicted residuals cross the threshold, the probability of the fault is 0.5.

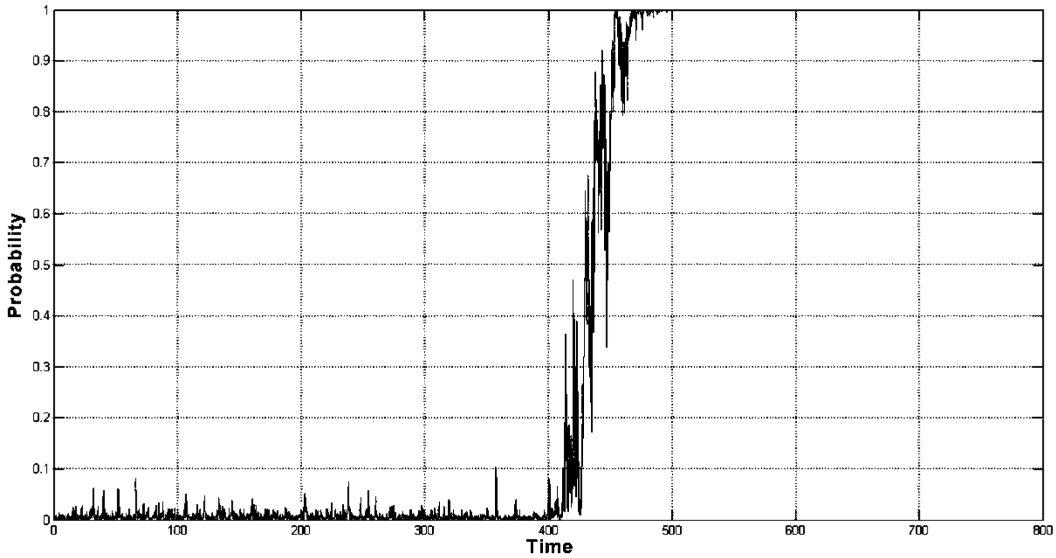

Figure 4.7. The probability of the fault for the predicted T_j residuals

Figure 4.8 shows the severity of the fault - increasing jacket temperature and sodium thiosulfate concentration. In this case study, the reactor contains sodium thiosulfate (with health rating 1–slight; flammability rating 0–none; reactivity rating 1–slight; contact rating 1–slight) [15]. Thus, if personnel use protective equipment such as safety glasses, lab coats, and proper gloves, the severity of the increase in temperature is much higher than that of the increase in concentration of sodium thiosulfate. Therefore, a associated with temperature is larger than a associated with concentration. In this case study, $a_{Tj} = 4$ and $a_{CA} = 1$.

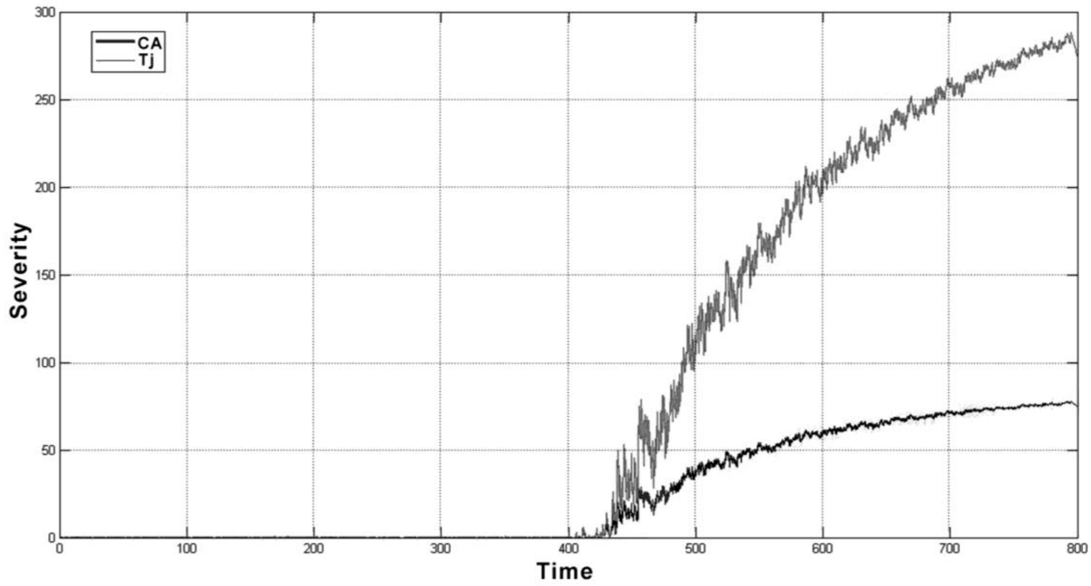

Figure 4.8. Severity of increasing jacket temperature (up) and sodium thiosulfate concentration (down)

The risk of the associated fault is obtained by combining the severity with the probability. Figure 4.9 shows the risk profile for the T_j residuals as well as the risk profile for the C_A.

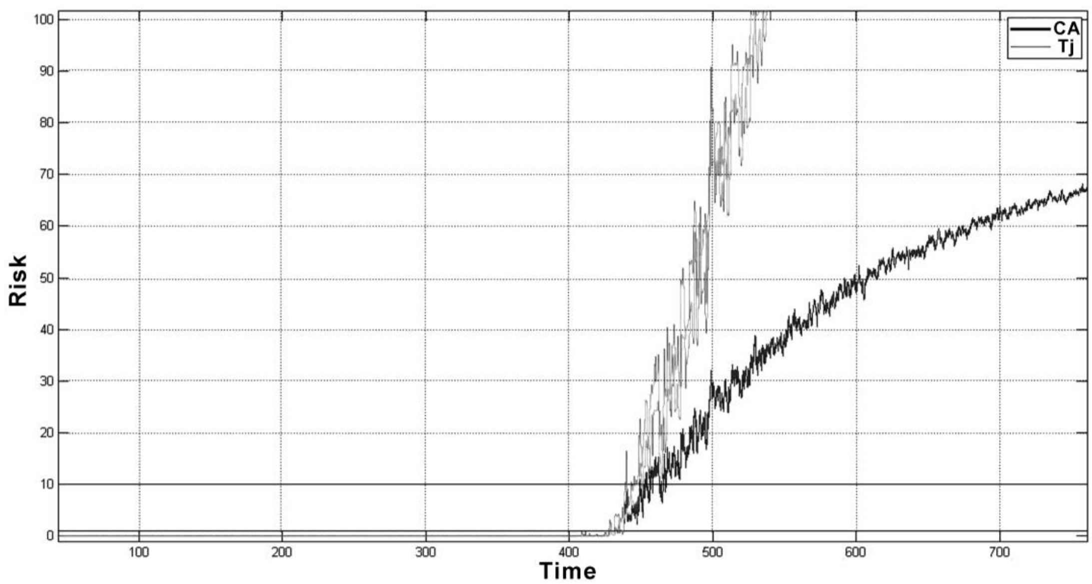

Figure 4.9. Risk profile for the predicted T_j (up) and C_A residuals (down)

The threshold for the risk signal is based on the acceptable risk criteria of the process system; for example, in a nuclear power plant the threshold will be very low, whereas in a water purification plant the acceptable risk may be relatively high. Figure 4.9 shows a few guiding principles of acceptable risk in process operation. The first risk threshold is at 1. It is used for fault detection in parallel with the warning for operators to initiate a priority response if the automatic systems failed [6, 8]. If no corrective action is taken, or if the corrective action fails to bring down the risk, and if the risk exceeds the second threshold, then the automatic safety system is activated (emergency shutdown). The application of the intensity coefficient improves the methodology by incorporating the impact of the fault on product quality, economics, and potential accidents.

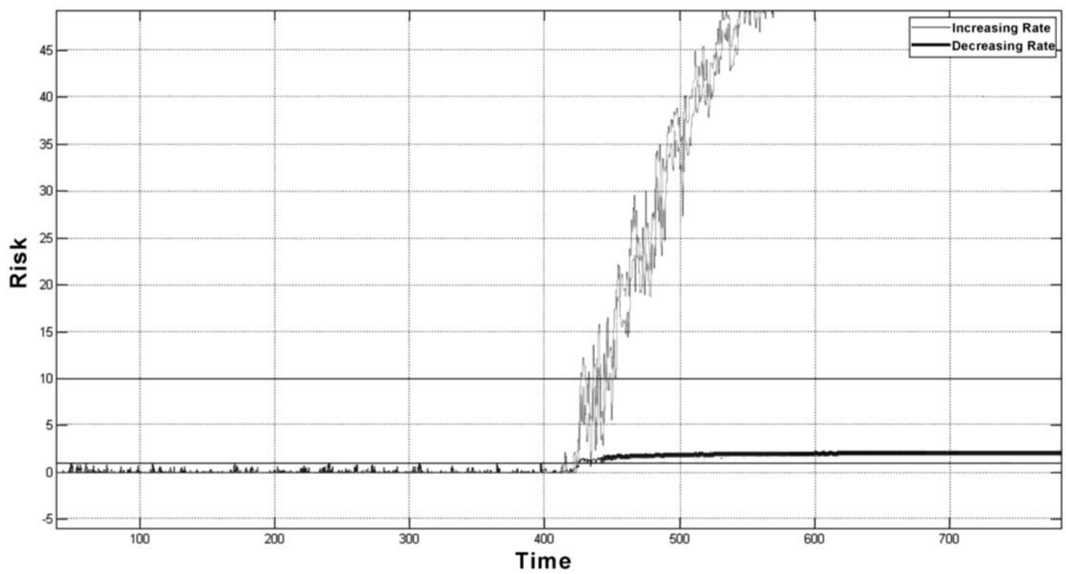

Figure 4.10. Risk of predicted T_j residuals in both increasing (up) and decreasing (down) cases

The effect of moderation coefficient b is illustrated in Figure 4.10. In case of decreasing T_j residuals, moderation coefficient b=0.02 moderates the severity of the fault. The risk associated with decreasing T_j is very low compared to the risk associated with increasing T_j.

Figures 4.5, 4.6, and 4.9 show how the application of risk-based fault detection provides early warnings prior to the fault affecting the system. The activation of an alarm is based on a risk of operation that exceeds the acceptable threshold (t=420s), rather than on residuals (t=430s) or signals crossing the threshold (t=470s).

4.5.2 Binary Distillation Unit

The second case study is a binary distillation column with 40 stages that separates a binary mixture with relative volatility of 1.5 into products of 96% purity (Figure 4.11) [14]. The assumptions considered to model the distillation unit are: binary mixture, equilibrium at all stages, constant pressure and relative volatility, total condenser, no vapor holdup, and linearized liquid dynamics.

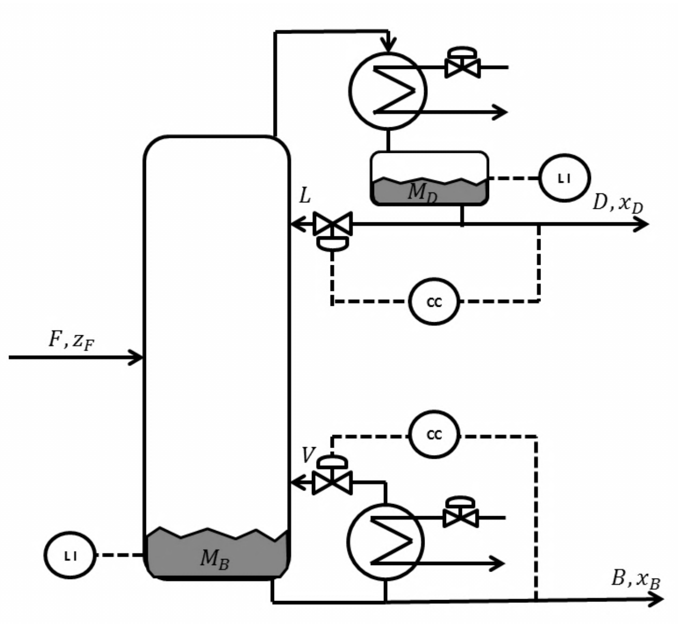

Figure 4.11. Binary distillation column

A nonlinear Simulink model is used to replicate the distillation unit [16]. The dynamic nonlinear model is linearized and subsequently discretized to a linear system with 82 states, 6 inputs, and 4 output variables. The first 41 states are compositions of light component with the reboiler as $x(1)$ and the condenser as $x(41)$. State $x(42)$ is holdup in the reboiler and $x(82)$ is holdup in the condenser. Inputs are reflux flow rate L, boilup flow rate V, top or distillate product flow D, bottom product flow B, feed rate F, and feed composition z_F. There are four sensors to measure top composition x_D, bottom composition x_B, condenser holdup M_D and reboiler holdup M_B. Many different kinds of fault have been considered in the distillation column [17]:

1. Process loads: feed flow rate, feed composition, top product flow rate, bottom product flow rate

2. Changes in heating of the reboiler and cooling of the condenser

3. Equipment fouling

In this study, a sudden increase in reboiler heat flow is considered as a fault. This fault may cause increasing in vapor flow rate with the time which would affect other process states, e.g., top and bottom concentrations. It is assumed that the vapor flow of the reboiler starts to increase at T=2000 min. The control inputs are all set to zero except the reflux flow rate, which is a pulse function. Figure 4.12 shows the fluctuation of top concentration in normal and faulty conditions.

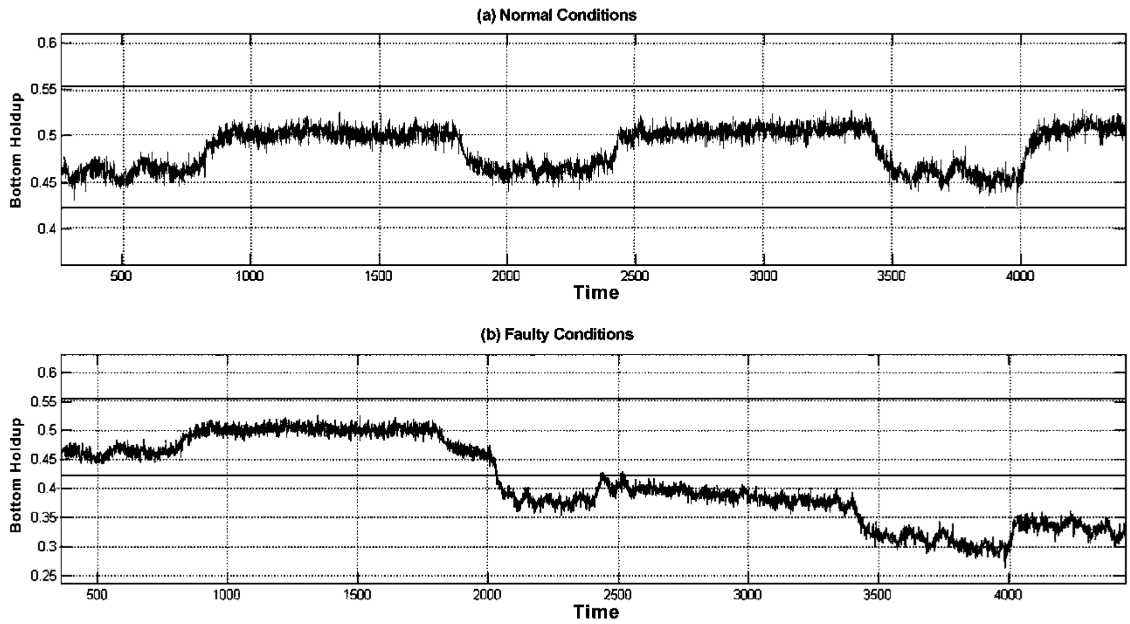

Figure 4.12. Bottom holdup fluctuation in (a) normal and (b) faulty conditions

A fault in the reboiler would affect top product concentration, which is directly monitored. Figure 4.13 shows the top concentration residuals.

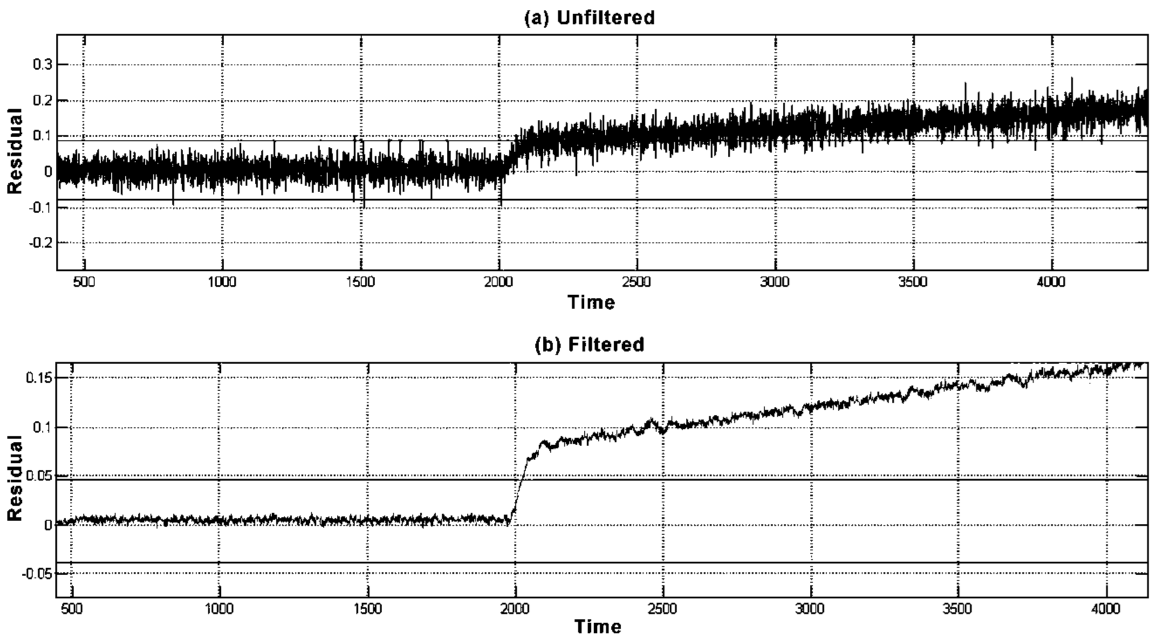

Figure 13: Top concentration residuals in faulty conditions (a) unfiltered (b) filtered

4.5.2.1 Dynamic Risk Assessment and Fault Detection

The risk of fault is obtained by applying the effect of severity on the probability of the fault. Figure 4.14 shows the risk profile of bottom liquid holdup as well as top concentration for the residuals shown in Figure 12. The first risk threshold is equal to 1 and may be used for fault detection in parallel with a priority warning for the operators to respond. The second threshold, placed at 10, would activate safety systems (emergency shutdown systems) [5, 8].

Since the severity of hazard associated with liquid holdup is generally higher than the severity of hazard associated with concentration, a higher-intensity coefficient is assigned for liquid holdup. In this case study, it is assumed that $a_{M_B} = 2$ and $a_{x_D} = 1$.

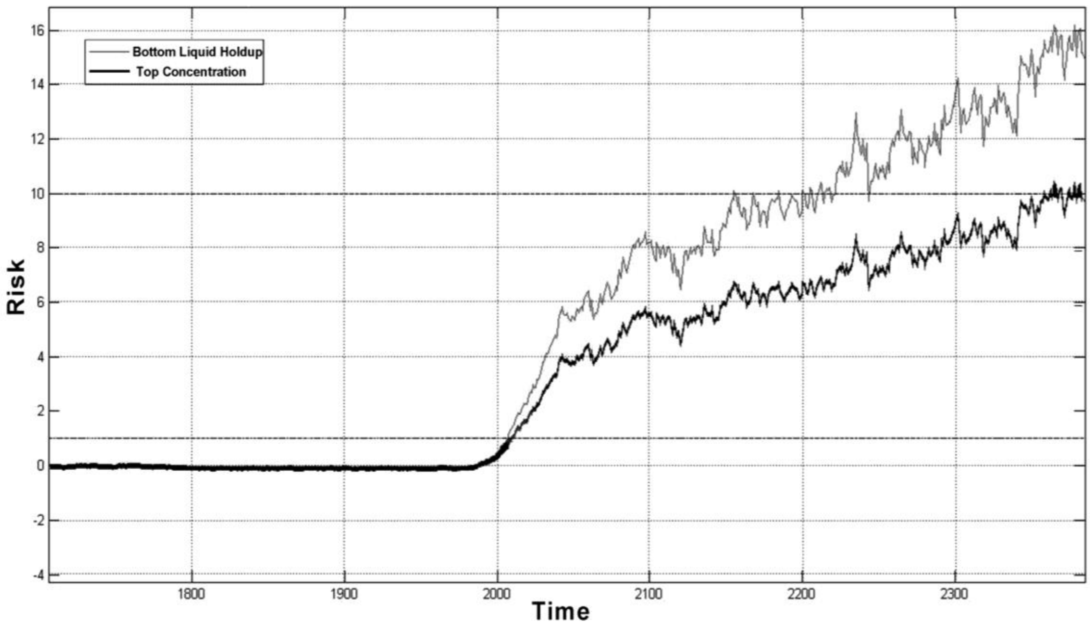

Figure 4.14. Risk profile of bottom liquid holdup (up) and top concentration (down) residuals

Increasing liquid holdup in each stage of the distillation column would be as hazardous as decreasing the liquid holdup. In this case study, the moderation coefficient is considered 1. As discussed in the first case study, the risk threshold may be defined based on the acceptable risk of each process system.

