# Dynamic Risk Assessment of Process Operations

By

©Sunday Adeshina Adedigba

A thesis submitted to the School of Graduate Studies

In partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

**Faculty of Engineering and Applied Science**

Memorial University of Newfoundland

**November 2017**

St. John's Newfoundland, Canada

*Dedication*

*This work is dedicated to almighty God for his favour and mercy in all ways, and to my wife Esther Oluwemiimo Adedigba and my lovely children: Esther Adeola Adedigba, and Samuel Olugbenga Adedigba.*

*I especially thank beloved late mother, Esther Abeke Adedigba. She labored and sacrificed tirelessly for me.*

# ABSTRACT

Process engineering systems have become increasingly complex and more vulnerable to potential accidents. The risks posed by these systems are alarming and worrisome. The operation of these complex process engineering systems requires a high level of understanding both from the operational as well as the safety perspective. This study focuses on dynamic risk assessment and management of complex process engineering systems' operations. To reduce risk posed by process systems, there is a need to develop process accident models capable of capturing system dynamics in real-time. This thesis presents a set of predictive process accident models developed over four years. It is prepared in manuscript style and consists of nine chapters, five of which are published in peer reviewed journals.  A dynamic operational risk management tool for process systems is developed, considering evolving process conditions. The obvious advantage of the developed methodologies is that it dynamically captures the real time changes occurring in the process operations. The real time risk profile provided by the methodologies developed serve as performance indicator for operational decision making.

The research has made contributions on the following topics: (a) process accident model considering dependency among contributory factors, (b) dynamic safety analysis of process systems using a nonlinear and non-sequential accident model, (c) dynamic failure analysis of process systems using principal component analysis and a Bayesian network, (d) dynamic failure analysis of process systems using a neural network and (e) an integrated approach for dynamic economic risk assessment of process systems.

# ACKNOWLEDGEMENTS

I sincerely express my profound gratitude to my supervisor, Dr. Faisal I. Khan for his love, patience and guidance during my doctoral degree program at Memorial University of Newfoundland. It is a great privilege to have worked under his supervision. I have immeasurably benefitted from his expertise in the field of safety and risk engineering. I would like to acknowledge my co-supervisor Dr. Ming Yang for his great support and encouragement throughout the research work.

I would like to extend my heartiest appreciation to my lovely wife Esther Oluwemimo Adedigba and my adorable children, Esther Adeola Adedigba and Samuel Olugbenga Adedigba. They have been a source of great inspiration and motivation. I greatly missed them for the first three years of this program. My wife ensured the secure welfare of our children so that I could fully concentrate on my PhD. I cannot thank her enough.

I am indebted to my colleagues at the Centre for Risk, Integrity and Safety Engineering (CRISE). I greatly appreciate their support. Also, I appreciate the prayers of all my brothers and sisters. Finally, I appreciate the support of all members of Deeper Life Bible Church, St. John’s.

# TABLE OF CONTENTS

# List of Tables

# List of Figures

## List of Acronyms and Symbol

| | |
|---|---|
| AL | Asset loss |
| ANN | Artificial neural network |
| B | Proportionality constant |
| BN | Bayesian network |
| BP | Back propagation |
| BT | Bow tie |
| CDU | Crude distillation unit |
| $C_i$ | Conditional probability of different states of failure nodes |
| CPT | Conditional probability table |
| CSB | Chemical Safety Board |
| DAG | Direct acyclic graph |
| DPB | Dispersion prevention barrier |
| ECC | Environmental cleanup cost |
| EMFPB | Emergency management failure prevention barrier |
| EML | Estimated maximum loss |
| EPB | Escalation prevention barrier |
| ETA | Event tree analysis |
| FMEA | Failure mode and effect analysis |
| FTA | Fault tree analysis |
| HAZOP | Hazard operability studies |
| HFPB | Human factor prevention barrier |
| HHL | Human health loss |

| | |
|---|---|
| HTHA | High temperature hydrogen attack |
| IBLF | Inverted beta loss function |
| IGLF | Inverted gamma loss function |
| INLF | Inverted normal loss function |
| $I_p$ | Conditional mutual information |
| IPB | Ignition prevention barrier |
| K | Maximum loss |
| K2 | Second loss coefficient |
| K3 | Third loss coefficient |
| K4 | Fourth loss coefficient |
| $l$ | Leak probability |
| L(y) | Actual loss at y |
| LFs | Loss functions |
| MAD | Mean absolute deviation |
| MINLF | Modified inverted normal loss function |
| MLP | Multi-layer perceptron |
| MORT | Management and oversight risk tree |
| MSE | Mean square error |
| n | Total number of components |
| OrFPB | Organization failure prevention barrier |
| P | Probability |
| $P((U\|E)$ | Conditional probability. |
| $P(C_K)$ | The occurrence probability of consequence in the event tree |
| P(U) | Joint probability distribution |
| $Pa(A_{i)}$ | Parent of variable $A_i$ |

| | |
|---|---|
| PCA | Principal Component analysis |
| $P_i$ | Failure probability of each component |
| PL | Production loss |
| PLS | Partial least square |
| QRA | Quantitative risk assessment |
| R | Real time probability |
| $R^2$ | Coefficient of determination |
| RKPCA | Recursive kernel principal components analysis |
| RPB | Release prevention barrier |
| $SB_k$ | Prevention barrier related to level k |
| SSE | Sum square error |
| SVD | Singular value decomposition |
| T | Target value |
| TAN | Tree augmented Naïve Bayes |
| X | Causal factors |
| $x_i$ | Failure probability of prevention barriers |
| Xp | Subset |
| Y | Common effect |
| Z | Training data |
| $\gamma$ | Shape parameter |

# Chapter 1

# 1.0 Introduction

## 1.1. Process Accident Modelling in the Process Industries

In recent decades, chemical process industries (CPI) have been dealing with several hazardous chemicals in various storage units, reactor systems and in other process operations. The complex and nonlinear interactions of process systems, which include equipment, operators, management and organization decisions, operating conditions, and external environmental conditions are the principal causes of accidents in chemical process industries. In most cases, these complex and nonlinear interactions are due to abnormal events and have caused devastating consequences, referred to as accidents (Meel & Seider 2006; Adedigba et al. 2016). This development led engineers to seek a more robust way of incorporating safety into the systems being built and the risk assessment of chemical process industries. Accident models are theoretical frameworks which typically show the relationship between causes and consequences and vividly explain why and how accidents occurred. Accident models are mainly used as techniques for risk assessment during the system development stage and for subsequent use as post hoc accident investigation tools to analyse the root causes of an accident (Qureshi 2008).

## 1.2. Element of Risk Analysis

Risk is defined as “a measure of the potential loss occurring due to natural or human activities. Potential losses are the adverse consequences of such activities in form of loss of human life, adverse effect, loss of property, and damage to the natural environment” (Modarres 2006).

Risk analysis is defined as “the process of characterizing, managing and informing others about the existence, nature, magnitude, prevalence, contributing factors, and uncertainties of the potential losses”(Modarres 2006). In other words, risk analysis is a methodical and scientific technique to predict and prevent the occurrence of an accident in a system. Its primary purpose is to avoid risk in a system. Risk analysis can be applied at different stages: design, development and construction and operation stages. The three principal elements of risk analysis are: risk assessment, risk management and risk communication. Interaction and overlap among the main elements of risk analysis are revealed by Figure 1.1. It is obvious that elements of risk analysis are intertwined and synergize one another. This synergy must be maintained for effective risk analysis.