Figures 4.12, 4.13, and 4.14 show how the application of the risk-based fault detection provides early warnings and early activation of safety systems prior to the fault affecting the system. Instead of an alarm based on residuals (t=2030min) or signals crossing the threshold (t=2062min), here an alarm is activated when the risk of operation exceeds the acceptable threshold (t=2009min). Safety systems are also activated at t=2221min, when the risk profile of bottom liquid holdup crosses the second threshold, but the risk profile of the top concentration crosses the second threshold 153min later, at t=2374min.

4.4 Conclusions

A multivariate risk-based fault detection and diagnosis technique targeting the safety issues of a process has been proposed here. In this method, an alarm is activated only when the risk of operation exceeds the acceptable threshold, instead of an alarm being generated based on residuals or signals crossing the threshold. Combining a multivariable residual generation process based on the Kalman filter with the risk-calculation procedure makes this method more robust to false alarms. This method has an increased capacity to distinguish between operational changes and abnormal conditions that have the potential to cause accidents. Other elements considered include the severity of the fault associated with different process variables and the increasing and decreasing rates of fault. The proposed technique is tested on two different process systems. The application of the technique provides early warnings prior to the fault impacting the system and furthermore, the risk-based approach may prioritize warning and recovery options.

Acknowledgments

The authors gratefully acknowledge the financial support provided by Vale Inc., Research & Development Corporation (RDC) and the Atlantic Canada Opportunities Agency (ACOA) under the Vale Research Chair Program.

References

[1] Dash S., Venkatasubramanian V., Challenges In the Industrial Applications of Fault Diagnostic Systems, Computers And Chemical Engineering **2000**, 24, 785-791.

[2] Park S., Himmelblau D. M., Fault Detection and Diagnosis via Parameter Estimation in Lumped Dynamic Systems, ndustrial and Engineering Chemistry Process Design and Development, **1983**, 22, 482-487.

[3] Chetouani Y., Mouhab N., Cosmao J., Stel L. E., Dynamic model-based technique for detecting faults in a chemical reactor, Process Safety Progress, **2003**, 22,3, 183–190.

[4] lsermann R., Fault detection methods for the supervision of technical processes. Process Automation, **1981**, 1, 36-44.

[5] lsermann R., Process Fault Detection Based on Modeling and Estimation Methods A Survey, Automatica, **1984**, 20, 4, 387-404.

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

[7] Misra M., Yue H., Qin S., Ling C., Multivariate process monitoring and fault diagnosis by multi-scale PCA, Computers & Chemical Engineering **2002**, 26, 9, 1281-1293.

[8] Zadakbar O., Imtiaz S., Khan F., Dynamic Risk Assessment and Fault Detection Using Principal Component Analysis, Industrial and Engineering Chemical Research, **2013**, 52, 2, 809–816.

[9] Isermann R., Model-Based Fault-Detection and Diagnosis – Status and Applications, Annual Reviews in Control, **2005**, 29, 71–85.

[10] Venkatasubramanian V., Rengaswamy R., Kavuri S.N., Yin K., A review of process fault detection and diagnosis, Part III: Process history based methods, Computers and Chemical Engineering **2003**, 27, 327-346.

[11] Hsoumi A., El Harabi R., Ali S., Abdelkrim M., Diagnosis of A Continuous Stirred Tank Reactor Using Kalman Filter, International Conference On Computational Intelligence, Modelling And Simulation, **2009**.

[12] Li W., Shah S.L., Xiao D., Kalman filters in non-uniformly sampled multirate systems: For FDI and beyond, Automatica **2008**, 44, 199–208.

[13] Guidelines For Chemical Process Quantitative Risk Analysis, Second Edition, Center For Chemical Process Safety Of The American Institute Of Chemical Engineers, Wiley Publication, **2000**.

[14] Skogestad S., Dynamics and control of distillation columns - A tutorial introduction, IChemE Transaction, **1997**, 75A, 539-562

[15] Sodium Thiosulfate; Material Safety Data Sheet No. S5230; Mallenckrodt Baker: Phillipsburg, NJ, 16 Sep. **2009**, http://www.jtbaker.com/msds/englishhtml/s5230.htm, (accessed April 2011).

[16] Skogestad S., Morari M., Understanding the Dynamic Behavior of Distillation Columns, Ind. & Eng. Chem. Research, **1988**, 27, 10, 1848-1862

[17] Elnemr H.A., Elewa M.M., Expert Failure Detection Technique for Distillation Column Proceedings of the 22nd International Conference on Industrial Electronics, Control, and Instrumentation, **1996**.

Chapter 5: Development of Economic Consequence Methodology for Process Risk Analysis†

Preface

A version of this chapter has been accepted for publication by the *Journal of Risk analysis*, in 2013. Zadakbar was the main lead on the work. The co-authors, Dr. Khan and Dr. Imtiaz supervised the work and helped to develop the methodology.

Abstract

A comprehensive economic consequence methodology with appropriate models for risk analysis of process systems is proposed. This methodology uses loss functions to relate process deviations in a given scenario to economic losses. It consists of four steps: definition of a scenario, identification of losses, quantification of losses, and integration of losses. In this methodology, the process deviations that contribute to a given accident scenario are identified and mapped to assess potential consequences. Losses are assessed with an appropriate loss function (revised Taguchi, modified inverted normal) for each type of loss. The total loss is quantified by integrating different loss functions. The proposed methodology has been examined on two industrial case studies. Implementation of this new economic consequence methodology in quantitative risk assessment will provide better understanding and quantification of risk. This will improve design, decision-making, and risk management strategies.

† *O. Zadakbar, F. Khan, S. Imtiaz, Development of Economic Consequence Methodology for Process Risk Analysis, Journal of Risk Analysis, 2013, In Press.*

5.1 Introduction

In recent years, process operations have become increasingly complex and thus more vulnerable to accidents [1]. As a result, significant amount of work is being undertaken to better monitor processes, evaluate risk, and develop safety systems. Over the years, quantitative risk assessment (QRA) has emerged as an acceptable framework for numeric evaluation of the overall risk of a process [2]. QRA was originally used in aerospace, electronics, and nuclear power industries; it is now being employed in the process industries [2-4]. QRA involves four major steps: hazard identification, consequence analysis, frequency assessment, and risk quantification [5]. Consequence analysis is an integral part of risk assessment. Consequence analysis is generally defined as an assessment of likely consequences if an accident is to occur. The complexity of a QRA depends on the scenario and the availability of data and consequence information [4]. Consequence assessment involves a wide variety of mathematical models, such as source models that predict the release rate of hazardous materials, fire and explosion models, impact intensity models, and toxic gas models [1].

Since 1970, several methodologies have been proposed for quantitative and qualitative risk assessment in process industries [1]. Most of these methodologies focus on prediction or assessment of failure probability [1-4], and few emphasize on consequence analysis in evaluating dynamic risk of a process integrating between process deviation and economic losses.

Greenberg *et al.* [5] listed the ten most important accomplishments in risk analysis in theory, methods and application. A comprehensive overview of methods to quantify risk was presented by Jonkman *et al.* [7] However, many researchers like M. Burgman *et al.* [8] stated the majority of applications of risk analysis consider the probability of events in detail, and pay almost no attention to the severity of consequences. In terms of determining quantitative consequence analysis, Lijian *et al.* [9] calculated personal injury and property damage consequences of vapor cloud explosion disasters of gas pipeline, by TNT equivalent method and composite energy method. Junior *et al.* [10] proposed a systemic accident analysis methodology based on the socio-technical principle of understanding the real operating conditions in which accidents take place. To achieve better eco-toxicological management, the quantitative structure-activity relationships QSARs proposed by Kar *et al.* [11] for risk assessment of chemicals. Aven *et al.* [12] covered issues related to risk assessment and appraisal with respect to petroleum operations. Si *et al.* [13] established a fire-explosion-poisoning quantitative probability model (FEPQPM) to analyze derivative accidents caused by hazardous chemicals leakage.

The present work aims to enhance current knowledge of consequence analysis by developing a means of quantifying consequences in economic terms. It also helps to integrate process information with loss functions for accident scenarios. This methodology, along with the proposed loss function model, will lead to a more accurate quantification of risk and, more importantly, to improved risk management decision-making.

5.1.1 Review of Loss Functions

Quality loss functions are designed to quantify losses associated with deviations of a product characteristic from the target value. Estimation of the potential cost savings resulting from the process improvements is the primary application of quality loss function. Loss function can also serve as measure of performance regardless of the method of quality improvement. It can also be used to determine if an investment to reduce variation is worth the cost. Roy (2010) stated loss functions can also be used to set the limit for the inspection of products [14].

Taguchi [15] proposed a quadratic form of loss function to quantify losses associated with deviations of a product characteristic from the target value. Taguchi's loss function has been widely used to determine engineering tolerance. However, due to the unbounded and symmetric characteristic of the Taguchi loss function, it is unrealistic to use this function in many manufacturing processes. [16] The boundlessness of the original Taguchi loss function has been improved by truncating the function at the points where it intersects the maximum loss. [10] In addition, asymmetric forms have been proposed [17-19]. Spring [16] proposed the reflected or inverted normal loss function (INLF) in response to criticisms of the quadratic loss functions. It is more flexible and provides a more reasonable assessment of the loss associated with deviations from target [20]. F. Sun *et al.* [21] revised the equation to simplify its application and to better employ its flexibility.

A revised inverted normal loss function (RINLF) was also proposed to quantify losses only out of the acceptable region between upper and lower thresholds. [22] In addition to these two major loss functions, a general loss function based on the inversions of gamma loss function has been proposed [20]. Bartholomew *et al.* [23] presented a complementary general class of inverted probability loss functions (IPLF) based on the inversion of common probability density functions, including the inverted normal loss function (INLF), inverted gamma loss function (IGLF), inverted beta loss function (IBLF), and their associated properties for the uniform, normal, gamma, and beta distributions. Although they recommended the normal distribution for the INLF, the gamma distribution for the IGLF and the beta distribution for the IBLF, they also suggested some other distributions to use, such as uniform distribution, when the process measurements do not follow exactly the corresponding conjugate distributions. These functions have been widely used in many other engineering areas, particularly in quality control, to quantify the economic losses associated with deviations from the target value of a product characteristic [20]. These loss functions are yet to be applied in process safety analysis.

5.1.2 Objectives of Current Work

This chapter introduces the idea of loss functions to quantitative risk assessment of process systems. It attempts to overcome the lack of integration between process deviation and economic losses. Loss functions are developed to quantify losses associated with the abnormal behavior of process parameters that may cause an accident [24]. The developed methodology will be able to assess losses at any point in time, based on the status of process variables.

5.1.3 Organization of the chapter

The remainder of the chapter is organized as follows. Section 5.2 presents the economic consequence methodology for the risk analysis of process systems. This methodology is based on loss functions assigned to each type of loss. In Section 5.3, the proposed methodology is applied to two process systems. The first case study describes the consequence analysis for a knock out drum, while the second case study analyzes the explosion in the ISOM unit in the BP Texas City refinery. The Chapter ends with concluding remarks in Section 5.4.

5.2 Consequence Assessment Methodology

Chemical processes are complex systems that may involve abnormal conditions, for various reasons. A process that deviates from normal operation may cause many direct and indirect consequences. The proposed methodology starts with defining the possible fault scenarios; it identifies different classes and subclasses of loss, quantifies each loss using an appropriate loss function, and finally integrates all losses to determine the overall loss. The overall methodology for developing an economic consequence methodology for risk analysis of process systems is shown in Figure 5.1.

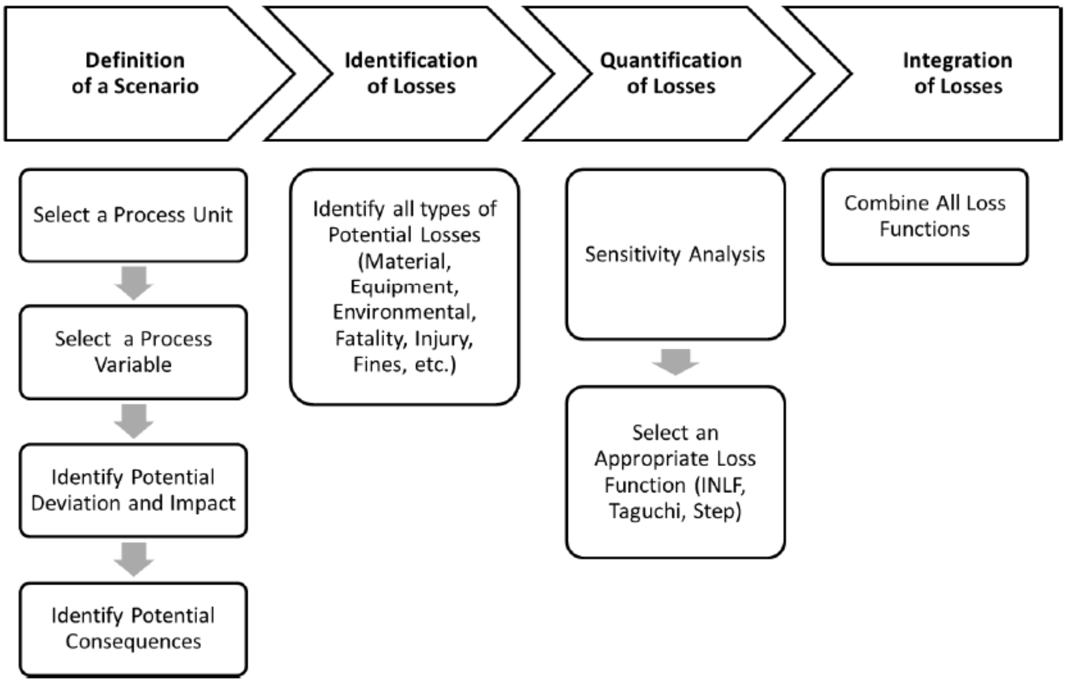

Figure 5.1. The flowchart of economic consequence analysis using loss functions

5.2.1 Defining a Scenario

The first step in the methodology is defining a scenario. In this step, a process unit is selected and the process variables related to the unit, such as temperature, pressure, concentration and level, are identified. The next step is to consider the pathways of a system's poor performance whereby the deviation of a process variable can cause incidents or accidents. For instance, in a pressure vessel containing hydrocarbon, the relevant process variables are pressure, temperature, and flowrate. Blocked outlets or a utility failure could cause the vessel's internal pressure to rise and exceed the maximum allowable working pressure. If relief systems fail to act, the failure could lead to release, fire, or explosion.

5.2.2 Identification of Losses

In this step, all significant types of losses due to potential deviation are identified. These losses include but are not limited to material loss, equipment loss, environmental loss, injury, fatality, and fines. These losses are represented by appropriate functions in subsequent steps. Identification of losses depends on the accident scenario. For each given scenario (process interruption/release/fire/explosion) the system is modeled using appropriate methods e.g. gas dispersion modeling to find the accident area. The model output is used to determine the number of people and/or equipment inside of this area or the amount of release leading to environmental loss.

5.2.3 Quantification of Losses

The quantification of losses is an important aspect of the overall methodology. Appropriate functions are chosen to quantify different type of losses. These functions are used particularly to quantify losses due to process malfunctions, releases, fires, and explosions. In this study, loss functions are divided into instant (step) and non-instant (Taguchi, inverted normal, etc.) loss functions.

5.2.3.1 Instant Losses

If the consequences associated with a process variable lead to an instant loss such as civil claims, injuries, or fatalities, a step function could be assigned to these types of losses. For injuries and fatalities, this function is developed based on the number of personnel present inside the accident area and their positions. For example, in case of a pressure vessel explosion near the control room with two operators and one field operator, the step function, as shown in Figure 5.2, can be assigned for fatality loss associated with the vessel's pressure. In case of an explosion at a low operating pressure, only the operator working around the vessel might receive fatal overpressure. However, with an explosion at a high operating pressure, the damaged area could be larger and two operators inside the control room may even be exposed to fatal overpressure.

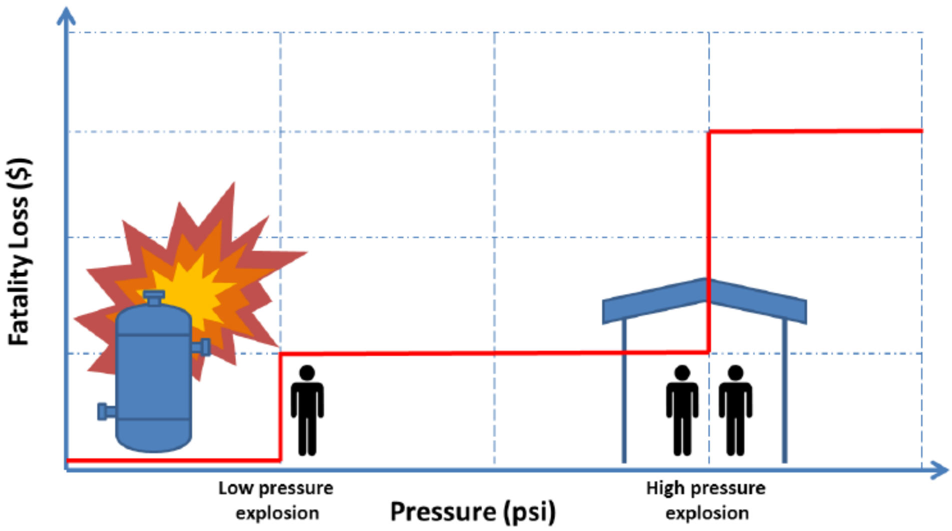

Figure 5.2. A step function assigned to fatality loss associated with process variable changes, based on the number of personnel exposed and their position

The step function assigned to the civil claims (Figure 5.3) consists of three steps: step 1 includes claims for individual minor injuries; step 2 includes individual major non-recoverable injuries, individual disabilities, and minor small group injuries (involving fewer than 10 people); and phase 3 includes significant losses, and large group injuries (involving more than 10 people). These three steps obviously depend on the number of people (workers and the public) involved and their positions.

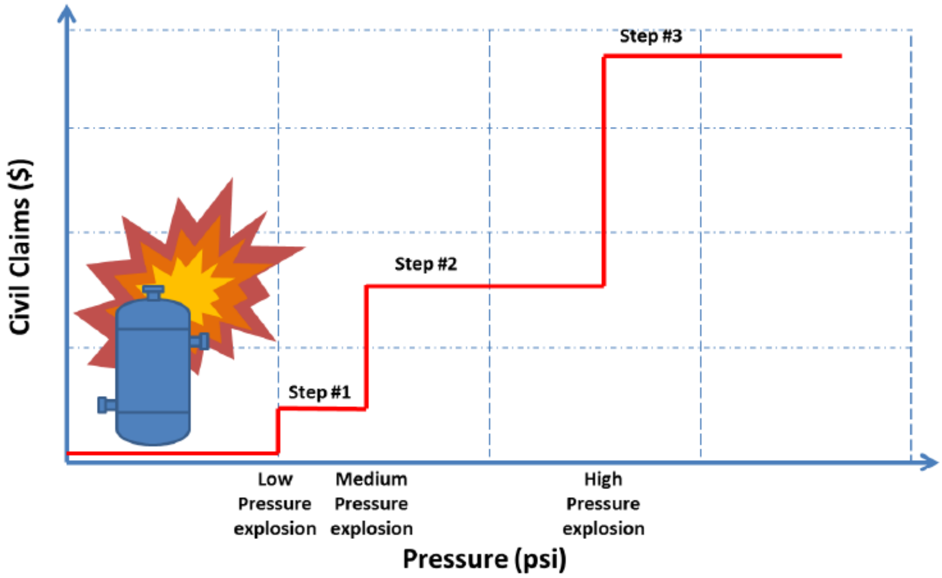

Figure 5.3. A step function assigned to civil claims associated with process variable change and the number of people (workers and the public) exposed

5.2.3.2 Non-Instant Losses

The concept of loss functions originated in quality control research. These functions are used to quantify losses associated with deviation from a set point. In the present study, loss functions are used to calculate the economic loss associated with deviations of a variable or parameter from the set point or from an acceptable range.

If the loss increases progressively (non-instantly) the inverted normal loss function (INLF) or Taguchi loss function can be assigned. In this step, an appropriate loss function for each type of loss is selected to quantify the economic losses. At any point, the loss function's input is measurements from sensors or calculated value from models.

5.2.3.2.1 Taguchi's Loss Function

Taguchi (1986) introduced the concept of quality loss functions. Taguchi's loss function is a quadratic function modified to assess losses associated with deviations of a product characteristic from the target. Taguchi defined the quadratic loss function as [15]:

$$Loss = k(y - T)^2 \tag{5.1}$$

where y is the quality characteristics or process variable, k is the coefficient of quality loss or the maximum loss, and T is the target value or the set point. However, this loss function is unbounded; the following revised Taguchi equation makes the function bounded: [17]

$$Loss = \begin{cases} k(y-T)^2 & if \quad y - T < \sqrt{k/B} \\ k & if \quad y - T > \sqrt{k/B} \end{cases} \tag{5.2}$$

where k is the maximum value of loss and B represents the coefficient of quality loss within the specified limits.

5.2.3.2.2 Inverted Normal Loss Function

Because of the limitations of Taguchi's loss function, an alternative function was developed by Spiring, [15] called inverted normal loss function (INLF). The INLF is bounded and provides a more reasonable assessment of the loss associated with deviations from the target by using the shape parameter [16]. Inverted normal loss function is a mirror image of a normal density function (on the horizontal axis). It has the minimum value zero at the target, and this increases as the variable moves away from the target and levels off at a maximum loss. The general form of inverted normal loss function is [17]:

$$Loss = K\left\{1 - \exp(-\frac{(y-T)^2}{2\gamma^2})\right\} \qquad \gamma = \frac{\Delta}{4} \tag{5.3}$$

where Δ is the distance from the target to the point where the maximum loss K first occurs. The function has been modified to simplify its application and make it more flexible. The modified form of inverted normal loss function is [21]:

$$Loss = \frac{K_\Delta}{1-\exp\{-1/2(\frac{\Delta}{\gamma})^2\}}\left\{1 - \exp(-\frac{(y-T)^2}{2\gamma^2})\right\} \tag{5.4}$$

where K_Δ represents the value of the loss at a specific distance Δ from the target. The modified inverted normal loss function presents a convenient form for fitting the loss function to the user's perception of loss [21].

5.2.3.3 Choice of Loss Function

The loss functions described in the previous section represent different types of losses. For non-instant losses, if a small deviation from set point causes a high or medium value of loss, the inverted normal loss function must be selected. However, if a large deviation from a set point causes a small value of loss, then the revised Taguchi loss function is a better choice (Figure 5.4).

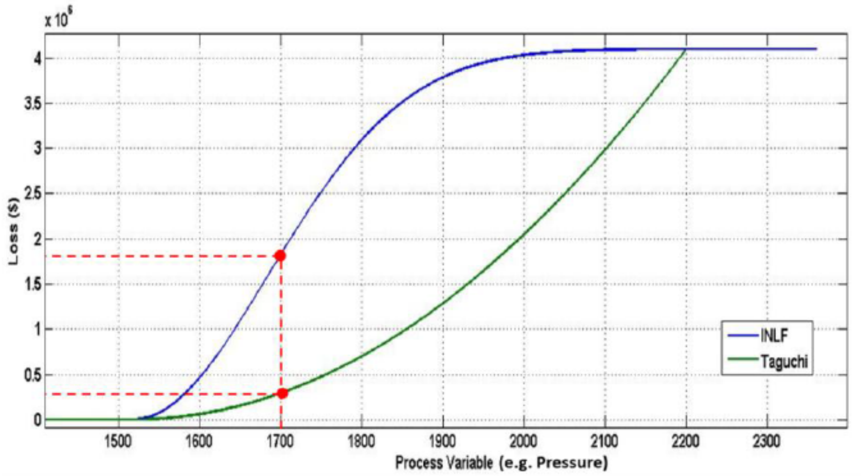

Figure 5.4. Comparing the sensitivity of modified INLF and revised Taguchi loss function

When $\gamma = \Delta/4$, the INLF weights small departures from the target more heavily than Taguchi's loss function. Therefore, the value of γ has important implications for the choice of loss function. If the value of γ is very high, the difference between the INLF and Taguchi's loss function reduces to near zero. The choice of $\gamma = \Delta/4$ is appropriate only if a small deviation from target results in a substantial loss.

5.2.4 Integration of Losses

Once the losses for each individual consequence have been modeled, an additive rule is applied to calculate the overall loss function. For example, in a given scenario, if two types of losses — material loss and equipment loss — are expected, the overall loss is calculated by adding the individual loss functions assigned to the material loss and to the equipment loss.

5.3 Case Studies

The proposed methodology is applied to two process systems. The first case study is a safety critical knock out drum near the flare system and the second one is the well-known BP Texas explosion.

5.3.1 Case Study 1 – Knock Out Drum

Flare systems in oil, gas, and petrochemical plants play critical roles in the safe operation of plants. Flare systems are designed to dispose waste gases and discharged liquids from process units safely. A flare system ensures maximum combustion of hydrocarbons while minimizing hazardous emissions into the atmosphere. Liquids and gases are separated and gases are burned off at the flare stack. Knock out drums are vessels that collect condensed liquids from the gas stream. These vessels are sized for containment of liquid carryover [25]. Condensed liquids are collected at the bottom of the drum and pumped off. Most flare knockout drums operate at relatively low pressures [26]. During flaring at high gas and liquid loads, the flare knock out drums must be able to prevent the development of waves or entrainment of disengaged liquid [27].

Based on API standard 521 [28], the risk of overfilling the flare knockout drum is assessed. Most flares are not designed to effectively combust liquid. Discharging liquid from the flare may cause flame-out, excessive smoke, unburnt hydrocarbon emission release, burning rain or pool fires around the flare stack. On the other hand, accumulation of excessive liquid in the knock out drum may cause overpressure inside the drum, which can lead to leakage, hydrocarbon release, and vapor cloud explosion.

5.3.1.1 Scenario Description

In this case study, the flare knock out drum in phase 19 of the South Pars gas field development, Tombak, Iran is considered. This knock out drum is designed and integrated into the system to remove entrained droplets from the vapor stream. In this drum, the liquid level is the most important process variable. If it crosses the high level, a duty pump is started to discharge the excess liquid. However, if the level crosses the high high level (HHL), a second duty pump is started to increase the discharge rate. In the meantime, crossing the HHL will initiate the emergency shutdown[1] (ESD) level 1[2]. This ESD level can be activated from the main control room [29].

Other process variables in this drum are the temperature, pressure, and feed flowrate. This drum is capable of handling the total plant flowrate, so the flowrate as a process variable could not lead to a hazardous condition. Thus, no control and monitoring instrument is needed for this process variable. If the temperature decreases, the built-in heater will be started to warm up the drum. This heater's power is low and will not generate enough heat to lead to a hazardous condition.

[1] ESD is a system designed to respond to a condition in the plant which may itself be hazardous or may lead to hazardous conditions if no action is taken.(30)

[2] ESD Level 1 typically shuts down the entire process facility.(30)

For this case study, a scenario is considered in which the liquid level in the knock out drum is rising. It could cause overpressure and leakage from flanges. This could lead to hydrocarbon release and vapor cloud formation. Ultimately, it may result into an explosion. Figure 5.5 graphically illustrates the rising level scenario which leads to an explosion.

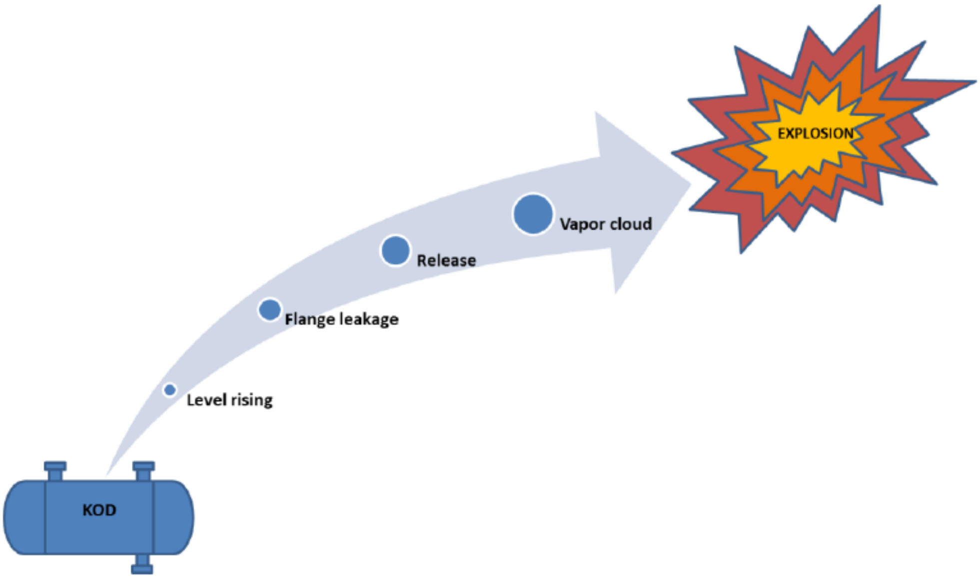

Figure 5.5. Illustration of the explosion scenario in case study 1

5.3.1.2 Consequence Analysis Using Loss Functions

Table 5.1 provides actions and consequences associated with a rising liquid level in the knock out drum that may potentially lead to explosion. The BP Texas City refinery accident is an example of a similar sequence of events that is studied in a subsequent section [25].

Table 5.1. Actions and consequences associated with rising levels in the knock out drum [29]

THRESHOLD	LEVEL	Actions and consequences
HL	1220 mm	Pump #1 is ON
HHL	1320 mm	Pump #2 is ON Emergency shutdown level 0 Control room permission Plant shutdown No production, lots of waste gas
>HHL	>1320 mm	Flange leakage release Vapor cloud Explosion Equipment loss (6 drums, 12 pumps) Fatality/Injury (2 persons) Production loss (1 month)

The explosion scenario has been modeled using PHAST [31]. DNV PHAST is a comprehensive hazard analysis tool for gas dispersion modeling. PHAST uses a proprietary dispersion model called a unified dispersion model (UDM) [31]. PHAST simulation results are shown in Figure 5.6.

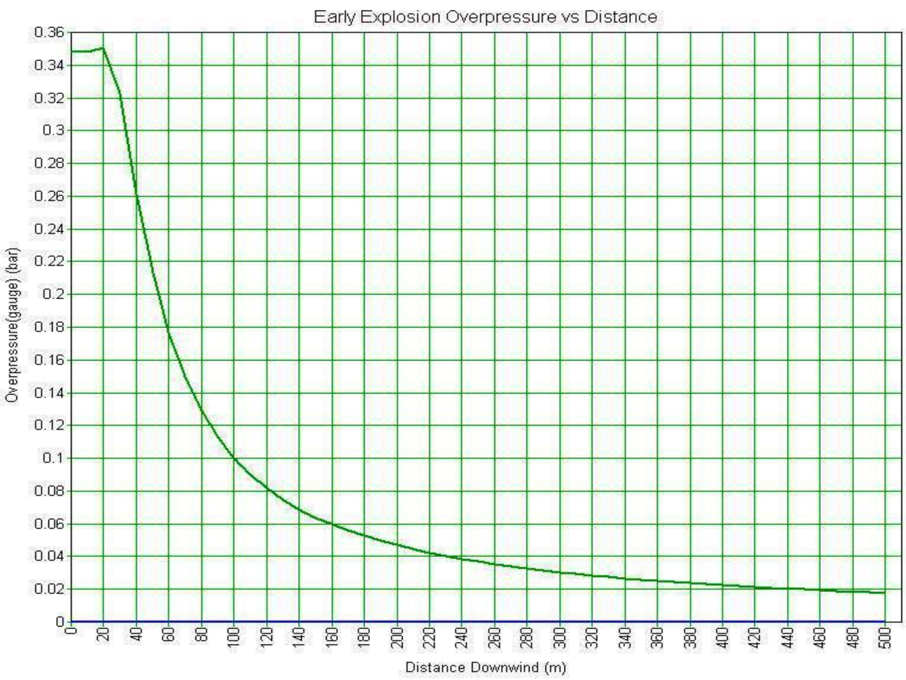

Figure 5.6. Explosion overpressure vs. distance from the drum[(29)]

Based on the results of the overpressure modeling, all equipment inside the area of 70 m from the explosion source, where the overpressure is more than 150 mbar, will be severely damaged [32].

Due to the high sensitivity of equipment losses to level fluctuation, inverted normal loss function is used to model equipment losses associated with a rise in liquid levels. In this case, the maximum loss is $4,000,000 and the shape factor is Δ/4=50:

$$\text{Equipment Loss} = \frac{K_\Delta}{1-\exp\{-1/2(\frac{\Delta}{\gamma})^2\}}\left\{1-\exp(-\frac{(y-T)^2}{2\gamma^2})\right\} == \frac{\$4{,}000{,}000}{1-\exp\{-1/2(\frac{200}{50})^2\}}\left\{1-\exp(-\frac{(Level(t)-1220)^2}{2(50)^2})\right\} \quad (5.5)$$

Figure 5.7 shows the level profile in the knock out drum, and Figure 5.8 shows equipment loss associated with rising levels in the knock out drum.

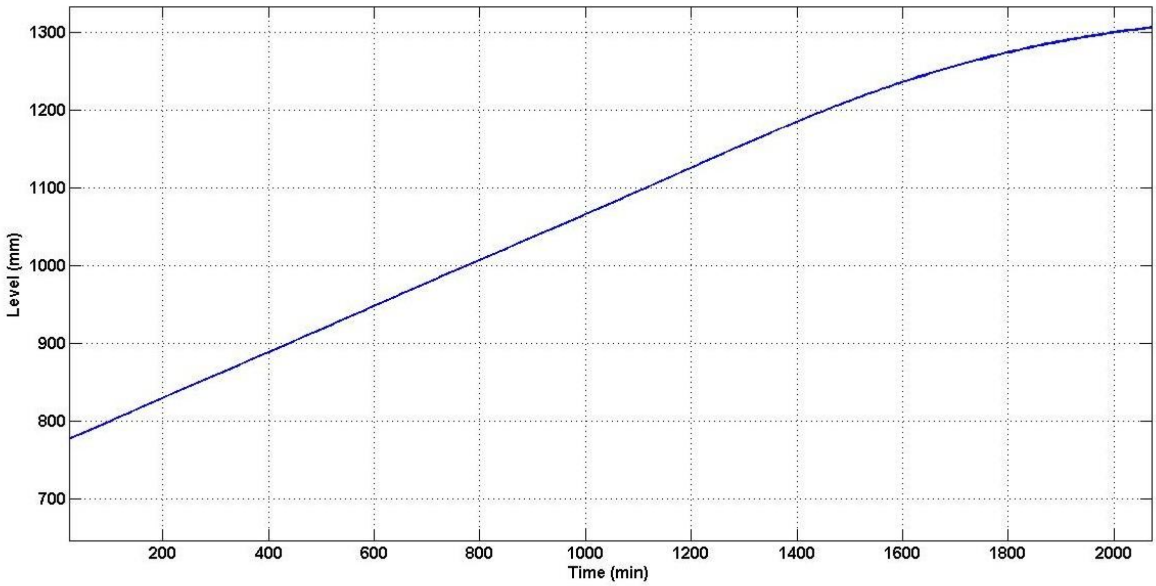

Figure 5.7. Level profile in the knock out drum

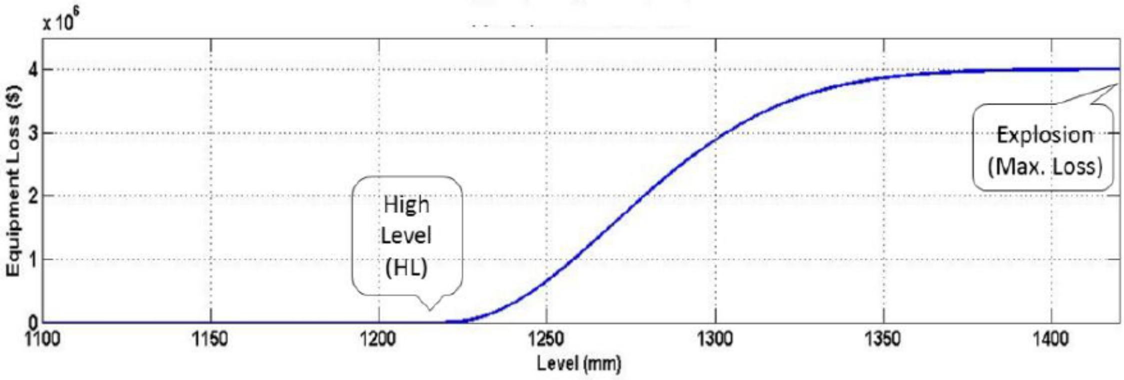

Figure 5.8. Equipment loss associated with rising liquid level in the knock out drum

The knock out drum is a part of the flare unit, which is the most safety critical part of the plant. In case of an explosion in this unit, the whole plant will be shut down, and it is assumed that one month's production will be lost. If we include hidden costs with material and production costs, the material loss can be modeled non-linearly using inverted normal loss function, with a maximum loss of $720,000,000 (given a total production rate of 1000 million ft^3/hr). Figure 5.9 shows material losses associated with the explosion.

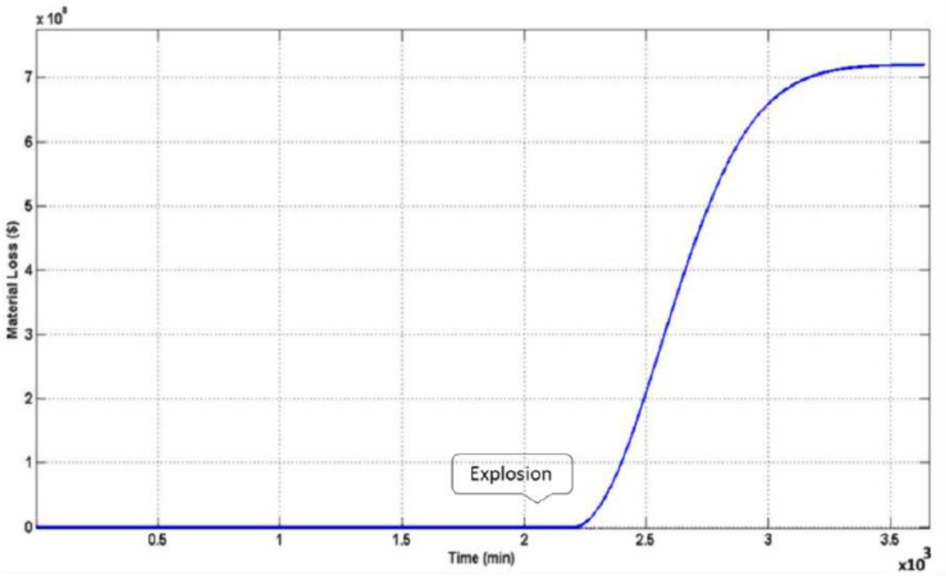

Figure 5.9. Material loss due to explosion originating in the knock out drum

In this unit, two field operators are available and an explosion may cause two fatalities. Figure 5.10 shows the fatality loss in case of explosion of the flare system's knock out drum.

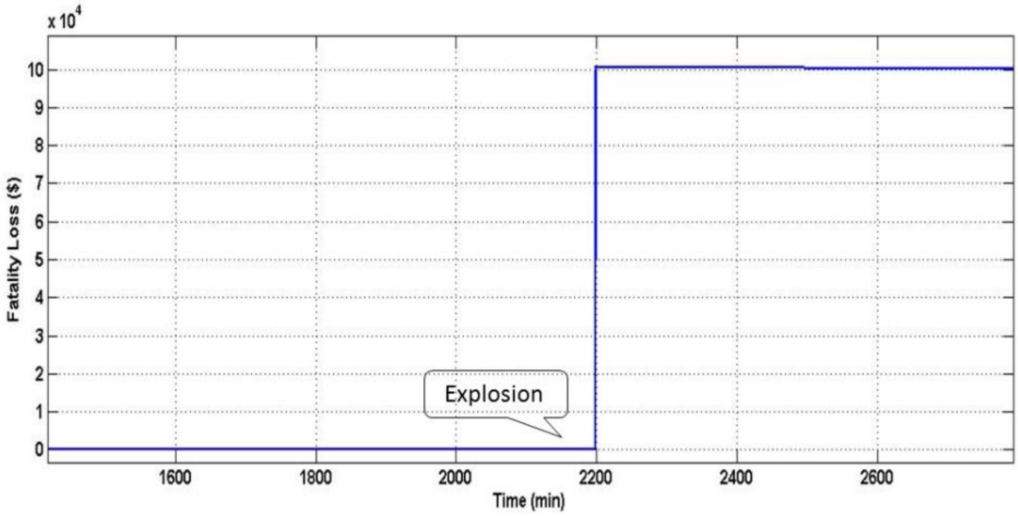

Figure 5.10. Fatality loss due to explosion originating in the knock out drum

Total loss is determined by combining fatality, equipment loss and material loss. Figure 5.11 illustrates the profile of total loss due to an explosion of the knock out drum in the flare unit. In this case study, material loss is the greatest contributor to total loss. Equipment loss is initiated by crossing the high level (HL); it reaches its maximum when the explosion occurs. The fatality loss is zero before the explosion; it suddenly (step function) reaches its maximum when the explosion takes place. On the other hand, material loss due to explosion and total plant shutdown is initiated after the high high level (HHL) is crossed.