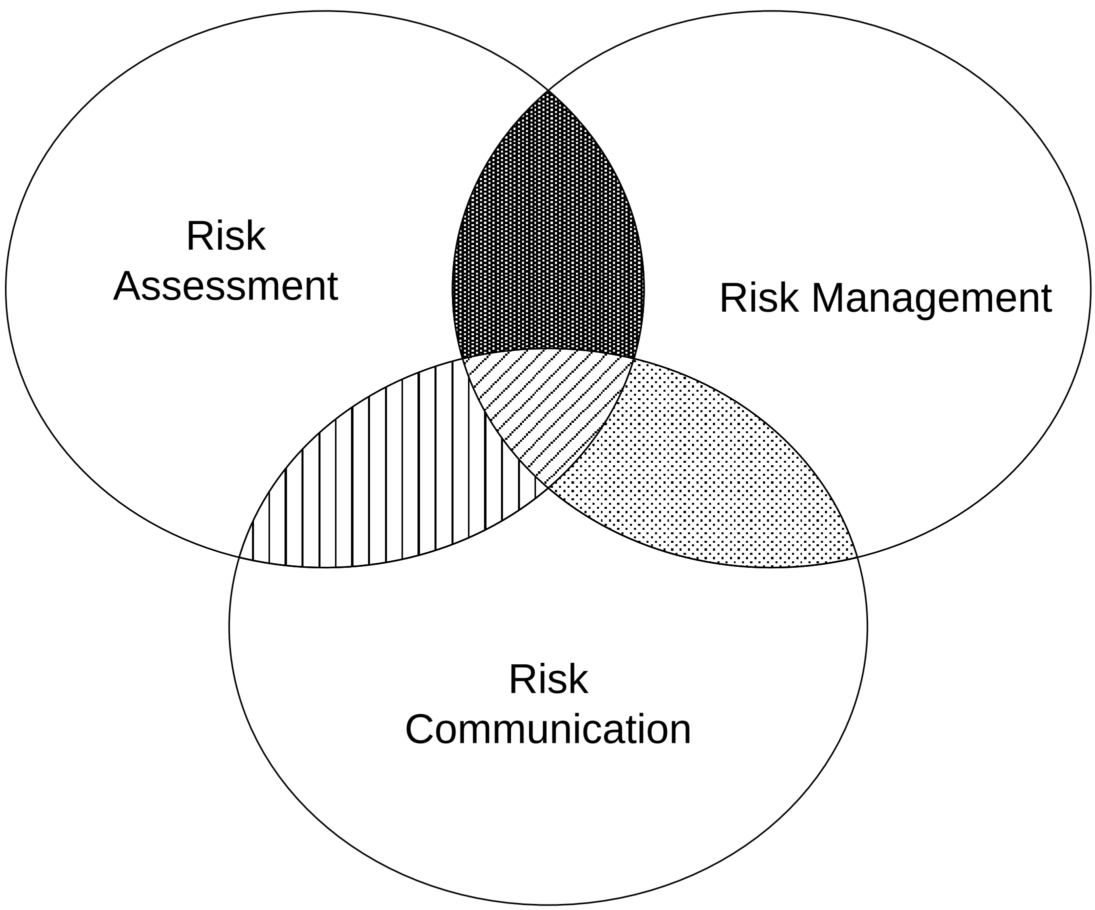



**Figure 1.1. Elements of Risk analysis.**

### 1.2.1. Risk Assessment

Risk assessment is a process or technique by which the probability of a loss in an engineering system is predicted and the magnitude (consequences) of the loss is also estimated. In broad terms, risk assessment addresses three main questions; (1) what can go wrong? (2) how likely is it? and (3) what are the losses (consequences)? The entire process of risk assessment involves incident identification and consequence analysis. Incident identification defines in detail how an accident process occurs and analyzes the probabilities. Consequence analysis expressly describes the expected damage. In general, risk assessment involves identification of potential scenarios, computing of their individual occurrence probabilities and explanation of the consequences that originated from each scenario. A risk assessment technique that estimates only probability is referred to as a probabilistic risk assessment (PRA), while the risk assessment process that estimates probabilities alongside with the consequences is termed quantitative risk analysis (QRA) (Crowl & Louvar 2001; Modarres 2006). Risk assessment is of paramount importance in estimating the safety, reliability and effectiveness of an engineering system (Khan et al. 2015; Villa et al. 2016).

### 1.2.2. Risk Management

Risk management is the procedure by which the likelihood of the magnitude and various risk contributors are predicted, appraised, minimized and controlled. It is a systematically coordinated procedure to avert, regulate, and minimize losses suffered as the result of risk exposure, weighing options and choosing suitable actions by taking into account risk values, legal and political issues and economic and technological constraints (Modarres 2006). In summary, in engineering terms, risk management should be considered as a " a control function focused on maintaining a particular hazardous, productive process within the boundaries of safe operation" (Rasmussen 1997).

Different types of risk management techniques have been developed and used to ensure the safety of chemical process systems. The main objective of these techniques is to identify process hazards, assess them, control the hazard and ultimately mitigate the residual risk at both the design and operational phases (Aven 2016). The main objective of risk management during the entire life cycle of a complex engineering system entails proactive decision making to:

- Frequently assess the risk
- Select which risks are significant
- Provide strategies to prevent or control the risks
- Frequently evaluate the efficiency of the strategies and review them, when necessary (Modarres 2006).

### 1.2.3. Risk Communication

Risk communication is defined as "the flow of information and risk evaluation back and forth between academic experts, regulatory practitioners, interest groups and the general public"(Leiss 1996). Risk communication basically updates, exchanges and transfers information and knowledge about risk, risk assessment outcomes and various risk management alternatives among analysts, decision makers and other stakeholders. Depending on the targeted audience, risk communications usually provide adequate information on the following specific areas: the nature of the risk, the nature of the benefits, risk management options and uncertainties in risk assessment (Modarres 2006).

Risk communication comprises perceptions of the risk and depends on the targeted audience; therefore, risk communications can be broadly classified into public, media and engineering community risk communications (Ayyub 2003).

## 1.3. Accident models

Accident models are theoretical frameworks which typically show the relationship between causes and consequences and explain in detail why and how accidents occurred. Accident models are mainly used as techniques for risk assessment during the system development stage and for subsequent use as a post hoc accident investigation to analyse the root causes of an accident (Qureshi 2008). Accident models systematically relate causes and consequences of the events and play a significant role in accident investigation and analysis. They tend to primarily address two major broad questions: (i) why accidents occur and (ii) how accidents occur. Classification of accident models can be done in several ways. Accident models are broadly categorized as either traditional or modern accident models. Traditional accident models are further sub grouped into sequential and epidemiological models. They are primarily descriptive models that lack predictive capacity and emphasize mainly human, organizational and management factors. Modern accident models can be sub classified into three sub categories: systematic, formal and dynamic accident models ( Al-shanini et al. 2014; Qureshi 2008).

Existing accident models have their own strengths and weaknesses and these depend mainly on the areas of their application, purpose and focus. The majority of the existing accident models are sequential accident models where accident processes from initiation to termination are considered as chains of independent events that occurred in a definite particular order. The severity of effects is presumed to progress through the sequential failure of independent events. These traditional models use a fault and event trees sequential approach to predict the cause-consequence relationship, which provides a sequential explanatory mechanism of accident propagation. However, in a real life situation, this may not be true.

Also, existing models are not capable of modelling multiple risk factors in process systems where interactions among systems are nonlinear and extremely complex, and they are not capable of using accident precursor data to evaluate risk and develop accident prevention strategies (Tan et al. 2013; Rathnayaka et al. 2011).

A thorough review of existing accident models reveals that the majority of the models belong to the class of sequential accident models, where the accident process is described as a chain of independent events that take place sequentially. Hence, the study is presenting dynamic accident model. A comprehensive review of the existing chemical process accident models shows that the current process accident models exhibit obvious weaknesses. These weaknesses are: (1) External hazards are not considered in the model. (2) The model presumes the causes of failure within safety barriers are independent, although in reality they are interdependent and this could significantly affect the results. (3) Provision is not made for other factors not accounted for in the fault tree model of prevention barriers. (4) Nonlinear interaction of various factors are not considered. (5) They do not capture evolving operational conditions, the time variant behaviour of process parameters and their dependent relationships.

The current study is an attempt to address some of these gaps and to contribute appreciable knowledge to this area of research.