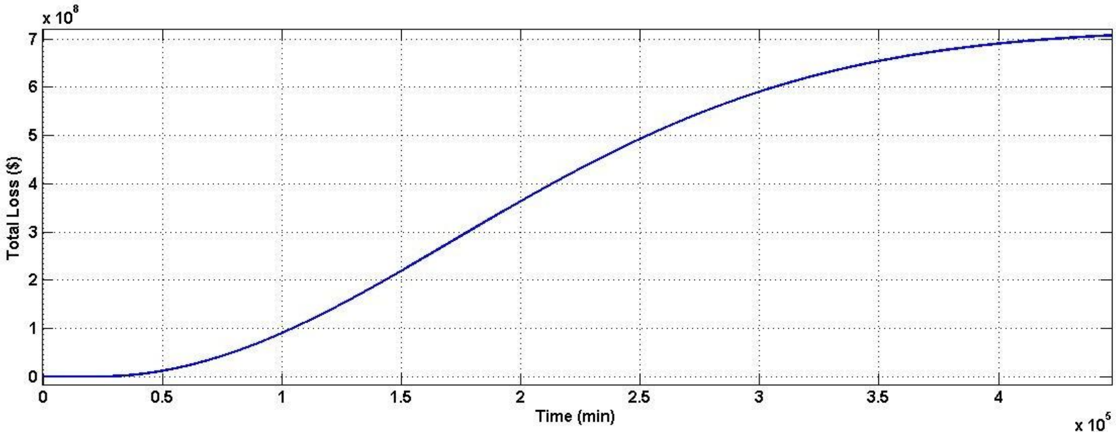

Figure 5.11. Total loss in case of explosion originating in the knock out drum

In conclusion, the scenario and potential sequences and consequences for this case study are as follows: liquid levels rising in the knock out drum, overpressure, leakage from flanges, hydrocarbon release, and ultimately vapor cloud explosion. The significant types of losses due to explosion are material loss, equipment loss, and fatality loss. The overall loss is determined by combining all loss functions, which provides a comprehensive economic consequence model for deviations of process variables. This economic consequence model can be used in conducting a dynamic risk assessment.

5.3.2 Case Study 2 – BP Texas City Accident

A retrospective economic consequence analysis is performed for the distillation tower flooding that led to the catastrophic explosion on March 23, 2005, at the British Petroleum (BP) refinery in Texas City, USA. In that accident, 15 people died, 180 were injured, the community was alarmed, and financial losses exceeded $1.5 billion [33-34].

5.3.2.1 Scenario Description

The isomerization (ISOM) unit of the Texas refinery converts low octane feed into higher octane components. The raffinate splitter tower of this unit is a distillation tower that takes non-aromatics feed from the aromatics recovery unit (ARU) and fractionates it into light and heavy components. The raffinate section of the ISOM unit is illustrated in Figure 5.12.

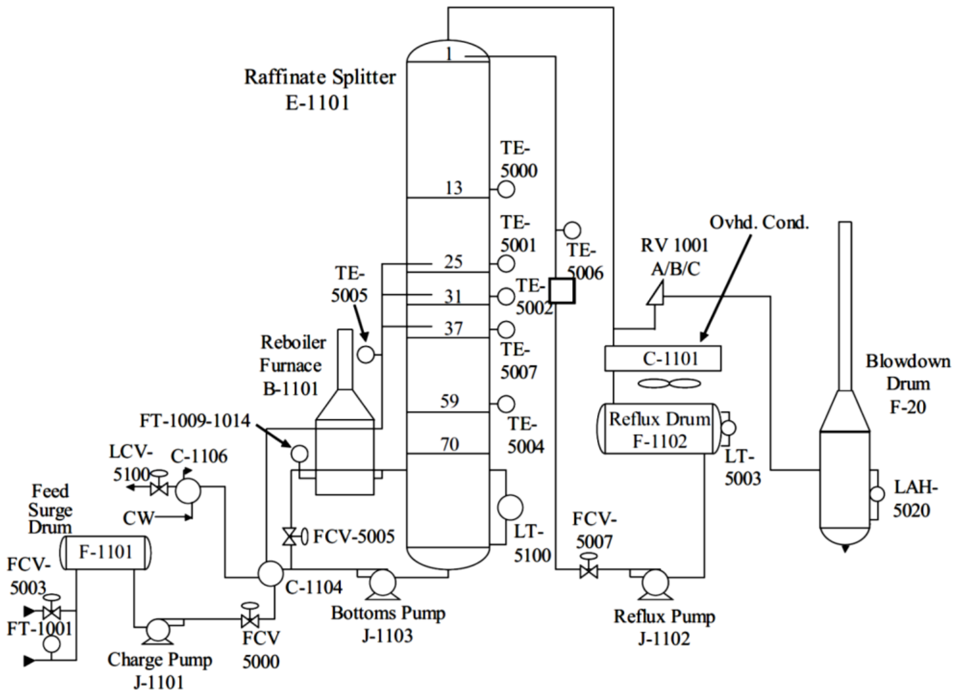

Figure 5.12. Raffinate section of the ISOM unit [35]

The sequence of events that led to the accident is as follows. The liquid level started rising in the raffinate splitter tower during the startup of the ISOM unit. As a result, the raffinate splitter tower was overfilled; pressure relief devices opened, and flammable liquid was released from a blowdown stack that was not equipped with a flare. The release of flammable liquid led to an explosion. The timeline and detailed descriptions of the events are given in Table 5.2 [36]. Information was combined from different sources to create a profile, shown in Figure 5.13, for levels in the raffinate splitter tower.

Table 5.2. Timeline of events from recommencing the ISOM unit startup to the vapor cloud explosion [36]

Time	Events
09:51 a.m.	Startup recommences
09:51 a.m.	Tower level is 2.7 m
11:16 a.m.	Four burners in the furnace are lit. The level transmitter reads 2.6 m but the level is actually 20 m.
11:50 a.m.	Fuel to the furnace is increased; the actual tower level is 30 m but the transmitter reads 2.6 m.
12:41 p.m.	The tower's pressure rises to 33 psig (228 kPa); operators open the 8-inch NPS chain valve to reduce pressure.
12:42 p.m.	Fuel gas to the furnace is reduced; the actual tower level is 43 m, but the transmitter reads 2.4 m.
12:59 p.m.	Heavy raffinate flows out of the tower and matches the feed flow.
	The actual level in the tower is 48 m but the transmitter reading is 2.4 m.
01:14 p.m.	Tower pressure spikes to 63 psig (434 kPa); all three relief valves open.
01:15 p.m.	Fuel gas to the furnace is reduced.
01:16 p.m.	The blowdown drum and stack are overfilled and the alarm fails to sound.
01:19 p.m.	Flammable hydrocarbon is released from the 34 m stack.
01:20 p.m.	Vapor cloud ignites and explodes.

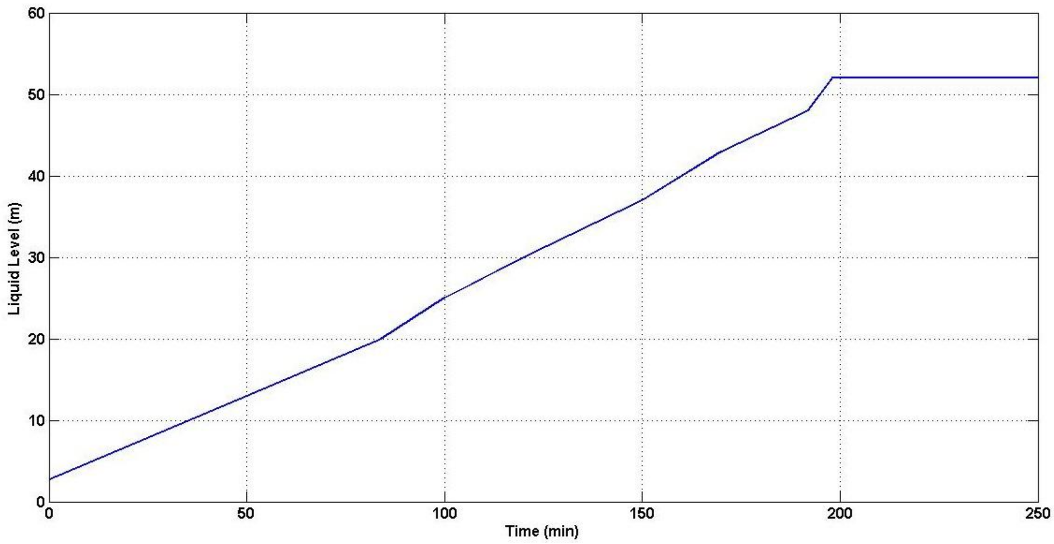

Figure 5.13. Level profile in the raffinate splitter tower

Based on the reports [27-29], major losses associated with that explosion include material loss, equipment loss, fines, and civil claims.

5.3.2.2 Consequence Analysis Using Loss Functions

The explosion caused $30 million in equipment and plant property damage at the BP refinery in Texas City [36]. Due to the high sensitivity of equipment losses to fluctuation in liquid levels, inverted normal loss function is used to model equipment losses associated with rising liquid levels. In this case, the maximum loss is $30,000,000 and the shape factor is Δ/4=12.32:

$$Equipment\ Loss = \frac{K_{\Delta}}{1-\exp\{-1/2(\frac{\Delta}{\gamma})^2\}}\left\{1-\exp(-\frac{(y-T)^2}{2\gamma^2})\right\} == \frac{\$30{,}000{,}000}{1-\exp\{-1/2(\frac{52}{12.32})^2\}}\left\{1-\exp(-\frac{(Level-2.7)^2}{2(12.32)^2})\right\} \quad (5.6)$$

Figure 5.14 shows equipment loss associated with rising liquid levels in the distillation tower.

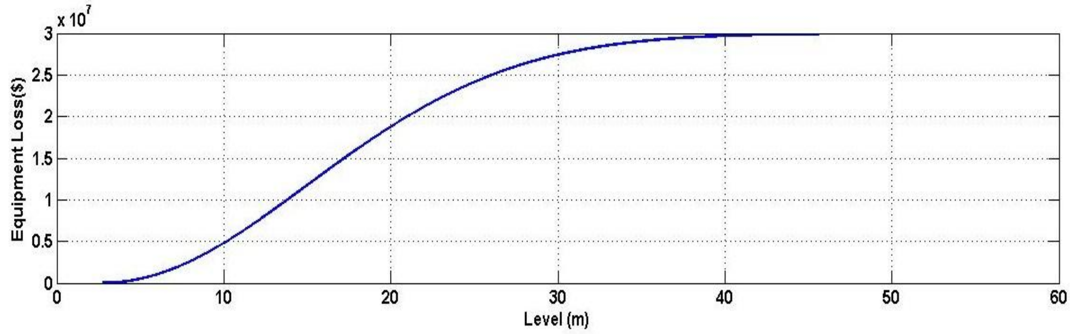

Figure 5.14. Equipment loss associated with rising level in the tower and explosion

Considering hidden costs along with material and production costs, the material loss can be modeled non-linearly using an inverted normal loss function. The maximum loss of $500,000,000 has been obtained by considering the refinery total capacity rate of 460,000 barrels per day. Figure 5.15 shows material losses associated with the explosion.

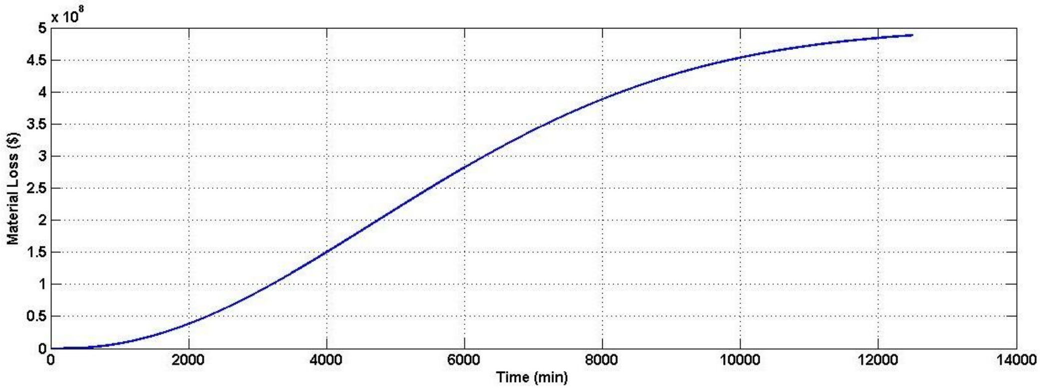

Figure 5.15. Material loss associated with rising level in the tower and explosion

The explosion caused more than \$21 million in fines and \$1 billion in civil claims [36-37]. BP agreed to pay \$21,361,500 in penalties and to abate all hazards for which they were cited by the Occupational Safety and Health Administration (OSHA). A summary of citations and proposed penalties can be found in OSHA national news release (USDL 05-1740, September 22, 2005) [38]. The Taguchi loss function is used to model fines associated with an explosion due to a rise in liquid levels. Figure 5.16 and 5.17 show fines and civil claim profiles respectively.

In this case study, the maximum fine is \$21 million and the coefficient $B = 8640.24$. Fines due to an explosion can be modeled using the following Taguchi loss function:

$$Fines = 8640.24(Level - 2.7)^2 \quad (5.7)$$

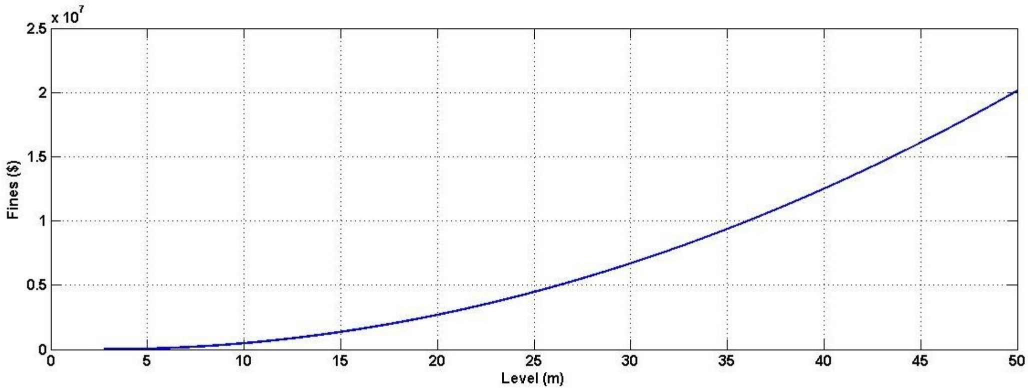

Figure 5.16. Fines associated with rising level in the tower and the explosion

Maximum civil claims of $1 billion can be modeled using the suggested step function (Figure 5.17).

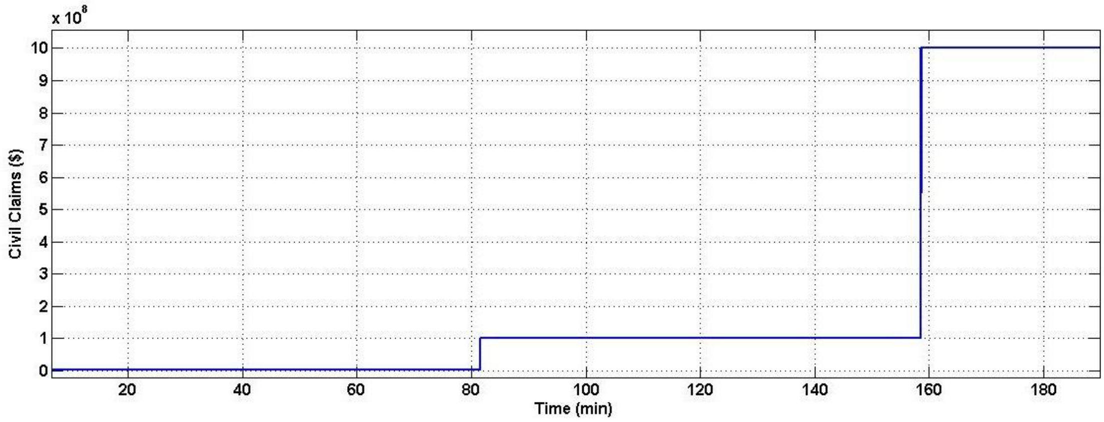

Figure 5.17. Civil claims associated with in the tower and the explosion

As discussed in the methodology, although fatalities and injuries are parts of civil claims, they can be modeled separately with a step function. Figure 5.18 shows fatality losses associated with the explosion. To precisely initiate each step in a fatality or injury loss profile, the operator's position around the explosion source is needed. In this case study, it is assumed 2 workers were in the control room and 13 workers were at the trailer near the ISOM unit [36].

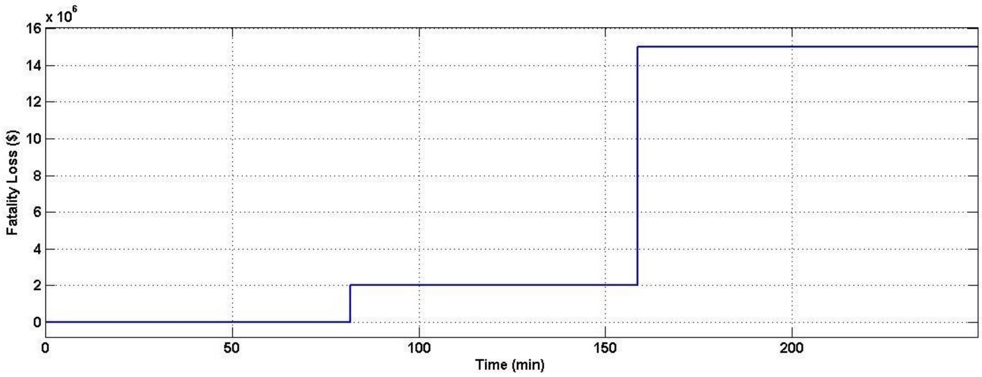

Figure 5.18. Fatality loss associated with rising level in the tower and the explosion

Total loss is determined by combining fines, civil claims, equipment loss, and material loss. Figure 5.19 illustrates the profile of total loss due to explosion in the ISOM unit.

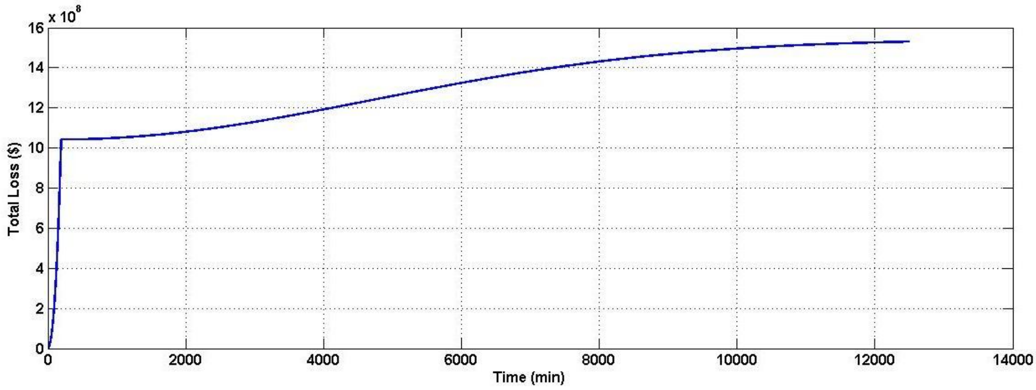

Figure 5.19. Total loss associated with rising level in the tower and explosion

A good agreement is observed between calculated losses and reported losses. Figure 5.19 shows that fines, civil claims, and equipment losses contribute 70% of the total economic losses in a short time. The main reasons these factors cause the majority of losses are lack of safety measures, failure of automated controls in the splitter tower, inadequate warning systems to warn overfilling of the distillation tower, lack of flare systems to safely burn flammable hydrocarbons entering the blowdown system, and the location of a temporary work area close to the process facility.

5.3.2.3 Evaluation of Safety Measures

Several possible modifications of the process and control system could have lowered potential loss. The proposed consequence methodology can be used to select the most feasible option. In this case study, a bypass line at the bottom of the distillation tower is added, as an additional safety measure, to simultaneously increase the discharge rate and decrease the rate of pre-heating. Figure 5.20 shows how the bypass line affects the rate of level increases inside the splitter tower.

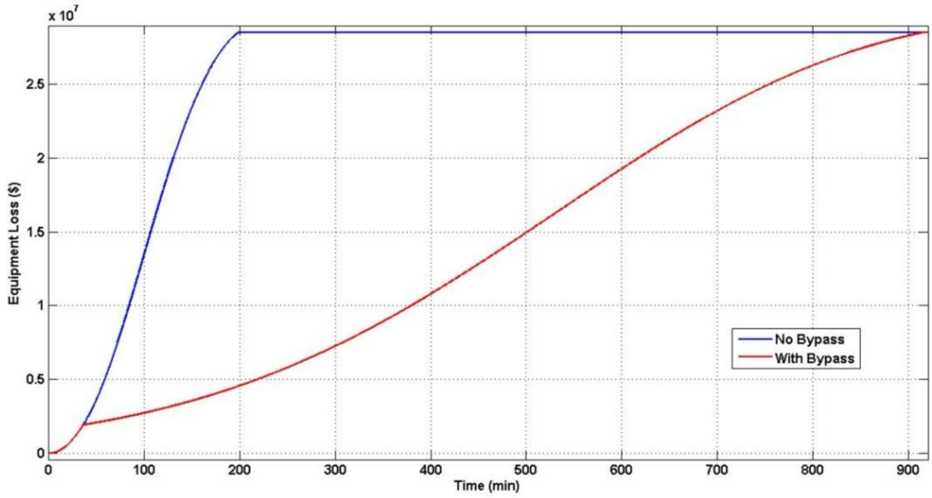

Figure 5.20. Effect of bypass line on rising liquid level inside the tower

Adding a bypass line at the bottom of the distillation tower may reduce the rate of level increase, which provides more time to detect, isolate, diagnose, and resolve the fault. Following this new level profile and performing the proposed economic consequence analysis may result in a new total loss profile, illustrated in Figure 5.21a.