## 1.4. Objectives of the Research

The primary objective of this research is to develop an integrated dynamic operational risk management tool for process operations. This key objective is divided into five sub-objectives:

- To develop an innovative predictive probabilistic model to assess hazardous process operation accident likelihood, such that accident occurrence probability bounds will be predicted. This will serve as an effective tool to facilitate risk assessment and management of process operations.

- To develop a dynamic nonlinear and non-sequential Bayesian network based process accident causation model for chemical process operations.
- To develop a dynamic Bayesian nonlinear model capable of predicting inter-dependency among process operations variables and subsequently predict and update risk profile dynamically, using Bayesian TAN algorithms.
- To developed a dynamic ANN model that is capable of predicting the risk profile from process monitoring data empirically and subsequently generalizing the ANN model developed to predict risk profile for chemical process operations.
- To develop an integrated approach for dynamic economic risk assessment of process systems.
- To test and verify the models developed with real life case studies.

## 1.5. Organization of the Thesis

Manuscript (paper) format is used in writing this thesis. The outlines of each chapter are presented below:

Chapter 2 presents the innovations and major contributions of this thesis to the dynamic risk assessment of chemical process operations.

Chapter 3 presents a thorough literature review relevant to the research. This includes a brief description of different accident models and risk assessment techniques.

Chapter 4 presents an innovative predictive probabilistic model to assess hazardous process operation accident likelihood such that accident occurrence probability bounds are predicted using a new non-sequential barrier-based process accident model.

This chapter is published in the *Journal of Process Safety and Environmental Protection 2016; 102: 633-647.*

Chapter 5 present a dynamic nonlinear and non-sequential Bayesian network based process accident causation model for dynamic risk prediction of chemical process operations.

This chapter is published in the *Journal of Chemical Engineering Research and Design 2016; 111: 169-183.*

Chapter 6 presents an integrated dynamic failure prediction analysis approach using principal components analysis (PCA) and Bayesian TAN algorithms.

This chapter is published in the *Journal of Industrial and Engineering Chemistry Research 2017; 56: 2094-2106.*

Chapter 7 presents an integrated ANN probabilistic approach capable of predicting the risk profile from process monitoring data empirically and subsequently generalizing the ANN model developed to predict risk profile for chemical process operations. This chapter is published in the *Journal of Process Safety and Environmental Protection 2017; 111: 529-542.*

Chapter 8 presents a dynamic economic risk assessment framework which integrates probability with consequences assessment. This model establishes the link between the process deviation with not only the probability estimation but also the potential loss prediction due to such deviations.

This chapter is published in the *Journal of Process Safety and Environmental Protection 2018; 116:312-325*

The logical relationship and progression among the chapters is represented by Figure 1.2.

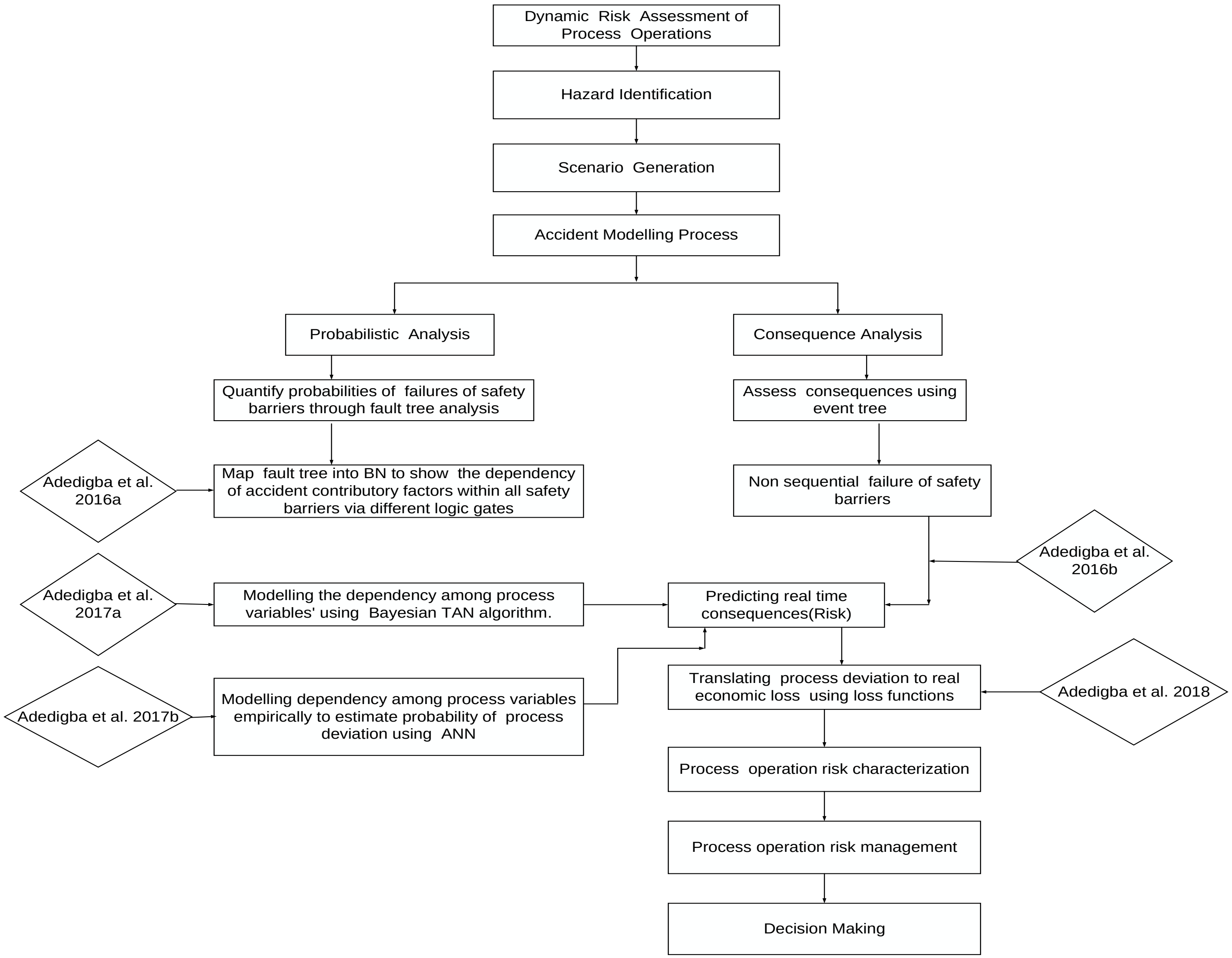



**Figure 1.2. Dynamic Operational Risk Management Tool for Process Systems.**

## References


Adedigba, S.A., Khan, F. & Yang, M., 2016. Process accident model considering dependency among contributory factors. *Process Safety and Environmental Protection*, 102, pp.633–647.

Al-shanini, A., Ahmad, A. & Khan, F., 2014. Accident modelling and safety measure design of a hydrogen station. *International Journal of Hydrogen Energy*, 39(35), pp.20362–20370.

Aven, T., 2016. Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253(1), pp.1–13.

Ayyub, B.M., 2003. *Risk Analysis in Engineering and Economics*,

Crowl, D.. A. & Louvar, J.F., 2001. *Title Chemical process safety:fundamentals with appplication* Third edit., Pearson Education.

Katsakiori, P., Sakellaropoulos, G. & Manatakis, E., 2009. Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. *Safety Science*, 47(7), pp.1007–1015.

Khan, F., Rathnayaka, S. & Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection*, 98, pp.116–147.

Leiss, W., 1996. American Academy of Political and Social Science Three Phases in the Evolution of Risk Communication Practice Three Phases in the Evolution of Risk Communication Practice. *Source:*

Meel, A. & Seider, W.D., 2006. Plant-specific dynamic failure assessment using Bayesian theory. *Chemical Engineering Science*, 61(21), pp.7036–7056.

Modarres, M., 2006. *Risk analysis in engineering: techniques, tools, and trends*,

Qureshi, Z.H., 2008. A Review of Accident Modelling Approaches for Complex Critical Sociotechnical Systems. , 86, p.72.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. *Safety Science*, 27(2-3), pp.183–213..

Villa, V. et al., 2016. Towards dynamic risk analysis: A review of the risk assessment approach and its limitations in the chemical process industry. *Safety Science*, 89, pp.77–93. 2017.

# Chapter 2

## 2.0. Novelty and Contribution

The novelties and main contributions of this doctoral research is in the area of dynamic safety and risk assessment of chemical process operations. The highlights of these contributions are stated below:

- An innovative probabilistic model to assess hazardous process operation accident likelihood. This work proposed a novel non sequential barrier based accident model, in which interdependency and nonlinear interaction among accident contributory factors within safety barriers are modelled for process accidents. This novel probabilistic approach is an effective tool to facilitate risk assessment and management of process hazards. This contribution is presented in chapter 4.
- An innovative predictive non sequential barrier based process model. The model account for non-linear interaction of accident contributory factors within safety barrier and subsequently allow the non-sequential failure of safety barriers to cause adverse event randomly. The model developed takes into consideration the complexity of process's operations and high level of interaction among sub-systems, thus accident causation is function of nonlinear interaction of various factors. This contribution is presented in Chapter 5.
- A novel probabilistic methodology that integrates principal Component analysis (PCA) and the Bayesian networks to detect fault and predict the probability of failure using real time process data. The key process variables that contribute the most to process performance variations are detected with PCA, while Bayesian network is adopted to model the

interactions among process variables to detect faults and predict time the time dependent probability of system failure. This contribution is presented in Chapter 6.

- An innovative Artificial neural network (ANN) data driven model. The model developed used a multi-layer perceptron (MLP) to define the relationship among process variables. The defined relationship is used to model a process accident considering logical and causal dependence of the variables. The predicted accident probability is subsequently used to estimate the likelihood of failure to the process unit. The model provide an efficient and effective way to estimate process accident probability as function of time and thus risks. This contribution is presented in Chapter 7.
- An innovative dynamic economic risk assessment framework which integrates probability with consequences assessment. The model developed link process deviation to accident probability and potential losses. The Bayesian Tree Augmented Naïve Bayes (TAN) algorithm is applied to model the precise and concise probabilistic dependencies that exist among key operational process variables to detect faults and predict the time dependent probability of system deviation. Loss function is used to define system economic losses as a function of process deviation. The time dependent probability of system deviation owing to an abnormal event is constantly updated based on the present state of the relevant process variables. This innovative contribution is presented in Chapter 8.

# Chapter 3

# 3.0. Literature Review

## 3.1. Chemical Process Industry (CPI)

The chemical process industry (CPI) uses extremely complex technological systems consisting of various equipment, operating procedures and control schemes. The plants in the chemical industry handle vast quantities of hazardous chemicals. The dynamic interactions among these various components: equipment, management and organizational (M&O) and human factors make CPI vulnerable to process deviations, which might eventually lead to failures if not correctly tackled (Al-shanini et al. 2014). In recent decades, notable devastating accidents such as the Piper and Alpha tragedy, the Bhopal toxic gas release disaster, the Nypro factory explosion at Flixborough, the Imperial sugar refinery dust explosion, BP's Texas city refinery explosion, and BP's Deepwater Horizon offshore drilling rig explosion are notable examples of complex system failure that caused devastating losses of human lives and properties (Rathnayaka et al. 2011).

Probabilistic Safety Assessment is a standard technique for safety assessment of complex and critical engineering systems. This technique is applicable to all phases of the engineering system life cycle: design, start up and different modes of operations. The primary objective of safety assessment techniques is identification of all potential hazards to prevent them and subsequently mitigate the residual risk. It is of paramount important to integrate management oversight and engineering analyses to formulate a comprehensive and systematic approach to effectively manage system risk (Cepin & Mavko 1997; Bahr 1977).

Process safety primarily focuses on prevention and mitigation of major process accidents such as toxic releases, fire, and explosions. The key steps in process safety assessment and management are hazard identification, risk assessment and management (Khan et al. 2015). Hazard identification steps primarily identify all potential process hazards and may analyze how these hazards can combine to cause accidents (Rathnayaka et al. 2010). Risk can be used as a parameter to measure process safety and it is quantitatively expressed as a product of probability and its consequences (Modarres 2006; Khan et al. 2015). Risk management involves systematic techniques to prevent, control and minimize losses suffered due to a risk exposure through the process of risk estimation, risk evaluation, risk based decision making and design improvement (Modarres 2006; Khan et al. 2015).

An effective means of combating process accidents is to develop an appropriate preventive measure focusing on the correct process plant components. Such approaches lie within the realm of accident modelling.

Accident models give detailed conceptualization of the characteristic accident, and essentially display the relationship between causes and effects. They are risk assessment technique to explain the causes of accidents (Qureshi 2007).

## 3.2. Classification of Accident models

Generally, accident models can be classified into two broad categories: traditional and modern accident models.

## 3.3. Traditional Accident models

Traditional accident models are broadly classified into sequential models and epidemiological models.

### 3.3.1. Sequential models

Sequential accident models are the most simplified types of accident models. They explain accident causation as the result of a chain of events that occurs in a definite order (Hollnagel 2002).

This models are not usually restricted to a single chain of events, and may be denoted in the form of hierarchies such as :petri networks, traditional event trees, Bayesian networks, fault trees and critical path models (Hollnagel & Goteman 1982). One famous sequential accident model is Domino theory, proposed by Heinrich in the 1940s. Domino theory describes an accident as a sequence of discrete events which occurred in a define temporal order, ending in an injury. Domino theory emphasizes that an accident can be prevented by eliminating any single factor from the accident sequence (Qureshi 2007; Rathnayaka et al. 2011). It has been modified by the International Loss Control Institute into a loss causation model (ILCI model) to predict how unsafe acts and conditions originate. The ILCI model shows a broader representation of accident propagation. Analysis in the ILCI model starts with the loss to people, property and the environment and propagate backwards through the chain of events that contributes to the loss individually. The immediate and the root cause of accidents in the ILCI model are described as personnel and job factors, management deficiencies and unsafe acts and conditions. However one disadvantage of Domino theory and the ILCI model is that the cause-consequence relationship among management, organization and the human level is not properly defined (Rathnayaka et al. 2011). Domino theory was widely adopted in various industries; however, many industrial accidents that occurred in the 1970s cannot be sufficiently described using a simple cause-consequence relationship. In general, sequential accident models are not capable of modeling nonlinear interaction among system components (Hollnagel & Goteman 1982; Qureshi 2007).

### 3.3.2. Epidemiological models

Epidemiological models try to describe causes of accidents in complex systems. As the name suggests, the model describes accident causation as analogous to the spreading of a disease, i.e., the result of a combination of both manifest and hidden factors that occur simultaneously in space and time. This terminology is defined as “ the unexpected, unavoidable unintentional act resulting from the interaction of host, agent and environmental factors within situations which involves risk taking and perception of danger”(Hollnagel 2002; Hollnagel & Goteman 1982). The view of epidemiological models is that an accident occurs due to a combination of “agents” and environmental factors that initiate an “unhappy setting”. Epidemiological models are important because they give the foundation for deliberating the complexity of accidents’ processes, making epidemiological models more advantageous than sequential models (Hollnagel 2002). One major epidemiological model is the Swiss cheese model proposed by Reason. The Swiss cheese model highlights how both human and organization failures initiate the accident process independently, taking the multi-causality of the accident into consideration (Underwood & Waterson 2014).

The Swiss cheese model has been adopted in many process industries to avert accidents due to human error. The Swiss cheese model places principal cheese slices sequentially along the accident path. The cheese slices represent the relevant safety barriers, while the holes denote the latent errors. The cheese slices act as protective (defensive) barriers against an incident or accident, while the holes are subjected to variation (change) based on the failure types. Once the holes are lined up, all safety barriers have failed; therefore, an accident will occur. The series of holes in the first cheese slice denote the hidden or latent failures. Unsafe acts are typically located in the last slices, while latent conditions are the holes through the cheese.( Katsakiori et al. 2009; Qureshi 2007; Rathnayaka et al. 2011).