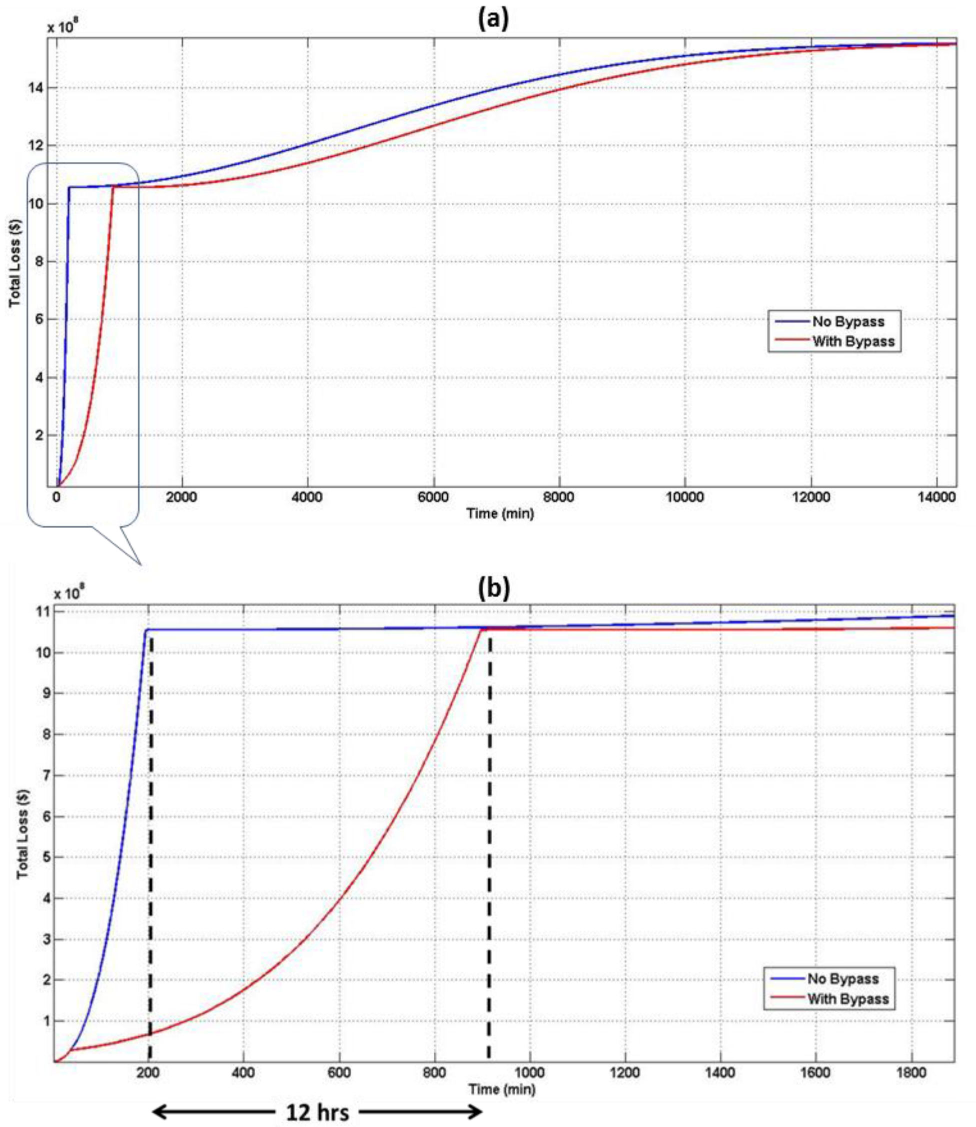

Figure 5.21. (a) Total losses associated with rising liquid level in the tower and the explosion in ISOM unit with and without bypass line; (b) Effect of bypass line on total loss associated with rising level in the tower and explosion in ISOM unit

The proposed methodology for consequence analysis using loss functions is capable of modeling the system after adding suggested safety measures. In this case study, the results of modeling the system with the proposed methodology show that the suggested bypass line could have offered operators an extra 12-hour window of time in which to detect faults and prevent the catastrophic explosion (Figure 5.21b).

Using the proposed methodology, losses resulting from this accident can also be modeled based on the temperature profile of the distillation tower, as indicated in Figure 5.22. Using the same methodology, the total loss associated with temperature fluctuations has been estimated and compared with the total loss associated with liquid level fluctuations with and without the bypass line (Figure 5.23).

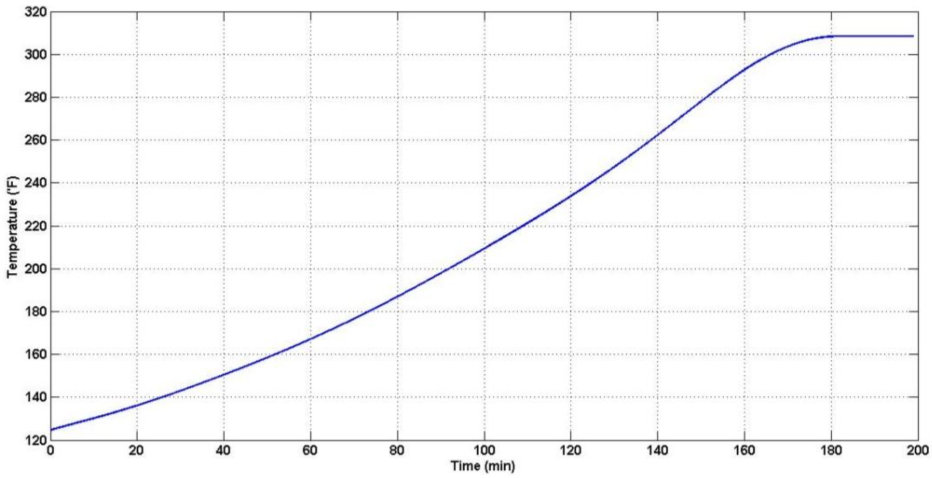

Figure 5.22. Temperature profile inside the tower of ISOM unit

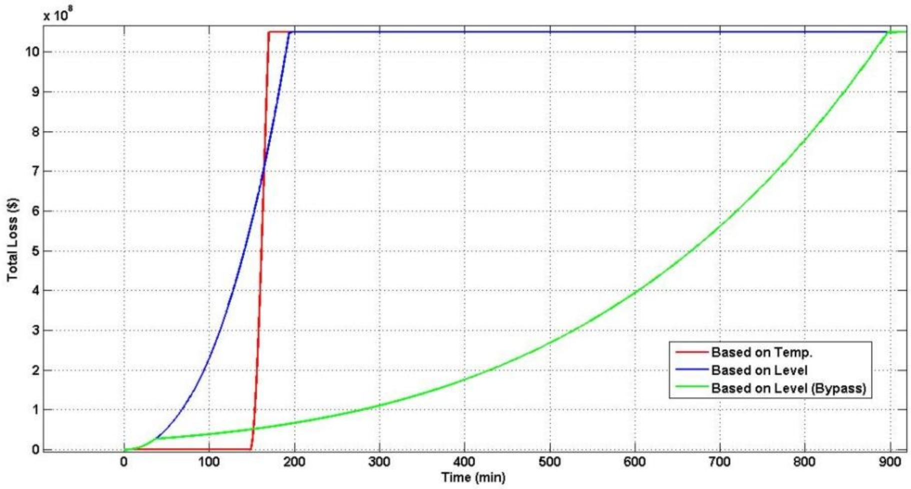

Figure 5.23. Comparison of total losses calculated based on liquid level and temperature profile of distillation unit for the period of 0-900 min

5.3.2.4 Summary of the Case Study

The scenario for this case study consists of following steps: rising level in the distillation tower leading to overpressure, release from the blowdown stack, and ultimately vapor cloud explosion. The significant types of losses due to explosion were identified as material loss, equipment loss, injuries and fatalities, fines, and claims. The inverted normal loss function was determined to quantify material and equipment losses, and Taguchi's loss function was used to quantify fines. Finally, the overall loss was determined through a combination of all loss functions, which provides a comprehensive economic consequence model based on process deviations. Meanwhile, the process system in the scenario was updated by adding a bypass line as a safety measure. The total loss associated with level (with and without the bypass line) has been compared. The second scenario was also defined based on rising temperature, overpressure, release from the blowdown stack, and ultimately vapor cloud explosion.

5.4 Conclusions

A new methodology for conducting consequence analysis in process industries is proposed in this Chapter. This methodology fills the gap in existing economic consequence models by relating economic losses to process variables. It quantifies economic losses associated with the deviation of operational variables from a target value/range. A series of loss functions are proposed as part of the methodology to quantify major losses such as material loss, equipment loss, environmental loss, fatalities and injuries, fines, and claims faced in the process industry.

The proposed method devised a technique to partition total losses to different categories; it illustrates the potential impacts of an accident. The method also shows the dynamic changes of loss under different abnormal conditions. The method is also useful for calculating the risk while the alternative safety measures are added. The proposed methodology is applied on two industrial case studies: the knock out drum of the flare unit in the South Pars gas processing plant, Iran, and the distillation tower of the ISOM unit in BP's Texas City refinery, USA. The losses were quantified and compared with the reported losses. A good agreement was observed between calculated losses and reported losses. Subsequently, additional safety measures were suggested and the overall consequences were also assessed. The results demonstrate the innovative ability of the methodology to use loss functions as a unique way of developing an economic consequence model in order to estimate major losses.

This economic model will be useful in dynamic risk assessment and in comparing alternative safety measures. Considering both the immediate and longer-term consequences, the results of analysis using this economic methodology improve risk assessment and enhance decision-making for process design and risk management.

Acknowledgment

The authors gratefully acknowledge the financial support provided by Vale Inc., Research & Development Corporation (RDC) and the Atlantic Canada Opportunities Agency (ACOA) under the Vale Research Chair Program.

References

[1] Khan F., Abbasi S. A., Techniques and methodologies for risk analysis in chemical process industries, Journal of Loss Prevention in the Process Industries. **1998**, 11, 261–277.

[2] Crowl D. A., Louvar J. F., Chemical Process Safety: Fundamentals with Applications, Prentice Hall PTR, Upper Saddle River, New Jersey, **2002**.

[3] Lees F. P., Loss Prevention in the Process Industries, Butterworth, New York, **1996**.

[4] CCPS/AIChE, Guidelines for Chemical Process Quantitative Risk Analysis, Wiley, New York, **2000**.

[5] Pula R., Khan F., Veitch B., Amyotte P., A Grid Based Approach for Fire and Explosion Consequence Analysis, Trans IChemE, Part B, Process Safety and Environmental Protection, **2006**, 84(B2), 79–91.

[6] Greenberg M., Haas C., Cox A., Lowrie K, McComas K, North W. Ten Most Important Accomplishments in Risk Analysis 1980–2010. Risk Analysis, **2012**, 32, 5, 771-781.

[7] Jonkman S. N., van Gelder M., Vrijling J. K., An overview of quantitative risk measures for loss of life and economic damage Journal of Hazardous Materials, **2003**, A99, 1–30.

[8] Burgman M., Franklin J., Hayes K. R., Hosack G. R., Peters G. W., Sisson S. A., Modeling Extreme Risks in Ecology, Risk Analysis, **2012**, 32, 11, 1293-1308.

[9] Lijian Y., Yingquan D., Lichen S., Lijun W., zongzhi W., Method for determining probabilities of leaks in quantitative risk assessment, Journal of Safety Science and Technology, **2007**, 3, 27-30.

[10] Junior M., Santosa M., Vidala M., Carvalho P., Overcoming the blame game to learn from major accidents: A systemic analysis of an Anhydrous Ammonia leakage accident, **2012**, 25, 1, 33–39.

[11] Kar S., Roy K., QSAR modeling of toxicity of diverse organic chemicals to Daphnia magna using 2D and 3D descriptors, Journal of Hazardous Materials, **2010**, 177, 344–351.

[12] Aven T., Renn O., On the Risk Management and Risk Governance of Petroleum Operations in the Barents Sea Area, Risk Analysis, **2012**, 32, 9, 1561-1575.

[13] Si H., Ji H., Zeng X., Quantitative risk assessment model of hazardous chemicals leakage and application, Safety Science, **2012**, 50, 1452–1461.

[14] Roy R. K., A Primer on the Taguchi Method. SME, **2010**.

[15] Taguchi G., Introduction to Quality Engineering: Designing Quality into Products and Processes. Kraus, White Plains, NY, **1986**.

[16] Spiring F. A.m The Reflected Normal Loss Function, Canadian Journal of Statistics. **1993**, 21, 321-330.

[17] Ryan T. P., Statistical Methods for Quality Improvement, Wiley, New York, NY, **1989**.

[18] Phadke M. S., Quality Engineering Using Robust Design, Prentice Hall, New York, NY, **1989**.

[19] Baker T. B., Engineering Quality by Design: Interpreting the Taguchi Approach, Marcel Dekker, New York, NY, **1990**.

[20] Spiring F. A., Yeung A. S., A General Class of Loss Functions with Industrial Applications, Journal of Quality Technology, **1998**, 30, 2, 152-62.

[21] Sun F., Laramee J., Ramberg J., On Spiring's Inverted Normal Loss Function, Canadian Journal of Statistics, **1996**, 24, 241-249.

[22] Pan J N., Wang J. H., A study of loss functions for product interference analysis, Industrial Engineering Research, **2000**, 2, 1, 80-100.

[23] Leung B., Spiring F. A., Some Properties of the Family of Inverted Probability Loss Functions, Quality Technology & Quantitative Management, **2004**, 1, 1, 125–147.

[24] Pan J., Pan J., A Comparative Study of Various Loss Functions in the Economic Tolerance Design, IEEE International Conference on Management of Innovation and Technology, **2006**.

[25] NORSOK standard S-001, Technical safety, Edition 4, **2008**.

[26] Mukhopadhyay R., Review unit-wide impacts on closed-drain drums, Hydrocarbon Processing, **2011**, 90, 5, 107-110.

[27] NORSOK standard P-100, Process systems, Edition 4, **2008**.

[28] ANSI/API Standard 521, Pressure-relieving and Depressuring Systems, Fifth Edition, **2007**.

[29] National Iranian Oil Company, Report SPP-SP19-140-P-101, South Pars Gas field Development, Phase 19.

[30] Shelley C. H., Cole A. R., Markley T. E., Industrial Firefighting for Municipal Firefighters, Fire Engineering Books, **2007**.

[31] Zadakbar O., Abbassi R., Khan F., Karimpour K., Golshani M., Vatani A., Risk Analysis of Flare Flame-out Condition in a Gas Process Facility, Oil & Gas Science and Technology, **2011**, 66, 3, 705-711.

[32] TOTAL, GS EP SAF 253A: Impacted Area, Restricted Area and Fire Zones. **2005**.

[33] Manca D., Brambilla S., Dynamic simulation of the BP Texas City refinery accident, Journal of Loss Prevention in the Process Industries, **2012**, 25, 950-957.

[34] Khan F. I., Amyotte P. R., Modeling of BP Texas City refinery incident, Journal of Loss Prevention in the Process Industries. **2007**, 20, 387–395.

[35] CSB (U.S. Chemical Safety and Hazard Investigation Board), BP Texas City refinery explosion and fire Investigation report, **2007**.

[36] Broadribb M. P., Lessons from Texas City A Case History, American Institute of Chemical Engineers, 40th Annual Loss Prevention Symposium, Orlando, Florida, April **2006**.

[36] Baker Panel, The Report of the BP U.S. Refineries Independent Safety Review Panel, **2007**. (http://www.bp.com/bakerpanelreport).

[36] National News Release: USDL 05-1740, OSHA Fines BP Products North America More Than $21 Million Following Texas City Explosion, Company Agrees to Make Extensive Plant-Wide Improvements, September 22, **2005**.

Chapter 6: Dynamic Risk Assessment using Nonlinear Non-Gaussian Fault Detection and Detailed Consequence Analysis[†]

Preface

A version of this chapter has been submitted to the *Canadian Journal of Chemical Engineering*, in 2014. Zadakbar was the main lead on the work. The co-authors, Dr. Khan and Dr. Imtiaz supervised the work and helped to develop the methodology.

Abstract

This chapter presents a dynamic risk assessment using a comprehensive economic consequence methodology in combination with a multivariate model based fault detection method. The proposed approach aims to calculate process risk dynamically at each sampling instant and also to identify and screen the faults that are not hazardous. The approach relies on a particle filter combined with a comprehensive economic consequence methodology. The fault detection module uses a state space model of the process plant and a particle filter algorithm that calculates the probability of the fault. The output of this module is then combined with the consequence module, which uses loss functions to relate process deviations to economic losses. The consequence module identifies, quantifies, and integrates losses for a given scenario. Combining the two modules for risk assessment makes this approach more reliable in the analysis of

[†] *O. Zadakbar, F. khan, S. Imtiaz, Dynamic risk assessment using non-linear non-Gaussian fault detection and detailed consequence analysis, Canadian Journal of Chemical Engineering, Submitted, 2014.*

realistic nonlinear process systems and improves decision-making for process design and risk management.

KEYWORDS

Dynamic Risk Assessment, Safety System, Fault Detection, Particle Filtering, Loss Functions

6.1 INTRODUCTION

The complexity of modern plants leaves processes more susceptible to unexpected failures that significantly affect safety, economics, and the environment [1], and faults must be detected and isolated as quickly as possible in order to maintain safety and reliability standards in process plants [2].

Over the last two decades, an impressive body of work has been devoted to fault detection, and various techniques have been proposed [3-12].

Fault detection methods based on models and on historical data have been used in chemical process plants. Although more success has been observed with multivariable fault detection techniques [11-16], none of the new approaches considers the consequences of fault on the process, equipment, or environment.

A risk-based fault detection method was earlier proposed by Bao *et al.* [17] to address safety concerns that were overlooked by other fault detection methods. The risk-based fault detection method issues a warning not when residuals or signals cross the threshold, but when the risk of a process exceeds the acceptable threshold. This is an important concept, because it eliminates non-hazardous faults and also gives a real time indication of the operational risk. Bao *et al.* [17] combined a univariate charting method with risk calculation. They calculated the probability of fault based on the deviation of a signal from the threshold. Univariate methods do not take into account changes in the input. Therefore, they are less able to distinguish between process faults and operational changes. This potentially limits the effectiveness of risk-based methods involving a univariate fault detection and diagnosis approach.

Zadakbar *et al.* [15] extended the method to a multivariate risk-based fault detection and diagnosis technique by combining a principal component analysis (PCA) with the risk assessment procedure. A principal component score was used as the main indicator for fault. The probability of fault was calculated based on the filtered score, and the severity of the fault was a weighted average of the consequences of each variable in the score, with the weights coming from the loadings of the principal component. Subsequently, the severity and probability of the fault were combined to permit the identification of risk at any sampling instant.

Zadakbar *et al.* [15-16] also developed another multivariable risk-based fault detection technique using a process model: a residual generation process based on the Kalman filter was combined with a risk assessment procedure. Residuals generated from a Kalman filter were used to calculate the probability of fault. The method proposed took advantage of the correlation between process input and output and was therefore better equipped to detect faults and calculate risk precisely. These methods, though successful in detecting and diagnosing faults, have limited use in real industrial processes, as they relied on linear models and assumed Gaussian noise, which is a common assumption in many fault detection and diagnosis methods [1,18]. Also, the consequence analysis of the fault was based on a relative scale, so the full impact of the method was not realized.

State estimations of nonlinear systems have been considered in the literature using extended Kalman filters (EKF) and unscented Kalman filters (UKF). These filters, however, assume Gaussian noise. The state estimations are often not satisfactory and lead to a high rate of false alarms [1, 19]. On the other hand, the particle filter (PF) is more general and can estimate states of nonlinear non-Gaussian systems [20].

In this Chapter, fault detection using particle filtering is combined with a newly developed economic consequence analysis methodology. This methodology makes the severity of the faults more meaningful by quantifying consequences in economic terms. This new economic consequence methodology helps to integrate real time process states to accident scenarios through loss functions. This methodology, along with the proposed loss function model, can quantify risk more accurately and, more importantly, improve decision-making about risk management.

6.1.1 Organization of the chapter

The remainder of the chapter is organized as follows. Section 6.2 introduces the risk-based fault detection method. This method combines particle filtering and risk assessment based on economic consequence methodology. The consequence analysis assigns loss functions to each type of loss. In Section 6.3, the proposed method is applied to two process systems. The first case study describes a dynamic risk assessment of a CSTR, while the second case study analyzes the explosion in a fluid catalytic cracker (FCC) unit. The Chapter ends with concluding remarks in Section 6.4.