## 3.4. Modern Accident models

Modern accident models are broadly classified into systemic models, formal models and dynamic sequential models (Al-shanini et al. 2014).

### 3.4.1. Systematic models

Systematic models explains the characteristic performance of the system as a whole rather than on the basis of the precise causes, the consequence mechanism or epidemiological factors. These models view accidents as emergent phenomena due to variability of the system (Stroeve et al. 2009). Systematic models have their origins in system theory, which comprises control theory, chaos models, coincidence models and stochastic resonance. All these models and theories are used to understand and predict complex interrelationships and interdependencies among system components, including human, technical, organizational and management factors (Hollnagel 2002; Qureshi 2007). Systematic accident models are broadly divided into theory system models and cognitive system models (Al-shanini et al. 2014). In `applying a systems theory approach to model an accident, the systems are viewed as comprising various interacting components which keep equilibrium by the use of feedback loops of information and control. The system theory approach sees the system as a dynamic process that is frequently adjusting to attain it primary objectives and responding to its changes and those in the environment. The design of the system imposes constraints on its performance characteristics for safe operation and at the same time must adjust to dynamic changes to preserve safety. In a systemic approach, accidents occur due to flawed processes involving complex interactions among engineering activities, people, organizational structure and physical and software system components (Hollnagel 2002; Qureshi 2007). Quite a number of systemic accident models have been developed. Two famous systemic accident models are Rasmussen's (1997) hierarchical socio-technical framework and Leveson's (2004)

STAMP (Systems Theoretic Accident Model and Processes). The STAMP model accidents occur as a result of inadequate control or inadequate implementation of safety associated constraints at the development, design and operation phases of the system and not because of independent component failures (Rathnayaka et al. 2011).

The cognitive system approach models the performance behaviour of a human-machine system from the perspective of the environment in which the work is taking place. Cognitive system models propose that we cannot comprehend what happens when thing go wrong without comprehending what happens when thing go well (Hollnagel & Wood 2005). The Joint cognitive systems explain "how humans and technology functions as a joint system rather than how humans interact with machines. Efforts to make work safe should start from an understanding of the normal variability of human and joint cognitive system performance rather than assumptions about particular but highly speculative error mechanism" (Qureshi 2007). Quite a few systemic accident models are based on the principles of Cognitive system engineering and include: Cognitive Reliability and Error and Analysis Method (CREAM), Functional Resonance Accident Model (FRAM) and Drivers Reliability and Error and Analysis Method (DREAM). CREAM models the cognitive features of human performance for the purpose of assessing the consequences of human error on system safety. FRAM describes in detail how components of the system may resonate and initiate hazards that can cascade into accidents. DREAM is another version of CREAM that is applied for analysis of traffic accidents (Hollnagel 2004; Hollnagel 1988). Contrary to sequential and epidemiological accident models, the systemic models do not depend on static cause - consequence relationships. Systemic models account for the dynamic, non-linear and perhaps the resonance resembling interactions that may likely cause accidents (Stroeve et al. 2009).

### 3.4.2. Formal models

Formal modes are built using mathematical techniques that offer a laborious and logical framework for design, specification and authentication of a computer system, including software and hardware. Fundamentally, formal methods of accident analysis use formal specification language consisting of three principal components: procedures for defining the syntax (grammatical well-formedness of sentences); rules for interpreting the semantics and subsequently rules for deducing important information from the proof theorems. This approach substantiates that the system has been designed correctly and demonstrates the features of the system without necessarily operating the system to know its behaviour (Van Lamsweerde 2000). Formal methods provide improvements to accident analyses by highlighting the significance of precision and in definitions and explanations, and subsequently giving symbols to explain certain parts of accidents (Qureshi 2007). Two famous formal models approach are: Why Because Analysis (WBA) and the probabilistic model of causality (Qureshi 2007). Detailed application of formal methods in both industry and research can be found in (Hinchey & Bowen 1995). WBA is built using formal semantics and logic. It applies deontic action logic as language for the construction of a formal model and is primarily focused on analyzing causality, subsequently permitting the unbiased evaluation of the events and conditions as causal factors. WBA starts with the reconstruction phase, where graphical formal notations are used to model the chains of events causing the accident. The important events and states are derived from accident investigation reports in their stint order. The event sequences are denoted in a logical form and are analyzed to find the reason for the accident using a graph called a WB graph. The WB graph is thoroughly examined to ascertain that causal relations in the graph satisfy the semantic causation relationship defined. The

WBA method has been adopted on a number of occasion to analyse aircraft and train accidents (Qureshi 2007).

Probabilistic causality models consider the relationship between cause and effect using probability theories. Many different methods have been adopted by various authors to develop probabilistic theories of causality. One principal contribution is the mathematical theory of causality propounded by ( Pearl 2000). This applies a structural causal model semantic and subsequently describes a probabilistic causal model as a pair. Figure 3.1 shows the classification of accident models.