6.2 Methodology

The algorithm of risk-based fault detection using particle filtering and economic consequence methodology is given in Figure 6.1 (Start time: k=0). The first step is to model the process to determine all measured and unmeasured variables. In the next step, the particle filter estimates all process states. After that, estimated state data is used for prediction of states for n steps ahead. In the next step, the probability of fault at each predicted point is calculated using equations proposed by Zadakbar *et al.* [15-17, 21]. In parallel, the economic consequence analysis was conducted in real time. It starts by defining the possible fault scenarios; it identifies different classes and subclasses of loss, quantifies each loss using an appropriate loss function, and finally integrates all losses to determine the overall loss. In the next step, the risk of each predicted point is calculated using probability and the severity of the fault. Finally, the risk profile is used to detect the fault and take necessary supervisory decisions to implement the appropriate safety measures.

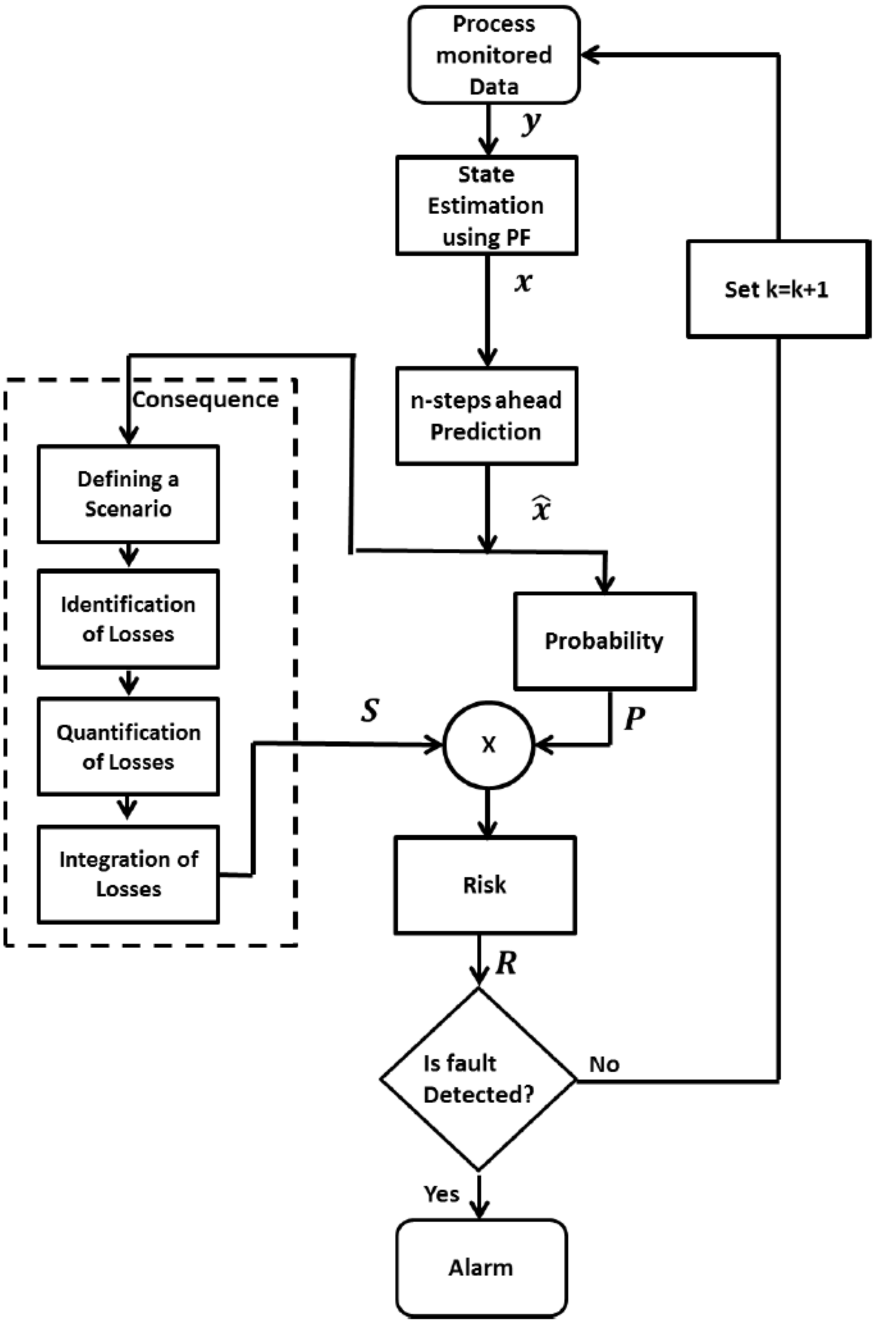

Figure 6.1. The algorithm of risk-based fault detection using particle filtering and economic consequence methodology

6.2.1 Fault Detection using Particle Filtering

Monte Carlo techniques for non-linear non-Gaussian state estimation have been proposed by Handschin [22]. Recently, the particle filter [23-24], an extension of the ideas in Handschin's paper [22], has attracted much attention [25-26].

The particle filter is a sequential Monte Carlo method based on point mass representations of probability densities. Several variants of the particle filters, such as Sequential Importance Re-Sampling (SIR), Auxiliary Sampling Importance Resampling (ASIR), and the Regularized Particle Filter (RPF), are the basis for most particle filters that have been developed [26-27].

We consider the following evolution of the state sequence of $(x_k, k \in \mathbb{N})$:

$$x_k = f_k(x_{k-1}, u_{k-1}, w_{k-1}) \tag{6.1}$$

$$y_k = h_k(x_{k-1}, v_{k-1}) \tag{6.2}$$

where k is the time index; x_k is the state vector; y_k is the measurement vector; $u_k$is the noise-free input; w_k is process noise representing disturbances, all un-modeled dynamics, and any mismatch between the process and states; and v_k is measurement noise representing inaccuracy in measuring.

From a Bayesian perspective, to estimate the states of nonlinear systems using a particle filter, one needs to recursively calculate some degree of belief in the state x_k at time k, given the data $y_{1:k}$ up to time k. Thus, the probability density function (PDF), $p(x_k| y_{1:k})$, needs to be constructed. It can be recursively obtained in two steps of prediction and update, assuming the initial condition $p(x_0| y_0) \equiv p(x_0)$ is known. In the first step, the prediction step, the Chapman–Kolmogorov equation [19] is used to calculate the prior PDF $p(x_k|\, y_{1:k-1})$, assuming the system model, $p(x_k|x_{k-1})$, and PDF $p(x_{k-1}|\, y_{1:k-1})$ are known at time $k-1$:

$$p(x_k|\, y_{1:k-1}) = \int p(x_k|\, x_{k-1})\, p(x_{k-1}|\, y_{1:k-1}) dx_{k-1} \tag{6.3}$$

Equation 1 represents a first order Markov process and there is only one past outcome for each transition; thus, $p(x_k|x_{k-1}, y_{1:k-1}) = p(x_k|\, y_{1:k-1})$.

In the second step, the update step, the measurement y_k obtained at the time k is used to update the prior PDF using Bayes rule:

$$p(x_k|\, y_{1:k}) = \frac{p(y_k|x_k)p(x_k|\, y_{1:k-1})}{p(y_k|\, y_{1:k-1})} \tag{6.4}$$

$$p(y_k|\, y_{1:k-1}) = \int p(y_k|\, x_k)\, p(x_k|\, y_{1:k-1}) dx_k \tag{6.5}$$

Equations 6.1 and 6.2 are used to define the PDF $p(x_k|\, y_{k-1})$ and the pdf $p(y_k|x_k)$ respectively.

To obtain the optimal Bayesian solution for the non-linear non-Gaussian systems, particle filters represent the required posterior density function through a set of random samples and their associated weights. An equivalent representation of the posterior PDF $p(x_{0:k}|\, y_{1:k})$ can be obtained if the number of samples becomes large enough; thus, the particle filter approaches the optimal Bayesian solution [20].

Consider that $\{x_{0:k}^i,\ W_k^i\}_{i=1}^{Ns}$ represents a random set characterizing the posterior PDF $p(x_{0:k}|\, y_{0:k})$, where $x_{0:k}^i$ represents a set of support points with associated weights of W_k^i, and $x_{0:k} = \{x_j\}_{j=1}^k$ represents the states up to time k. The following equation provides a discrete weighted approximation of the posterior PDF $p(x_{0:k}|\, y_{0:k})$,

$$p(x_{0:k} | y_{1:k}) \approx \sum_{i=1}^{N_s} W_k^i \delta (x_{0:k} - x_{0:k}^i) \tag{6.6}$$

where δ is the Dirac measure. The weights can be obtained based on the principle of importance sampling,

$$W_k^i = p(y_k, x_k^i) W_{k-1}^i \tag{6.7}$$

Thus, this SIS algorithm propagates the weights and particles recursively as each measurement is received sequentially.

A re-sampling step, involving generating a new set $\{x_{0:k}^{i*}\}_{i=1}^{Ns}$ by resampling and replacement N_s times, is added to the algorithm. In this case, the sample is independent and distributed identically from the discrete density $p(x_k | y_{1:k})$. After re-sampling, the weights of the particles are reset to $1/N_s$. This is an optional step to avoid collapsing of particles to one point [20, 27-28].

At any given time k, the weights W_k^i are calculated using the likelihood function $p(y_k | x_k)$, and this information is transferred to the samples by re-sampling and progressed to the next time, $k + 1$. Systematic re-sampling is easy to implement and minimizes the Monte Carlo variation [24]. Figure 6.2 shows the schematic diagram explaining the implementation of the SIR algorithm.

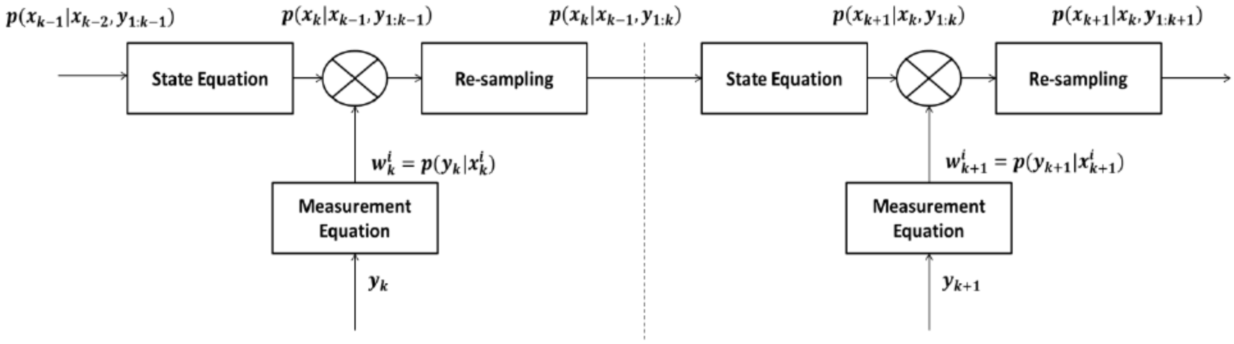

Figure 6.2. Schematic diagram explaining the implementation of the SIR algorithm [28]

The estimated states from particle filtering are used for fault detection. Finally, the particle filter is used to make a four step ahead prediction at any given time k. The filtered predicted signal is used for fault detection and dynamic risk assessment.

6.2.2 Dynamic Risk Assessment

Risk is defined as a measure of likely harm or economic loss caused by the fault if a corrective action is not taken. Risk, R, depends on the probability of occurrence of a fault leading to an unwanted event, P, and the severity of losses associated with that event, S [16]:

$$R = P \times S \tag{6.8}$$

6.2.2.1 Probability of Fault

The probability of a fault is calculated using the following equations by Zadakbar *et al.* [16]:

$$for\ x_{mk} > \mu_m \dashrightarrow P = \varphi\left(\frac{x_{mk}-(\mu_m+3\sigma_m)}{\sigma_m}\right) = \int_{-\infty}^{x_{mk}} \frac{1}{\sqrt{2\pi}\sigma_m} e^{\frac{(x_{mk}-(\mu_m+3\sigma_m))^2}{2\sigma_m^2}} dx \tag{6.9}$$

$$for\ x_{mk} < \mu_m \longrightarrow\ P = 1 - \varphi\left(\frac{x_{mk}-(\mu_m-3\sigma_m)}{\sigma_m}\right) = 1 - \int_{-\infty}^{x_{mk}} \frac{1}{\sqrt{2\pi}\sigma_m} e^{\frac{(x_{mk}-(\mu_m-3\sigma_m))^2}{2\sigma_m{}^2}} dx \quad (6.10)$$

where $m = 1, \cdots, n$ and n is the number of process states, and μ and σ represent mean and variance respectively.

6.2.2.2 Consequence Analysis

The severity of fault is calculated with the economic consequence methodology for process risk analysis proposed by Zadakbar *et al.* [16-17]. In this methodology, the process deviations that contribute to a given accident scenario are identified and mapped to assess potential consequences. Losses are assessed with an appropriate loss function (revised Taguchi, modified inverted normal) for each type of loss. The total loss is quantified by integrating different loss functions.

The first step of this methodology is the selection of a process unit. For each process variable in this unit, all potential deviations, impacts, and consequences, such as process interruption, release, fire, and explosion are identified. The next step is to identify all potential losses, such as material and equipment losses, injury, and fines. To quantify these losses, an appropriate loss function (Taguchi, inverted normal) or a step function is assigned based on the sensitivity of losses to the deviation of the selected process variable. Assigned functions for each type of loss are integrated to quantify the total loss [21].

6.3 Case Studies

The application of the proposed methodology is demonstrated on two process systems.

6.3.1 CSTR

For this case study, a non-linear model of the CSTR presented by Zadakbar *et al.* [16-17] was developed. The reaction in the system is the oxidation of sodium thiosulfate by hydrogen peroxide, which is an irreversible exothermic oxido-reduction reaction [29]. The concentration of sodium thiosulfate, C_A, the temperature in the reactor, T_R, and the temperature in the coolant jacket, T_J, are selected as the elements of process state matrix. It is assumed that there are two sensors to measure the concentration and the jacket temperature. The control variables of the process are feed to the reactor, F, and cooling water flow rate, F_W. Both variables could be manipulated. The feed concentration and cooling water inlet temperature T_{Jin} are uncontrolled inputs in the process; thus, they may be considered as disturbances.

The reactor temperature is modeled as an unmeasured process state using a particle filter algorithm. Figure 6.3 shows the predicted estimation of reactor temperature.

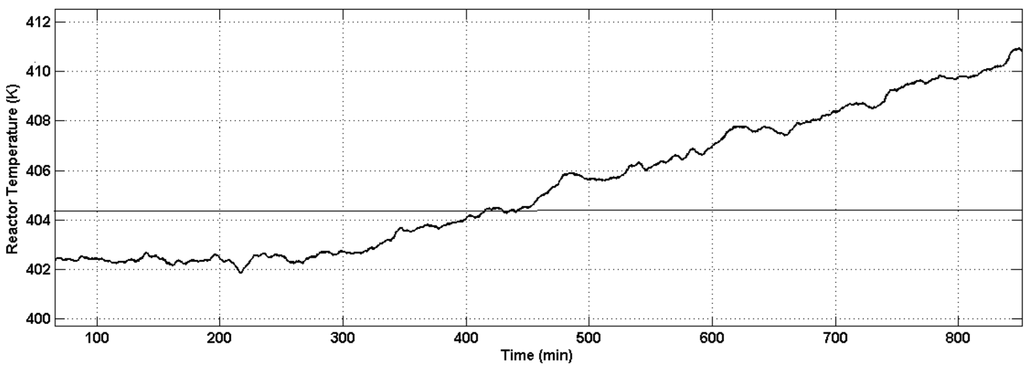

Figure 6.3. The filtered predicted estimation of reactor temperature.

This filtered signal is used to calculate the probability of the fault using Equations 6.9 and 6.10 proposed by Zadakbar *et al.* [21]. Figure 6.4 shows the probability of the fault based on the unmeasured reactor temperature signal.

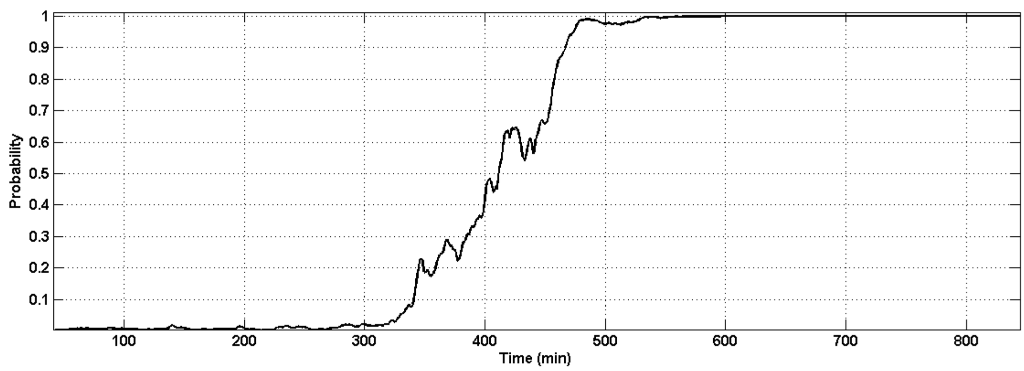

Figure 6.4. The probability of fault based on an unmeasured reactor temperature signal

The next step is consequence analysis. The scenario for this case study is increasing reactor temperature leading to overpressure and explosion. Significant types of losses due to explosion are identified as equipment loss and injuries. The overall loss is determined through a combination of all losses. The losses are calculated using economic consequence models that are functions of process deviations.

Because of the high sensitivity of equipment to temperature, an inverted normal loss function is used to model equipment losses, assuming the maximum loss is $30,000. Figure 6.5 shows equipment losses.

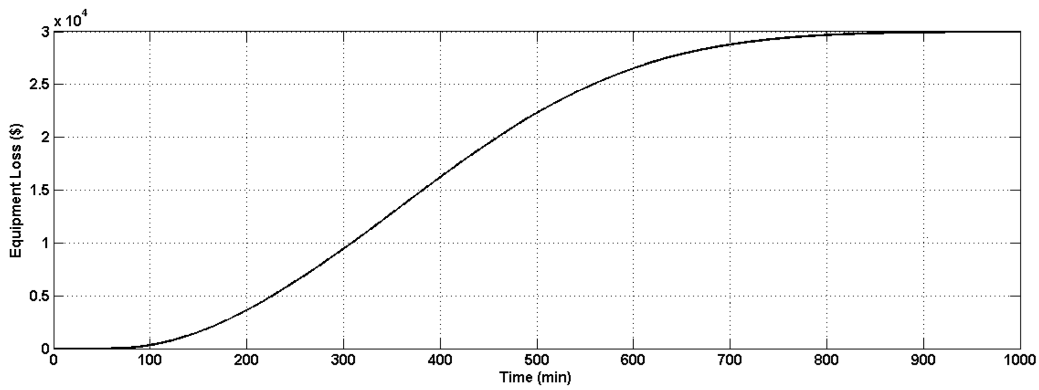

Figure 6.5. Equipment losses based on reactor temperature

In this case, assuming one operator works with the reactor, the step function assigned to the injury consists of one step for individual major non-recoverable injuries. Figure 6.6 shows injury losses resulting from uncontrolled reactor temperature.

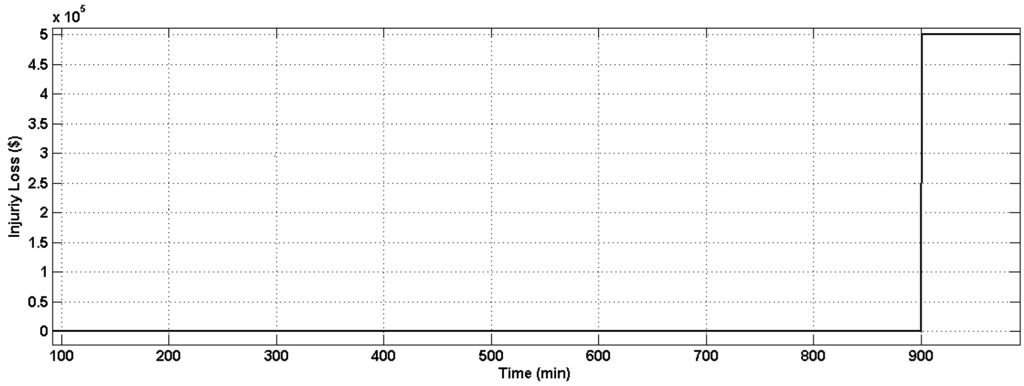

Figure 6.6. Injury loss based on reactor temperature

All losses can be integrated to calculate the total losses. Figure 6.7 shows the total losses due to an explosion in the CSTR.