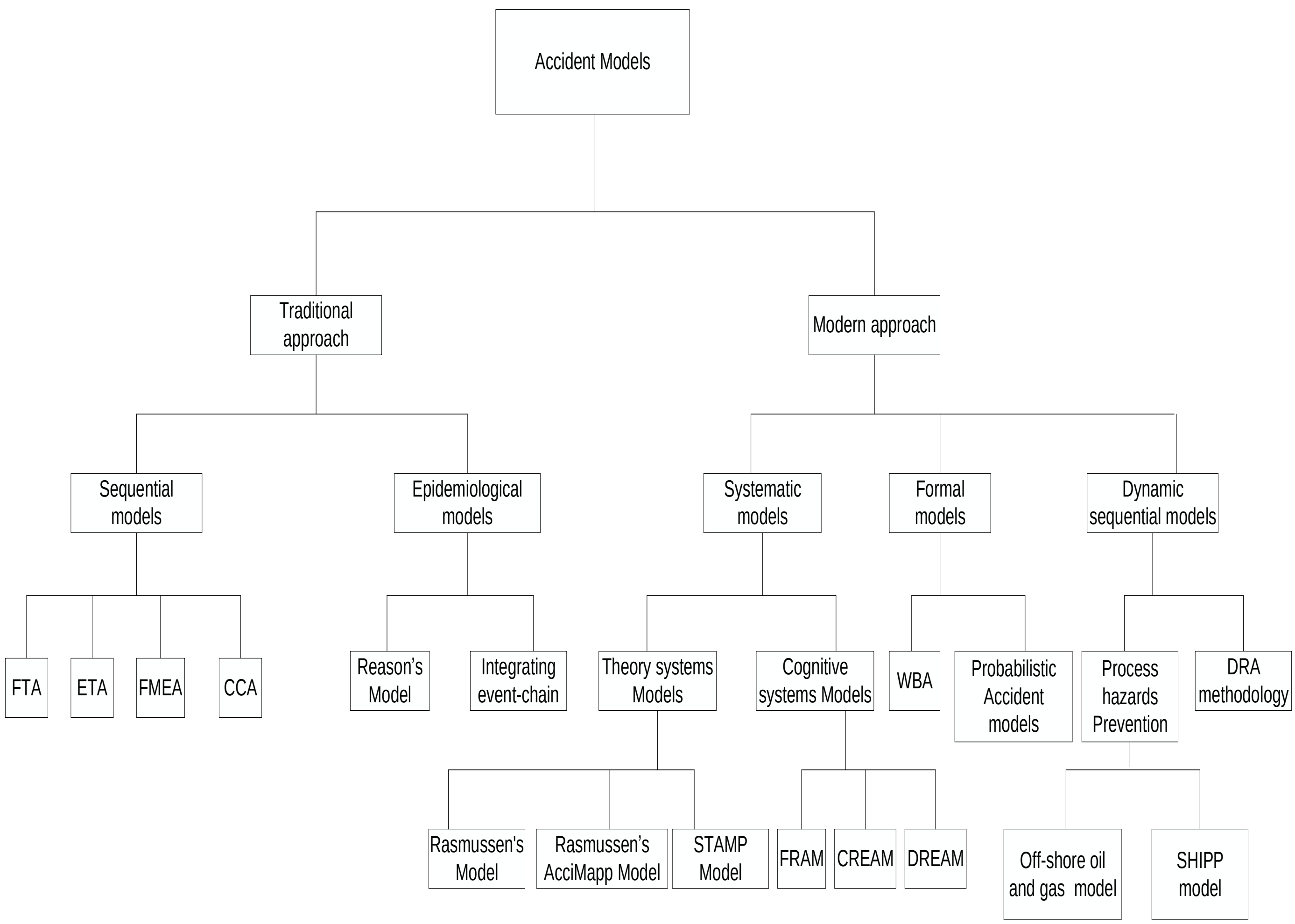



**Figure 3.1. Classification of accident models.**

The existing accident models belong to the class of sequential accident models, where the accident process is described as a chain of independent events that take place sequentially. They do not capture evolving operational conditions, the time variant behaviour of process parameters and their dependent relationships. The current study present dynamic accident models to address some of the gaps of sequential models.

## 3.5. Quantitative Risk Analysis

Risk is defined as a "measure of human injury, environmental damage, or economic loss in terms of both the incident likelihood and the magnitude of the loss or injury" (CCPS 1999).

Quantitative Risk Analysis (QRA) deals with the quantitative estimate of risk using mathematical techniques based on engineering evaluations for combining estimates of incident consequences and frequencies (CCPS, 1999). The most widely used techniques are fault tree analysis, event tree analysis and a combination of both fault tree and event tree analysis, which is known as the Bow-tie technique. Nevertheless, these conventional risk assessment methods are static in nature, failing to capture the variation of risks as an operation fluctuates (Ferdous et al. 2011; Khakzad et al. 2012).

Recently, BNs have gained much attention because they can accommodate different kinds of statistical dependencies that cannot be easily included in other accident analysis techniques. A brief description of these conventional risk analysis techniques is given below.

### 3.5.1. Fault Trees

A fault tree is a deductive, graphic methodology used to determine failure probability of a complex system. The top event in the fault tree represents a major accident initiating hazard. The top event is placed at the top of the fault tree and the fault tree is graphically modelled downward to allow the visualization of all possible combinations of malfunctions and wrong actions that could initiate the top event. Fault trees are usually constructed from events and logic gates (Khakzad et al. 2011). The underlying technical failures that lead to accidents are usually denoted as basic events. The logic gates in the fault tree represent numerous ways by which machines and human error interact to cause the accident. AND and OR gates are the commonly used logic gates in the fault tree. Analysis using the fault tree can proceed both qualitatively and quantitatively (Nivolianitou et al. 2004). In the AND gate, process components interact in parallel structure and process failure requires the simultaneous failure of all components in parallel. The failure probability of the top event in parallel structure (AND gate) is calculated by equation 3.1. Also, in the OR gate, process components interact in a series structure and failure of any single component in the series leads to failure of the process. The failure probability of the top event in a series structure (OR gate) is calculated by equation 3. 2.

$$\boldsymbol{P = \prod_{i=1}^{n} P_i} \qquad (\mathbf{3.1})$$

$$\boldsymbol{P = \prod_{i=1}^{n} (1 - P_i)} \qquad (\mathbf{3.2})$$

### 3.5.2. Event Trees

The event tree is an inductive systematic technique that starts with a specified accident initiating event and terminates with all the feasible consequences, normally called the "end state consequences" of the event tree. Event tree techniques are widely used to denote incident scenarios. They describe a probable sequence related to an accident initiating event that transits through successive prevention barriers and terminates with ultimate consequences (Nývlt & Rausand 2012). The likelihoods (probabilities) of end state consequences P($C_k$) are quantify by equation 3.3.

$$\mathbf{P(C_K) = \prod_{j \in SB_k} x_i^{\theta_{i,k}} (1 - x_i)^{1-\theta_{i,k}}} \qquad (3.3)$$

where $SB_k$ represents the prevention barrier related to level k; and $\theta_{i,k} = 1$ whenever a level k failure transits through the failure branch of safety ( prevention) barrier i; $\theta_{i,k} = 0$ whenever a level k failure transits through the success branch of safety( prevention) barrier i. $x_i$ is the failure probability of prevention (safety) barriers (Adedigba et al. 2016; Rathnayaka et al. 2010).

### *3.5.3.* Bow-Tie (BT) Analysis

Bow-tie analysis (BT) is a technique that combines fault tree analysis and event tree analysis. In this approach, the top event in the fault tree analysis serves as an initiating event of the event tree analysis. This technique clearly analyzes root causes and resultant consequences of an accident process. A BT is graphical logical relationship among several causes, denoted as basic events on one side and potential consequences on the other side, via prevention barriers (safety barriers). Bow-tie analysis combines the advantages of FT and ET and has been widely applied in various fields of science. The BT risk analysis technique has been applied to a dust explosion accident in a sugar refinery (Khakzad et al. 2012). BT, like Fault tree and event tree analysis, shows similar limitations and deficiencies of independency assumptions and is extremely difficult to apply for complex system analysis.

### *3.5.4.* **Bayesian Network**

The Bayesian network (BN) is a graphical technique; it provides a robust probabilistic technique of reasoning under uncertainty. BN techniques have been extensively used in risk and safety analysis based on probabilistic and uncertain knowledge. BN (also known as a probabilistic dependence graph) is a direct acyclic graph with numerous nodes representing variables and arcs signifying direct causal relationships among the linked nodes. A conditional probability table (CPT) is assigned to the various nodes to denote conditional dependencies among the linked nodes (Bobbio et al. 2001; Khakzad et al. 2013). Based on both conditional independence and the chain rule, the BN represents the joint probability distribution P(U) of a set of discrete random variables, U= $\{A_1\ldots,A_n\}$, incorporated in the network as:

$$\mathrm{P(U)} = \prod_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{P(A_i|\ P_{a(A_i}))} \tag{3.4}$$

where Pa ($A_i$) is the parent of variable $A_i$ and P (U) is the joint probability distribution of variables (Pearl 1998; Jensen & Nielsen 2007).

The BN makes use of Bayes theorem to update the prior occurrence probability of events to give consequence probability (posterior) provided that new information called evidence is given. The following equation is used to estimate posterior probability:

$$P(\mathrm{U|E}) = \frac{\mathrm{P(U,E)}}{\mathrm{P(E)}} = \frac{\mathrm{P(U,E)}}{\sum_{\mathrm{U}} \mathrm{P(U,E)}} \tag{3.5}$$

## References

Adedigba, S.A., Khan, F. & Yang, M., 2016. Chemical Engineering Research and Design Dynamic safety analysis of process systems using nonlinear and non-sequential accident model. *Chemical Engineering Research and Design*, 111, pp.169–183.

Al-Muslim, H. & Dincer, I., 2005. Thermodynamic analysis of crude oil distillation systems. *International Journal of Energy Research*, 29, pp.637–655.

Al-shanini, A., Ahamad, A. & Khan, F., 2014. Accident modelling and analysis in process industries. *Journal of Loss Prevention in the Process Industries*, 32, pp.319–334.

Bahr, N.., 1977. *System Safety Engineering and Risk Assessment: A Practical Approach, Second ... - Nicholas J. Bahr - Google Books*, Taylor & Francis, NY.

Bobbio, a. et al., 2001. Improving the analysis of dependable systems by mapping Fault Trees into Bayesian Networks. *Reliability Engineering and System Safety*, 71, pp.249–260.

CCPS, 1999. *Guidelines for Consequence Analysis of Chemical Releases*, New York: AMerical Institute of Chemical Engineers.

Cepin, M. & Mavko, B., 1997. Probabilistic safety assessment improves surveillance requirements in technical specifications. *Reliability Engineering & System Safety*, 56, pp.69–77.