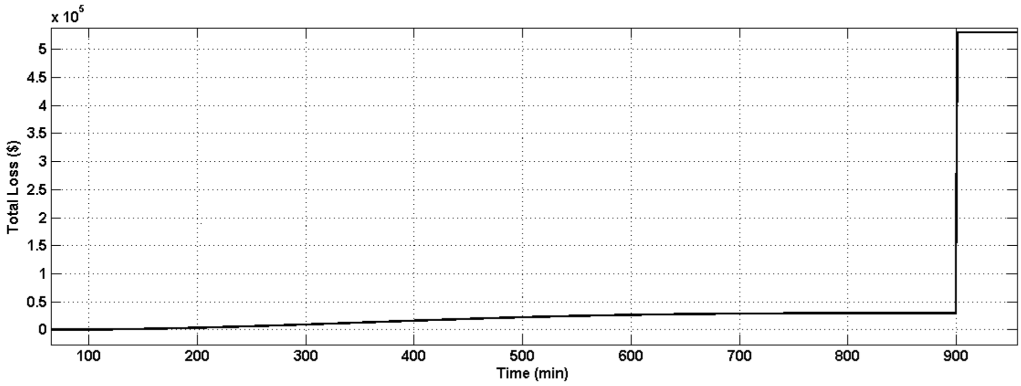

Figure 6.7. Total loss based on reactor temperature for the explosion scenario

A risk profile is generated by multiplying the probability of occurrence of the fault at any particular instant by its economic consequences. Figure 6.8 shows the risk profile for increasing reactor temperature leading to overpressure and explosion.

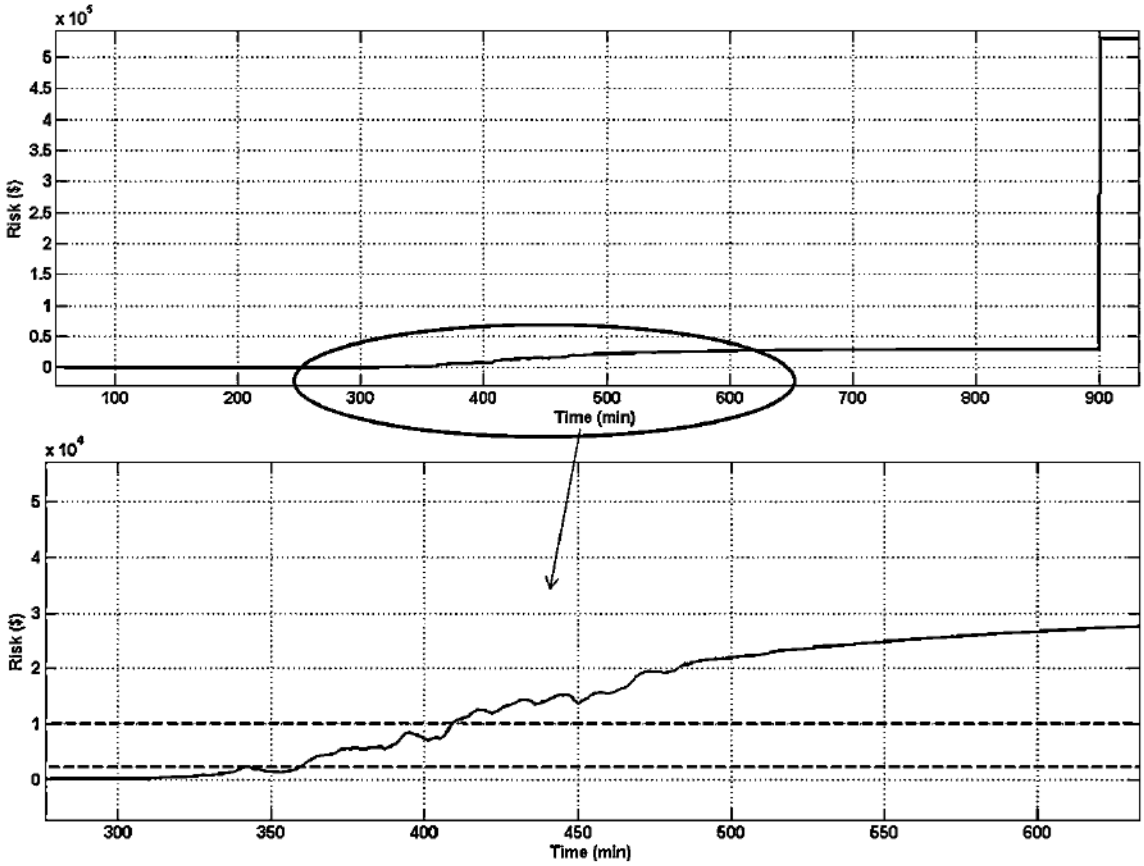

Figure 6.8. Risk profile based on increasing the reactor temperature

Fault detection using this risk profile can prioritize safety measures and recovery actions and therefore improve decision making for risk management. The criteria for acceptable economic risk in each scenario could determine the risk thresholds. Fault can be detected at t=410 min with a threshold of $1,000. In this case, if the risk profile crosses the $10,000 threshold, the system can be shut down to prevent the accident.

6.3.2 Fluid Catalytic Cracking Unit (FCCU)

This case study considers the fluid catalytic cracking unit (FCCU) shown with its PI regulatory controllers in Figure 6.9. The inlet stream is a mixture of preheated fresh feed and hot slurry recycle from the bottom of the fractionator. This mixture enters the reactor riser, where the cracking reactions take place, to mix with the hot regenerated catalyst. As a result of the reactions, coke is deposited on the surface of the catalyst. Spent catalyst is transported to the regenerator through the spent catalyst U-bend. Gaseous product from the reactor is passed to the fractionator to separate into various product streams [30]. Two typical faults in the FCCU are low quality and yield (related to low conversion, high dry gas production, low production of gasoline, and low octane in gasoline) and high temperature in the riser and the regenerator [31-32].

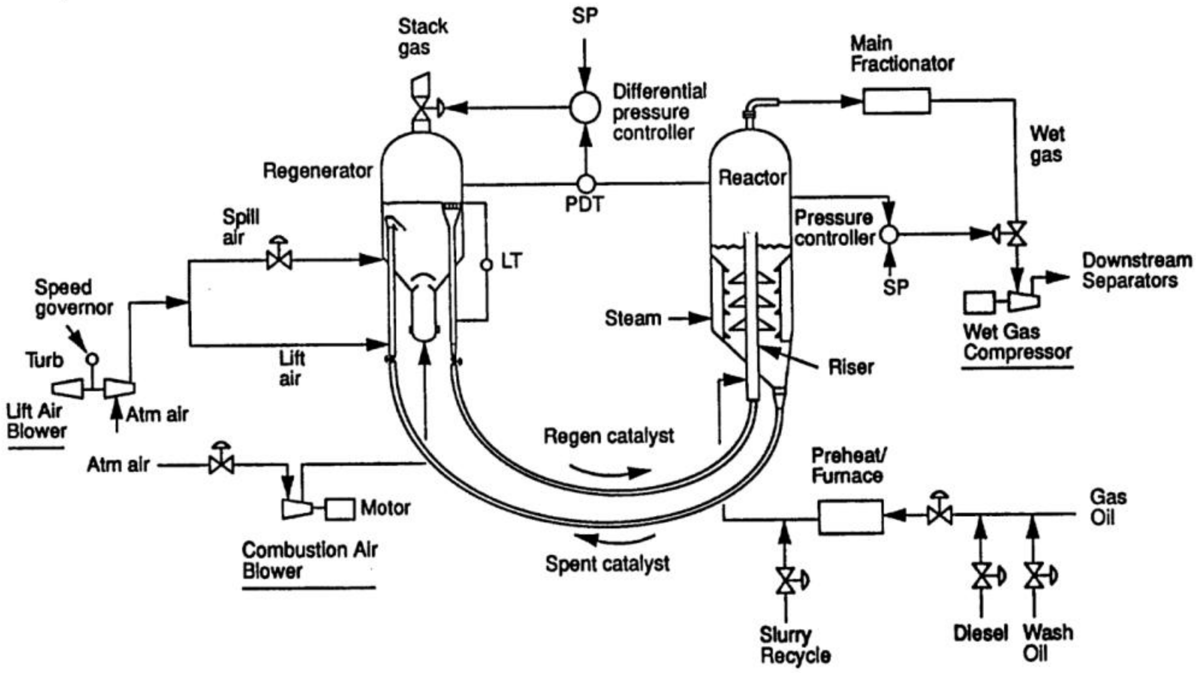

Figure 6.9. The schematic of the FCCU [30]

6.3.2.1 Particle filtering

A non-linear state space model developed by McFarlane *et al.* [30] dynamically simulates this unit. A particle filtering algorithm based on sequential importance re-sampling (SIR) is used for state estimation and fault detection. Table 6.1 shows process states and measurements.

Table 6.1. Process states and measured variables

Variables	States	Measurements
y(1)=Reactor Pressure (P4, psi)	•	•
y(2)=Differential Pressure (DP, psi)	•	
y(3)=Air Flow Rate (Fair, mole/s)	•	•
y(4)=Regenerator Pressure (P6, psi)	•	•
y(5)=Furnace Temperature (T3, F)	•	•
y(6)=Fresh Feed Temperature (T2, F)	•	•
y(7)=Riser Temperature (Tr, F)	•	•
y(8)=Regenerator Temperature (Treg, F)	•	•
y(9)=Spent Catalyst level (splev, ft)	•	
y(10)=Cyclone Temperature (Tcyc, F)	•	•
y(11)=Differential Cyclone temperature (DT, F)	•	
y(12)=Carbon monoxide concentration (XCo, ppm)	•	
y(13)=Oxygen concentration (XO, ppm)	•	
y(14)=Coke wt fraction in spent catalyst (Csc)	•	
y(15)=Coke wt fraction in regenerated cat.(Crgc)	•	
y(16)=Air blower flow inlet-surge	•	•
y(17)=Wet compressor inlet suction flow	•	•
y(18)=Combustion air suction flow	•	•
y(19)=Combustion air suction pressure (P1, psi)	•	•
y(20)=Combustion air discharge pressure (P2, psi)	•	•

A gradual increase in the composition of heavy components in the feed to the FCCU is introduced as a fault at t=50min. This fault affects several process states, but it is more rapidly detected by monitoring the unmeasured state of coke weight fraction in the spent catalyst [31]. Figure 6.10 shows the estimated state using particle filtering.

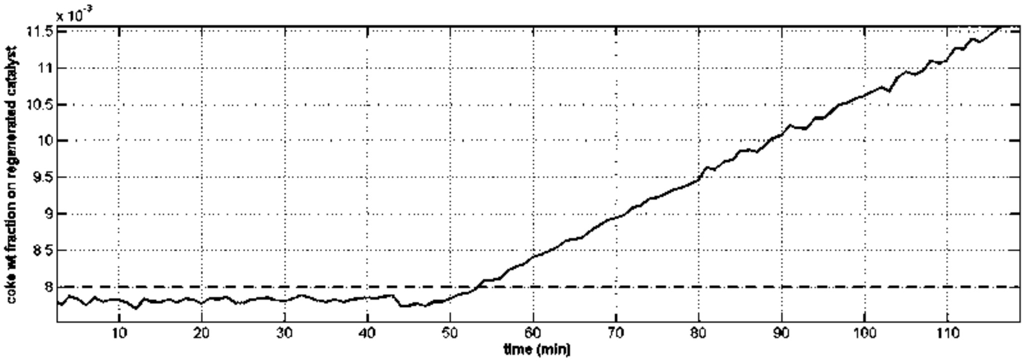

Figure 6.10. Filtered prediction of the unmeasured state of coke weight fraction in spent catalyst

6.3.2.2 Risk Assessment

An increase in the composition of heavy components in the feed rapidly increases coke deposition in the riser and the concentration of coke on spent catalyst, causing a higher combustion rate in the regenerator and an increase in regenerator temperature. Since this increased combustion rate is not enough to burn off all coke on the catalyst, the amount of carbon on regenerated catalysts increases as well. Because of the constant air flow, the concentration of oxygen decreases and the stack carbon monoxide concentration increases. The combustion of additional carbon monoxide raises temperatures in the reactor, the regenerator, and the riser; thus, wet gas production increases. The reactor pressure controller opens the wet gas suction valve. However, when the valve saturates, pressure in the reactor starts to increase. If the amount of coke on spent catalyst continues to grow, the temperature in the riser can reach its maximum metallurgical limit of 995°F. Over-pressurization could lead to leaks, release, and a vapor cloud explosion. Figure 6.11 shows the probability of the fault, calculated using Equations 6.9 and 6.10.

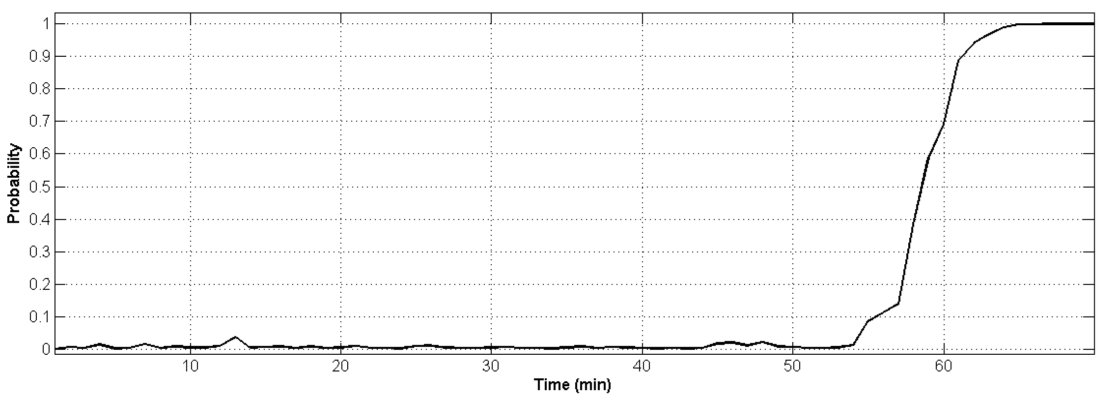

Figure 6.11. Probability of the fault based on coke weight fraction in spent catalyst

The scenario for this case study consists of following steps: an increasing coke weight fraction in spent catalyst leading to overpressure, release, and ultimately a vapor cloud explosion. The significant types of losses due to explosion are material loss, equipment loss, injuries and fatalities, fines, and claims. Each loss is modeled by a loss function. The overall loss is determined through a combination of all loss functions, which provides a comprehensive economic consequence model based on process deviations.

Some data from the report of explosion in FCC Unit, La Mède, France, Total, 1992 [33] is used for the implementation of the economic consequence analysis. Based on the results of the overpressure modeling, all equipment within about two hectares of this refinery, where the overpressure is more than 150 mbar, could be severely damaged. Due to the high sensitivity of equipment losses, an inverted normal loss function is used to model equipment losses. In this case, the maximum loss is $30,000,000. Figure 6.12 shows equipment losses due to the explosion.

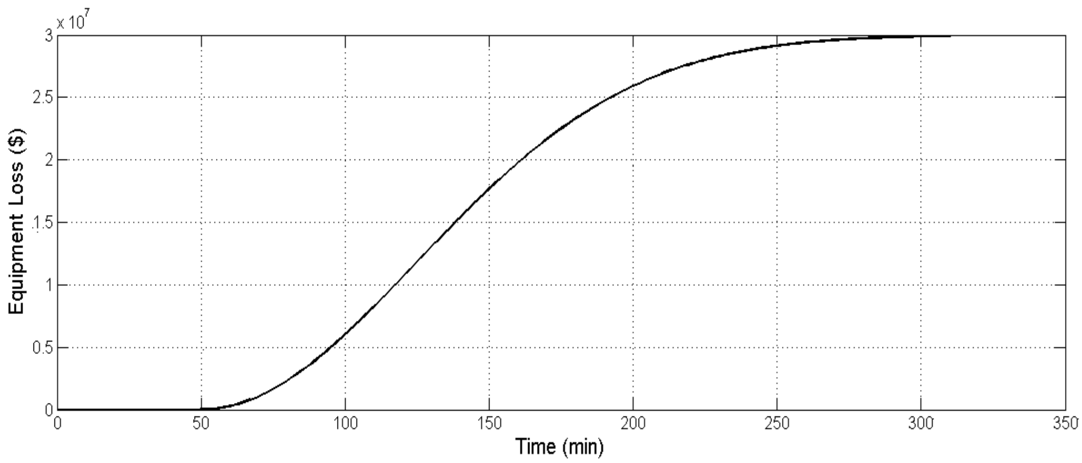

Figure 6.12. Equipment losses due to explosion in the FCCU

The number of reported injuries was 38. Damage to property was reported in the town of Martigues, 4.5 km away. In this case study, the maximum fine was $2,000,000. Fines due to the explosion are modeled using the Taguchi loss function. The step function assigned to the civil claims and juries consists of three steps: step 1 includes claims for individual minor injuries; step 2 includes individual major non-recoverable injuries, individual disabilities, and minor small group injuries (involving fewer than 10 people); and phase 3 includes significant losses and large group injuries (involving more than 10 people). Figures 6.13, 6.14, and 6.15 show losses due to fatalities, injuries and civil claims, and fines.

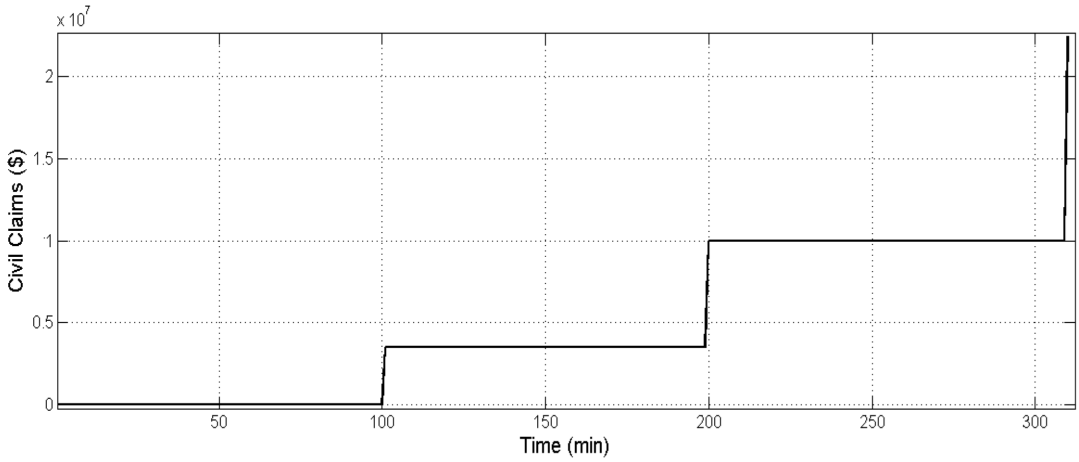

Figure 6.13. Economic losses due to injuries and civil claims

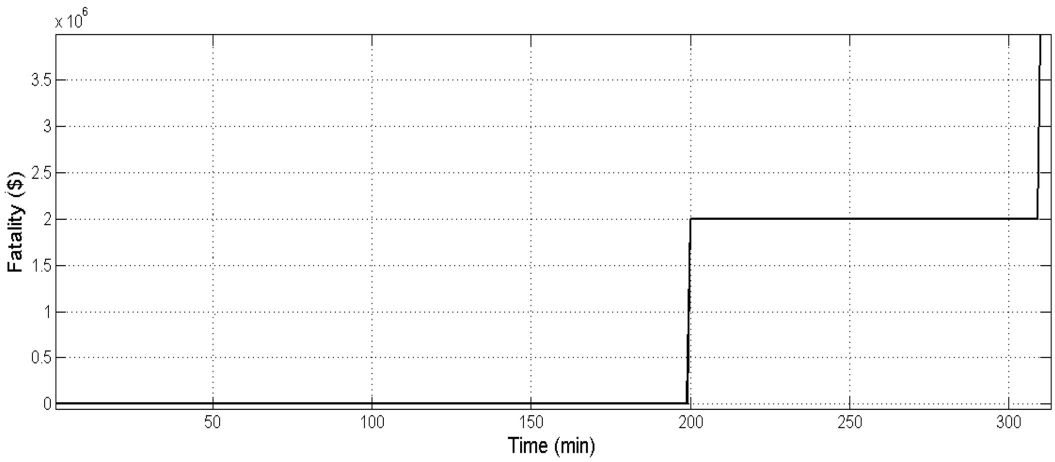

Figure 6.14. Economic losses due to fatalities

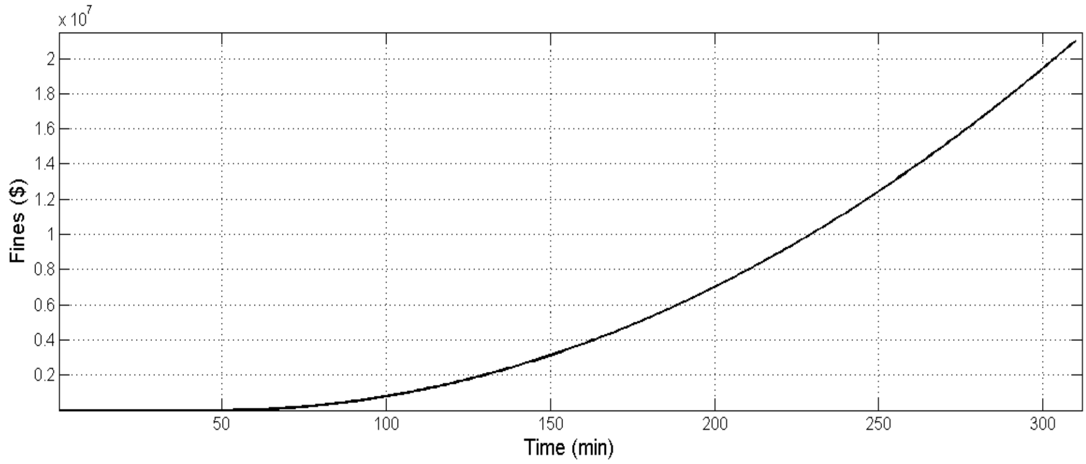

Figure 6.15. Economic losses due to fines

The FCCU and surrounding process units are severely damaged by the explosion, resulting in the whole refinery being shut down for six months. Material losses can be modeled using an inverted normal loss function, with a maximum loss of $180,000,000 (Figure 6.16).