Ferdous, R. et al., 2011. Fault and Event Tree Analyses for Process Systems Risk Analysis: Uncertainty Handling Formulations. *Risk Analysis*, 31(1), pp.86–107.

Hinchey, M.G. (Michael G. & Bowen, J.P. (Jonathan P., 1995. *Applications of formal methods*, Prentice Hall.

Hollnagel, E., 2004. *Barrier and Accident prevention*, Hampshire, Ashgate: Elsevier.

Hollnagel, E., 1988. *Cognitive Reliability and Error Analysis Method*, Elsevier.

Hollnagel, E., 2002. Understanding accidents-from root causes to performance variability. *Proceedings of the IEEE 7th Conference on Human Factors and Power Plants*, pp.1–1–1–6.

Hollnagel, E. & Goteman, Ö., 1982. The Functional Resonance Accident Model. , 94, pp.176–196.

Hollnagel, E. & Wood, D.., 2005. *Joint Cognitive Systems : Foundation of Congitive Systems Engineering*, New York: Taylor & Francis, NY.

Jensen, F.. & Nielsen, T.., 2007. *Bayesian Networks and Decision Graphs* 2 edition., Springer, New York.

Katsakiori, P., Sakellaropoulos, G. & Manatakis, E., 2009. Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. *Safety Science*, 47(7), pp.1007–1015.

Khakzad, N., Khan, F. & Amyotte, P., 2012. Dynamic risk analysis using bow-tie approach. *Reliability Engineering & System Safety*, 104(null), pp.36–44.

Khakzad, N., Khan, F. & Amyotte, P., 2013. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection*, 91(1-2), pp.46–53.

Khakzad, N., Khan, F. & Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety*, 96(8), pp.925–932.

Khan, F., Rathnayaka, S. & Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection*, 98, pp.116–147.

Van Lamsweerde, A., 2000. Formal Specification: a Roadmap. In *Procceeding of the Conference on The Future of Software Enginering , Limerick, Ireland.* ACM Pres, pp. 147–159.

Leveson, N., 2004. A new accident model for engineering safer systems. *Safety Science*, 42(4), pp.237–270. Available at: http://linkinghub.elsevier.com/retrieve/pii/S092575350300047X [Accessed March 12, 2017].

Modarres, M., 2006. *Risk analysis in engineering: techniques, tools, and trends,*

Nivolianitou, Z.S., Leopoulos, V.N. & Konstantinidou, M., 2004. Comparison of techniques for accident scenario analysis in hazardous systems. *Journal of Loss Prevention in the Process Industries*, 17, pp.467–475.

Nývlt, O. & Rausand, M., 2012. Dependencies in event trees analyzed by Petri nets. *Reliability Engineering and System Safety*, 104, pp.45–57.

Pearl, J., 2000. Causality: Models, Reasoning, and Infeerence. *Econometric Theory*, 19, pp.675–685.

Pearl, J., 1998. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmannn, San Francisco, California.

Qureshi, Z.H., 2007. A Review of Accident Modelling Approaches for Complex Socio-Technical Systems. In *12th Austrialian Conference on safety Related progarammable Systems*. pp. 47–60.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. *Safety Science*, 27(2-3), pp.183–213.

Rathnayaka, S., Khan, F. & Amyotte, P., 2010. SHIPP methodology : Predictive accident modeling approach . Part II . Validation with case study. *Process Safety and Environmental Protection*, 89(2), pp.75–88.

Rathnayaka, S., Khan, F. & Amyotte, P., 2011. SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection*, 89(3), pp.151–164.

Stroeve, S.H., Blom, H.A.P. & (Bert) Bakker, G.J., 2009. Systemic accident risk assessment in air traffic by Monte Carlo simulation. *Safety Science*, 47(2), pp.238–249.

Underwood, P. & Waterson, P., 2014. Systems thinking, the Swiss Cheese Model and accident analysis: A comparative systemic analysis of the Grayrigg train derailment using the ATSB, AcciMap and STAMP models. *Accident Analysis & Prevention*, 68, pp.75–94.

# Chapter 4

## 4.0 Process accident model considering dependency among contributory factors

### Preface

*A version of this chapter has been published in the* ***Journal of Process Safety and Environmental Protection 2016; 102: 633-647.*** *I am the primary author. Co-author Faisal Khan provided fundamental understanding, assisted in developing the conceptual model and subsequently translated this to the numerical model. Co-author Ming Yang provided much needed support in implementing the concept and testing the model. I carried out most of the data collection and analysis. I prepared the first draft of the manuscript and subsequently revised the manuscript, based on the feedback from co-authors and also a peer review process. The two co-authors assisted in developing the concept and testing the model, reviewed and corrected the model and results. They also contributed to the review and revision of the manuscript.*

### Abstract

With the increasing complexity of the hazardous process operation, potential accident modeling is becoming challenging. In process operation accidents, causation is a function of nonlinear interactions of various factors. Traditional accident models such as the fault tree represent cause and effect relationships without considering the dependency and nonlinear interaction of the causal factors.

This paper presents a new non-sequential barrier-based process accident model. The model uses both fault and event tree analysis to study the cause-consequence relationship. The dependencies and nonlinear interaction among failure causes are modelled using a Bayesian network (BN) with

various relaxation strategies. The proposed model considers six prevention barriers in the accident causation process: design error, operational failure, equipment failure, human failure and external factor prevention barriers. Each barrier is modeled using BN and the interactions within the barrier are also modeled using BN. The proposed model estimates the lower and upper bounds of prevention barriers failure probabilities, considering dependencies and non-linear interaction among causal factors. Based on these failure probabilities, the model predicts the lower and upper bounds of the process accident causation probability. The proposed accident model is tested on a real life case study.

**Keywords:** Accident Modelling, Risk Assessment, Accident prediction, Bayesian network analysis

## 4.1. Introduction

In recent times, chemical process industries (CPI) are dealing with highly hazardous chemicals at different stages of their process operations. The dynamic technological complexity of process systems which include equipment, management and organisation decisions, operators, operating conditions, external environmental conditions and their various interactions are major causes of accidents in process industries. This complexity has numerous dimensions; interactive complexity is on the increase in systems currently being built. Process systems now contain large amounts of dynamically interacting components. In the current complex system, humans interact with technology and produce an outcome due to their collaboration which cannot be accomplished either by technology or humans operating independently. Therefore, safe operation of the modern complex system demands a thorough understanding of interactions and interrelationships between, human, technical, environmental and organizational phases of the system (Qureshi 2008; Leveson 2004b).

Recent analysis of CPI accidents has shown an increase in the frequency of accidents in most regions of the world, probably due to these complex interactions (Kidam et al. 2014; Khan & Abbasi 1999). Process accidents are normally due to a chain or sequence of failure of events caused by failure of one or several physical components and abnormalities of process parameters (Tan et al. 2013).

Process accident models give detailed features of accidents and clearly express the relationship between causes and effects. They provide an adequate explanation of why accidents occur and they are a very useful technique for process risk assessment. Process accidents normally follow three steps: initiation, propagation and termination (Crowl & Louvar 2001) and any of these steps could lead to hazardous events.

Accident models systematically relate causes and consequences of the events and play a significant role in accident investigation and analysis. Accident models primarily tend to answer two major broad questions: (i) why accidents occur and (ii) how accidents occur. Classification of accident models can be done in several ways. Accident models are broadly categorised as either traditional or modern accident models. Traditional accident models are further sub-grouped into sequential and epidemiological models. They are primarily descriptive models that lack predictive capacity and emphasize mainly human, organisational and management factors. Modern accident models can be sub classified into three sub- categories: systematic, formal, and dynamic accident models ( Al-shanini et al. 2014; Qureshi 2008).