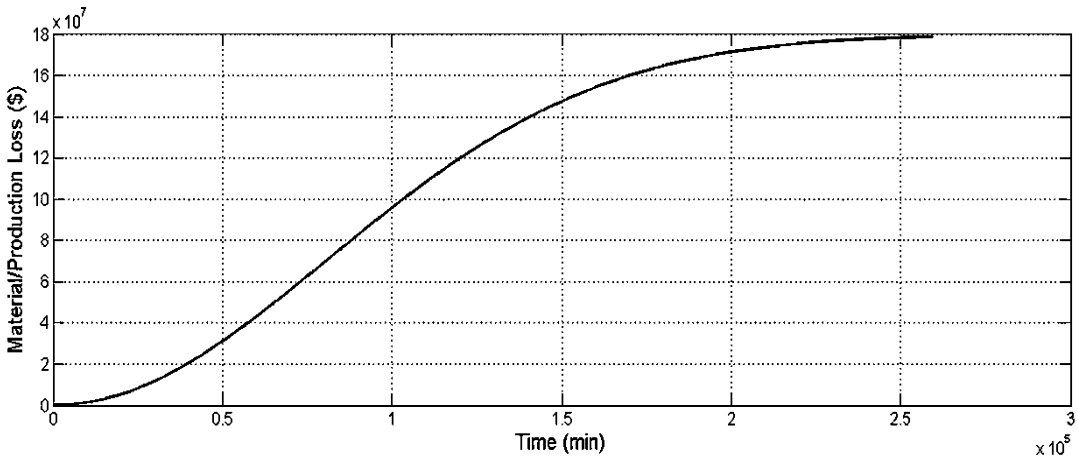

Figure 6.16. Economic losses due to material losses

Thus, all losses could be integrated to calculate the total loss. Figure 6.17 shows the total loss due to the explosion in the FCCU.

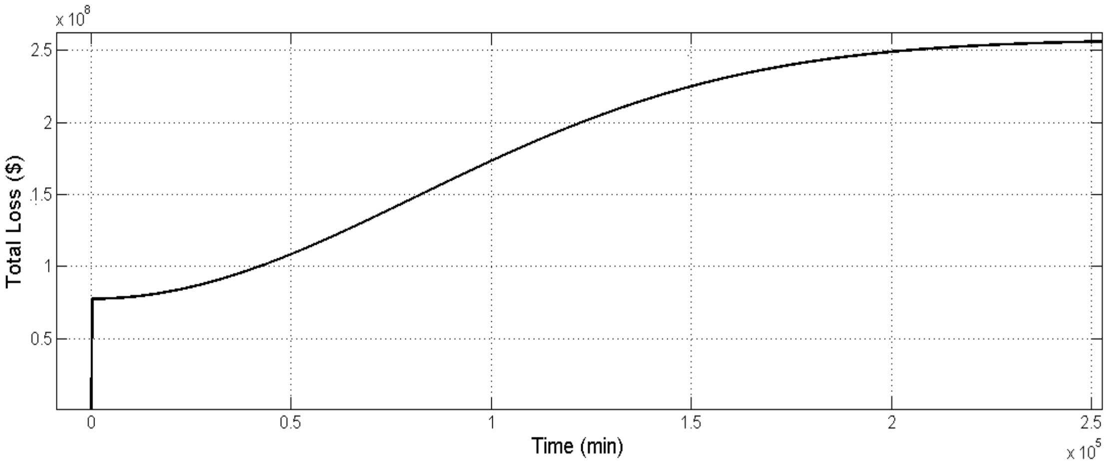

Figure 6.17. Total loss due to the explosion in the FCCU

The risk profile is now generated by combining the probability of occurrence of the fault and its economic consequences. Figure 6.18 shows the risk profile for increased weight fraction of coke on spent catalyst leading to overpressure, release, and explosion.

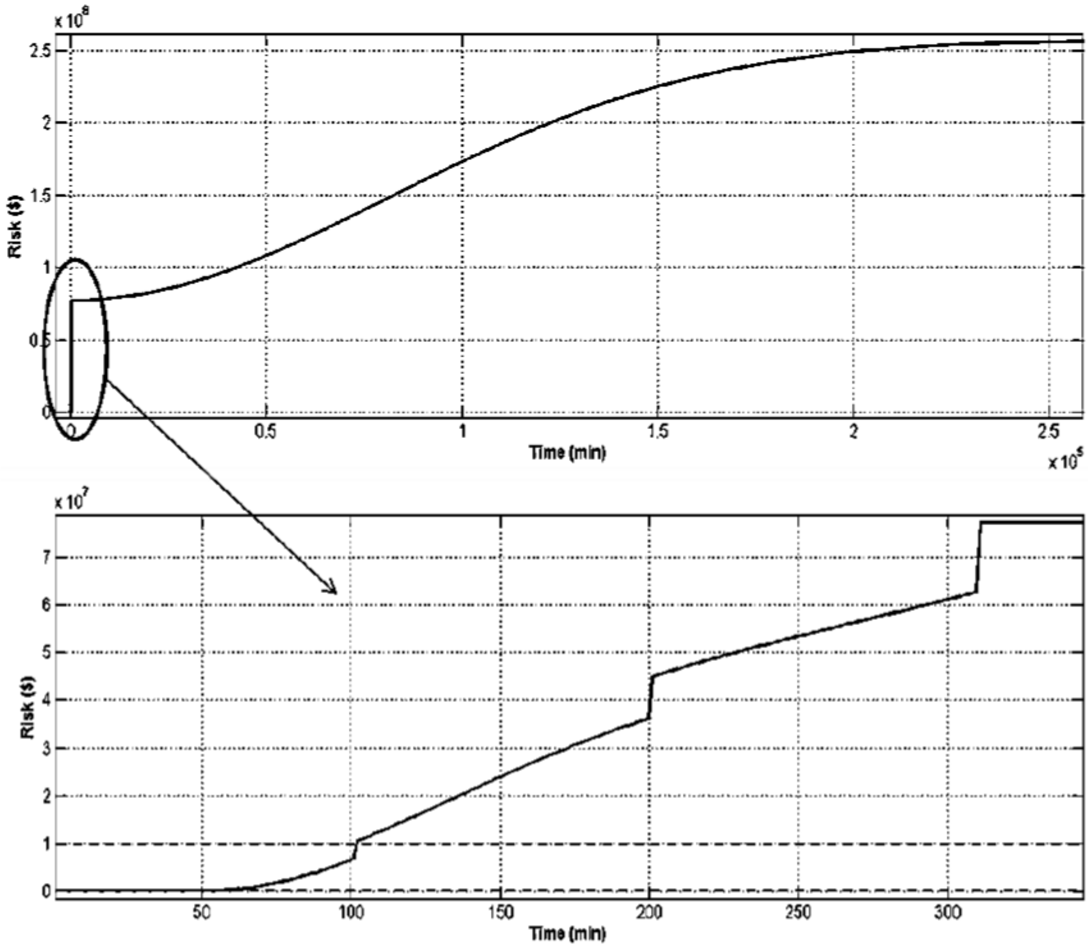

Figure 6.18. Risk profile based on increasing the weight fraction of coke on spent catalyst leading to explosion

When risk profiles are based on an economic model that considers both the immediate and longer-term consequences (such as material losses), fault detection can aid decision-making for risk management by enhancing the prioritization of safety measures and recovery actions based on the risk of each process state. The risk thresholds could be selected based on the criteria for the accepted economic risk in each scenario. In this case, fault is detected at t=53 min if the risk profile crosses the threshold of $1 M; if risk exceeds the threshold of $10 M, a safety instrumented system might be activated to reduce total fresh feed flowrate. This would help bring the riser temperature back to its constraint limit.

6.4 Conclusions

A Particle filter based fault detection has been extended to a risk-based fault detection framework targeting the safety issues of a process system. Combining a particle filter with a risk calculation procedure that uses loss functions to quantify major losses (such as material loss, equipment loss, fatalities and injuries, fines, and claims faced in a process industry) makes this method more realistic for real-industrial implementation. Risk-based fault detection based on a combination of PF with a detailed consequence analysis offers significant advantages:

- It allows monitoring in non-linear non-Gaussian systems.
- It uses different economic consequence models that are properly assigned to represent different types of loss.
- It gives real-time estimation of the loss.
- It is more reliable in analyzing realistic nonlinear process systems.
- It enhances clarifies decision-making for risk management by enhancing prioritization of safety measures and recovery actions.
- It has been demonstrated on the CSTR and FCCU case studies.

Acknowledgements

The authors gratefully acknowledge the financial support provided by Vale Inc., Research and Development Corporation of Newfoundland (RDC), and the Atlantic Canada Opportunities Agency (ACOA) under the Vale Research Chair Program. The authors also acknowledge the financial support from NSERC through a collaborative research grant.

References

[1] Alrowaie F., Gopaluni R.B., Kwok K.E., Fault detection and isolation in stochastic non-linear state-space models using particle filters, Control Engineering Practice, **2012**, 20, 10, , 1016–1032.

[2] Basseville M., Detecting changes in signals and systems - A survey. Automatica, **1988**, 24, 3, 309-326

[3] Willsky A. S., Detection of Abrupt Changes in Signals and Dynamical Systems, Lecture Notes in Control and Information Sciences **1986**, 77, 27-49

[4] Sorsa T., Koivo H. N., Neural Networks in Process Fault Diagnosis, IEEE Transactions on Systems, Man, and Cybernetics, **1991**, 21, 4.

[5] Maki Y., Loparo K., A neural-network ... IEEE Transactions on Control Systems Technology, **1997**, 6, 5, 529-541

[6] Isermann R., Balle P., Trends in the application of model-based fault detection, Control Engineering Practice, **1997**, 5, 5, 709–719.

[7] Isermann R., Integration of fault-detection and diagnosis methods, IFAC Symposium Control Engineering Practice, **1997**, 5, 5, 639-652.

[8] Chiang L. H., Braatz R. D., Russell E. L., Fault detection and diagnosis in industrial systems, Springer, **2001**.

[9] Venkatasubramanian V., Raghunathan R., Kewen Y., Kavuri S. N., A review of process fault detection and diagnosis Part I: Quantitative model-based methods, Computers and Chemical Engineering, **2003**, 27, 3, 293-311

[10] Venkatasubramanian V., Raghunathan R., Kewen Y., Kavuri S. N., Review of Process Fault Detection and Diagnosis - Part II: Qualitative Models and Search Strategies, Computers and Chemical Engineering, **2003**, 27, 3, 313-326.

[11] Venkatasubramanian V., Raghunathan R., Kewen Y., Kavuri S. N., Review of Process Fault Detection and Diagnosis - Part III: Process History Based Methods, Computers and Chemical Engineering, **2003**, 27, 3, 327-346.

[12] Isermann R., Model-based fault-detection Annual Reviews in Control **2005**, 29, 1, 71–85.

[13] Isermann R., Fault Detection Methods for the Supervision of Technical Processes, Process Automation, **1981**, 1, 36-44.

[14] lsermann R., Process Fault Detection Based on Modeling and Estimation Methods a Survey, **1984**, Automatica 20, 387.

[15] Zadakbar O., Imtiaz S., Khan F., Dynamic Risk Assessment and Fault Detection Using Principal Component Analysis, **2013**, 52, 2, 547-990.

[16] Zadakbar O., Imtiaz S., Khan F., Dynamic Risk Assessment and Fault Detection Using a Multivariate Technique, **2013**, 32, 4, 365-375

[17] Khan F., Bao H., Iqbal T., Chang Y., Risk-based Fault Diagnosis and Safety Management for Process Systems, Process Safety Progress, **2011**, 30, 6.

[18] Frank P. M., Ding X., Survey of robust residual generation and evaluation methods in observer-based fault detection systems. Journal of Process Control, **1997**, 7, 6, 403–424.

[19] Alrowaie F., Gopaluni R. B., Kwok K. E., An Algorithm for Fault Detection in Stochastic Non-Linear State-Space Models Using Particle Filters, Advanced Control of Industrial Processes (ADCONIP), **2011**, 60-65.

[20] Arulampalam M. S., Maskell S., Gordon N., Clapp T., A Tutorial on Particle Filters for Online Non-linear/ Non-Gaussian Bayesian Tracking, IEEE Transaction on Signal Processing, **2002**, 50, 2, 174-188.

[21] Zadakbar O., Imtiaz S., Khan F., Development of Economic Consequence Methodology for Process Risk Analysis, Risk Analysis, In Press, **2013.**

[22] Handschin J. E., Mayne D. Q., Monte Carlo Techniques to Estimate the Conditional Expectation in Multistage Nonlinear Filtering, International Journal of Control, **1966**, 9, 5, 547-559

[23] Gordon N. J., Salmond D. J., Smith A. F. M., A novel approach to nonlinear and non-Gaussian Bayesian state estimation, IEE Proceedings-F: Radar, Sonar and Navigation, **1993**, 140, 107-113.

[24] Kitagawa G., Monte Carlo Filter and Smoother for Non-Gaussian Non-Linear State Space Models, Journal of Computer Graphics and Statistics, **1996**, 5, 1, 1–25.

[25] Kitagawa G., A Self-Organizing State-Space Model, Journal of the American Statistical, **1998**, 93, 443, 1203–1215.

[26] A. Doucet, J. F. G. Freitas, N. J. Gordon, An introduction to sequential Monte Carlo methods, in Sequential Monte Carlo Methods in Practice, New York: Springer-Verlag, **2001**.

[27] Douc R., Cappé O., Comparison of resampling schemes for particle filtering in: Image and Signal Processing and Analysis, **2005**, 64-69.

[28] Imtiaz S., Roy K., Huang B., Shah S. L., Jampana P., Estimation of States of Nonlinear Systems Using a Particle Filter, ICIT **2006**, 2432-2437.

[29] Hsoumi A., El Harabi R., Ali S. B. H., Abdelkrim M. N., Diagnosis of a Continuous Stirred Tank Reactor Using Kalman Filter, Computational Intelligence, Modeling and Simulation, **2009**, 153-158.

[30] McFarlane R. C., Reineman R. C., Bartee J. F., Georgakis C., Dynamic Simulator for A Model IV Fluid Catalytic Cracking Unit, Computers & Chemical Engineering, **1993**, 17, 3, 275–300.

[31] Sadeghbeigi R., Fluid Catalytic Cracking handbook, 2nd edition, Houston (USA): Gulf Professional Publishing Company, **2000**.

[32] Agudelo C., Anglada F. M., Cucarella E. Q., Moreno E. G., Integration Of Techniques for Early Fault Detection and Diagnosis for Improving Process Safety:

Application to a Fluid Catalytic Cracking Refinery Process, Journal of Loss Prevention in the Process Industries, **2013**, 26, 660-665

[33] Ministère chargé de l'environnement, Explosion de gaz dans les unités craquage catalytique et gas plant d'une raffinerie, le 9 novembre **1992**, Aria - fiche juin 2008

(http://www.aria.developpement-durable.gouv.fr/wp-content/files_mf/FD_3969_La_Mede_1992_fr.pdf)

Chapter 7: Summary, Contributions, and Recommendations

7.1 Summary

This work extended the univariate risk-based fault detection method to a multivariate approach by combining both multivariate history based and model based fault detection methods with dynamic risk assessment.

A multivariate history based method, principal components analysis (PCA), was combined with the risk assessment procedure. Principal components score was used as the main indicator for fault. The probability of fault was calculated based on the filtered score; and severity of the fault was a weighted average of severities based on the score and loading of the principal components. Subsequently severity and probability of the fault were combined together to get the risk at any point in time.

In addition, another multivariable risk-based fault detection technique was developed using a process model. Residual generation process based on the Kalman filter has been combined with a risk assessment procedure. Residuals generated from a Kalman filter were used to calculate probability and severity of the fault. Subsequently, the model based approach was extended to a nonlinear and non-Gaussian model for analysis of realistic industrial problems. The estimated states from particle filtering were used for fault detection.

A comprehensive economic consequence methodology was also proposed to fill the gap in existing economic consequence models by relating economic losses to process variables. To calculate the severity of the fault, one could use this methodology for quantifying economic losses associated with the deviation of operational variables from a set-point. The proposed method devised a technique to partition total losses to different categories such as material loss, equipment loss, environmental loss, fatalities and injuries, fines, and claims faced in a process industry. Loss functions were used as a unique way of developing an economic consequence model in order to estimate major losses.

Finally, fault detection using particle filtering was combined with a newly developed economic consequence analysis methodology. For fault detection and probability calculation, process states were estimated by a sequential Monte Carlo method based on point mass representations of probability densities. For severity calculation the new economic consequence analysis was used. This approach allows real-time monitoring in non-linear non-Gaussian systems while making the severity of the faults more meaningful by quantifying consequences in economic terms.

7.2 Contributions

The main conclusions of this study are as follow:

Dynamic risk assessment and fault detection: This work proposed a dynamic risk assessment and fault detection framework targeting the safety issues of a process. In this framework a warning system is activated when the risk of operation exceeds the acceptable threshold.

Early fault detection: Case studies showed how the application of multivariate risk-based fault detection provides early warnings and early activation of safety systems prior to the fault affecting the system. Instead of an alarm activation based on residuals or signals crossing the threshold, an alarm is activated when the risk of operation exceeds the acceptable threshold.

Reducing False Alarms: The proposed approaches have an increased capacity to distinguish between operational changes and abnormal conditions. Combining a multivariable residual generation process based on the Kalman filter, and reliable estate estimation based on particle filtering, and also score based calculations in PCA approach, the risk-calculation procedure makes this method more robust to false alarms.

Distinguish between different faults: Other elements considered include the severity of the fault associated with different process variables and the increasing and decreasing rates of fault. To distinguish between different types of faults two new coefficients (intensity coefficient and moderation coefficient) were introduced in the severity equations. Intensity coefficient indicated the intensity of the severity of the fault associated with each different process variable. Since the severity of the fault in the case of a decreasing rate may not be equal to the severity of the fault in an increasing rate, a new coefficient was introduced to consider this effect and moderate the severity of the outcome.

Risk management and prioritization of safety measures: Using the approaches along with the proposed economic consequence analysis based on loss functions, leads to a more accurate quantification of risk and, more importantly, improved risk management decision-making. In other words, multivariate dynamic risk assessment and fault detection enhanced decision-making for risk management by enhancing prioritization of safety measures and recovery actions.

Multivariate Approach: The proposed methods (based on PCA and Kalman filter) took advantage of the correlation between process input and output and were therefore better equipped to detect faults and calculate risk precisely. These methods though successful in detecting and diagnosing faults have limited use in real industrial process, as they relied on linear models and assumed Gaussian noise which is a common assumption in many fault detection and diagnosis method.

A More realistic approach: A model based approach using a particle filter was developed to deal with nonlinear and non-Gaussian process models. This approach is more reliable in analysis of realistic nonlinear process systems.

Comprehensive economic consequence analysis: A new methodology for conducting consequence analysis in process industries was proposed in this research. This methodology fills the gap in existing economic consequence models by relating economic losses to process variables. It quantifies economic losses associated with the deviation of operational variables from a target value/range. A series of loss functions are proposed as part of the methodology to quantify major losses such as material loss, equipment loss, environmental loss, fatalities and injuries, fines, and claims faced in the process industry.

In this research, fault detection using particle filtering was combined with a newly developed economic consequence analysis methodology. This methodology makes the severity of the faults more sensible by quantifying consequences in economic terms. This new economic consequence methodology helps to integrates real time process state to accident scenarios via loss functions.

7.3 Recommendations

This study can be further extended as follow:

Development of a qualitative approach: In this work a quantitative approach has been taken for dynamic risk assessment and fault detection. One may consider development of a qualitative approach.

More work on history based fault detection methods: This work proposed combination of both model based and history based methods with dynamic risk assessment. Model based methods were considered for both linear and nonlinear problems. However, one may develop some history based methods to deal with nonlinear problems.

Development of a universal univariate model for economic consequence analysis: This work presented a new univariate methodology for economic consequence analysis by assigning an appropriate loss function to each type of loss. However, one may develop a universal model/equation applicable to all types of loss caused by a fault.