One principal limitation of these accident models is that they are usually case-specific, commonly descriptive, qualitative and merely conventional models that cannot utilize accident precursor data to develop prevention strategies. Those that have quantitative units had limitations of data scarcity and uncertainty. However, dynamic accident models have a great benefit of simplicity because of their sequential arrangement or layout and because non-linear interactions can be represented within the main framework. The dynamic accident model is predictive and uses real time precursor data to evaluate the likelihood of all available end- states (Al-shanini et al. 2014).

Kujath et al (2010) developed a process accident model for offshore oil production to prevent offshore process accidents using the concept of safety barriers. Five major prevention barriers were connected alongside the accident propagation path to prevent and mitigate the consequences of hydrocarbon release. Fault tree analysis was used to analyse the failure of prevention barriers, and consequences were analyzed using an event tree. The end state precursor data in the event tree analysis were used to update the failure probabilities of safety barriers via the Bayesian theorem. Despite the application of this model to the Piper Alpha (1988) and BP's Texas city refinery (2005) the model still exhibits some limitations, which are : (1) There is only provision for operational and technical failures; all other accident contributory factors such as human and organisational errors were not part of the model; and (2) Other accident initiating events such as an explosion were not considered (Rathnayaka et al. 2011).

In order to overcome the weakness in Kujath's model, Rathnayaka et al. (2011) provided an extension of this model by incorporating other factors ( i.e., management and organisational factors) that were neglected by Kujath into a new accident model called System Hazard Identification, Prediction and Prevention (SHIPP) methodology. All accident contributory factors were modeled into seven prevention barriers. In this model, accident precursor data were used to update the failure probabilities of every barrier with the Bayesian updating technique. The SHIPP model was validated for two LNG facilities effectively and the results obtained were highly promising (Rathnayaka et al. 2012; Rathnayaka et al. 2010).

However, in spite of the promising results obtained with the use of SHIPP methodology, the model still has some weaknesses that may affect the accuracy of the results obtained. These weakness are: (1) External hazards are not considered in the model. (2) The model presumed the causes of failure within safety barriers were independent, although in reality they are interdependent and this could grossly affect the results. (3) Provision was not made for other factors that were not accounted for in the fault tree model of prevention barriers. (4) Nonlinear interaction of various factors were not considered.

This paper proposes a novel non sequential barrier based accident model, in which interdependency and nonlinear interaction among accident contributory factors within safety barriers are modelled for process accidents. This work also proposes major influencing factors of process accidents. Considering dependencies and non-linear interaction among causal factors, the proposed model is capable of estimating the lower and upper boundary of prevention barrier failure probabilities. The remaining parts of this paper are organised as follows. Section 4.2 provides a brief description of basic characteristic of BNs. Section 4. 3 presents canonical models based on the assumption of independence of causal influence. Section 4. 4 presents the proposed accident model. Section 4. 5 demonstrates the application of the proposed model using the Richmond refinery accident. Section 4.6 presents the results and discussion. Finally, Section 4.7 provides the conclusion.

## 4.2. Bayesian Network

Bayesian networks (BN) are direct acyclic graph (DAG) with various nodes representing variables and arcs which represent direct dependencies among the variables. A BN usually consists of both qualitative and quantitative parts. The qualitative part is an acyclic directed graph naturally showing the causal structure of the domain; the other quantitative part denotes the joint probability distribution of its variables. All variables in a BN are adequately represented in a conditional probability table (CPT). A CTP provides complete specification of probabilistic interaction that has the capability to model any type of probabilistic dependence between a discrete node and its parents. The probabilities in the CTP denote the probabilities of each state given the state of the parent variable. However, if a variable in BN does not have parent variables, the CPT denotes the prior probability variable (Kraaijeveld & Druzdzel 2005).

A Bayesian network represents the joint probability distributions for a set of discrete random variables X, where X is given as

$$X = \left(X_1, X_2, ..., X_n\right) \tag{4.1}$$

where n is finite in this case. Equation (4.1) can be decomposed into products of conditional probability distributions for each of the variables provided their parent is known. In the case of a root node with no parents, prior probability is used instead. The joint probability distribution for a set of discrete random variables $X = \left(X_1, X_2, ..., X_n\right)$ can be calculated by taking the product of all the priors and their conditional probability distribution (Kraaijeveld & Druzdzel 2005). Mathematically this is given by

$$P(x_1, x_2, \dots, x\ n) = \prod_{i=1}^{n} P(x_i | pa_{(xi)}) \tag{4.2}$$

## 4.3. Canonical Probabilistic Models

Canonical models are advantageous because they make the construction of a probabilistic model easy and also reduce the computation time. One foremost challenge in using the BN model to model practical problems is the difficulty that arises in obtaining the numerical parameters that are required to fully quantify it. Discrete joint probability distributions are generally represented as CTPs, which are a collection of discrete probability distributions of a variable conditional on its given parents in the BN. The size of CPTs increases exponentially with the number of parents in a BN. Therefore, it is extremely difficult to build CTPs for variables having many parents. This is because these numerical parameters in CPTs are obtained from a data base or from human expertise (Oniśko et al. 2001; Diez & Druzdzel 2007).

One way of overcoming the challenge of obtaining these numerical probabilities is to apply the canonical models. Canonical models permit building of probability distribution from a fewer number of parameters (Bobbio et al. 2001). Noisy-OR and Leaky Noisy-OR are typical examples of a canonical model.

### 4.3.1. NOISY-OR GATE

The Noisy-OR gate belongs to the family of models widely referred to as independent of causal influences (ICI). A Noisy-OR gate is one form of canonical interaction that is extensively used in Bayesian networks.

The Noisy-OR gate is applicable when there are numerous possible causes $X_1$, $X_2$, $X_3$…, $X_n$ of an effect variable Y. The model has two assumptions: (1) Each of the causes $X_i$ has a probability pi strong enough to cause Y, when other causes are absent. (2) Each of the causes $X_i$ influences Y independently from each other. The noisy model requires specification of n parameters $p_1$, $p_2$…, $p_n$. $p_i$ is the probability that effect Y is true given that cause $X_i$ is true and all other causes $X_j$, $j \neq i$, are false (Oniśko et al. 2001). Therefore,

$$p_i = \Pr\left(y \middle| \bar{x}_1, \bar{x}_2, \ldots, x_i, \ldots, \bar{x}_{n-1}, \bar{x}_n\right) \tag{4.3}$$

The two outcomes of variable X are represented by $x$ and $\bar{x}$. The probability of y provided a subset $X_p$ of the $X_i$s that are true is given by the following formula, from which the complete CPT of Y conditional on its parents $X_1$, $X_2$…., $X_n$ can be derived.

$$\Pr\left(y \middle| X_P\right) = 1 - \prod_{i:X_i \in X_P} \left(1 - p_i\right) \tag{4.4}$$

The use of the model results in a substantial reduction in the number of probabilities needed to quantify the cause-effect interaction. The model only needs "n" probabilities whereas the unrestricted model needs $2^n$ probabilities (Heckerman & Breese 1996).

### 4.3.2. Leaky Noisy-OR

The Noisy-OR gate does not consider the situation where a subsystem could fail though all of its components are functional. Leaky Noisy-OR considers a situation where the effect variable can be true though all of its causes are false. The Leaky model presumes a positive probability called leaky probability (*l*). Leaky probability is the probability that effect Y will occur spontaneously though all its causes are false. The model is applicable where it is impossible to capture all potential causes that could make effect Y occur. The effect of leaky probability could be easily modeled by the influence between $X_i$ and Y that has changed due to the addition of an unknown Parent F (Bobbio et al. 2001; Wasyluk 2001; Zagorecki & Druzdzel 2004). Therefore, the leaky Noisy-OR gate formula that can be used to calculate the probability of Y given the subset $X_P$ of the $X_i$ which are true is

$$\Pr\left(Y \middle| X_P\right) = 1 - \left[\left(1 - l\right) \prod_{i:X_i \in X_P} \left(1 - p_i\right)\right] \tag{4.5}$$

## 4.4. The Proposed Accident Model methodology

This section describes vividly each step of the proposed accident model. Figure 4.1 presents the flow chart of the model.

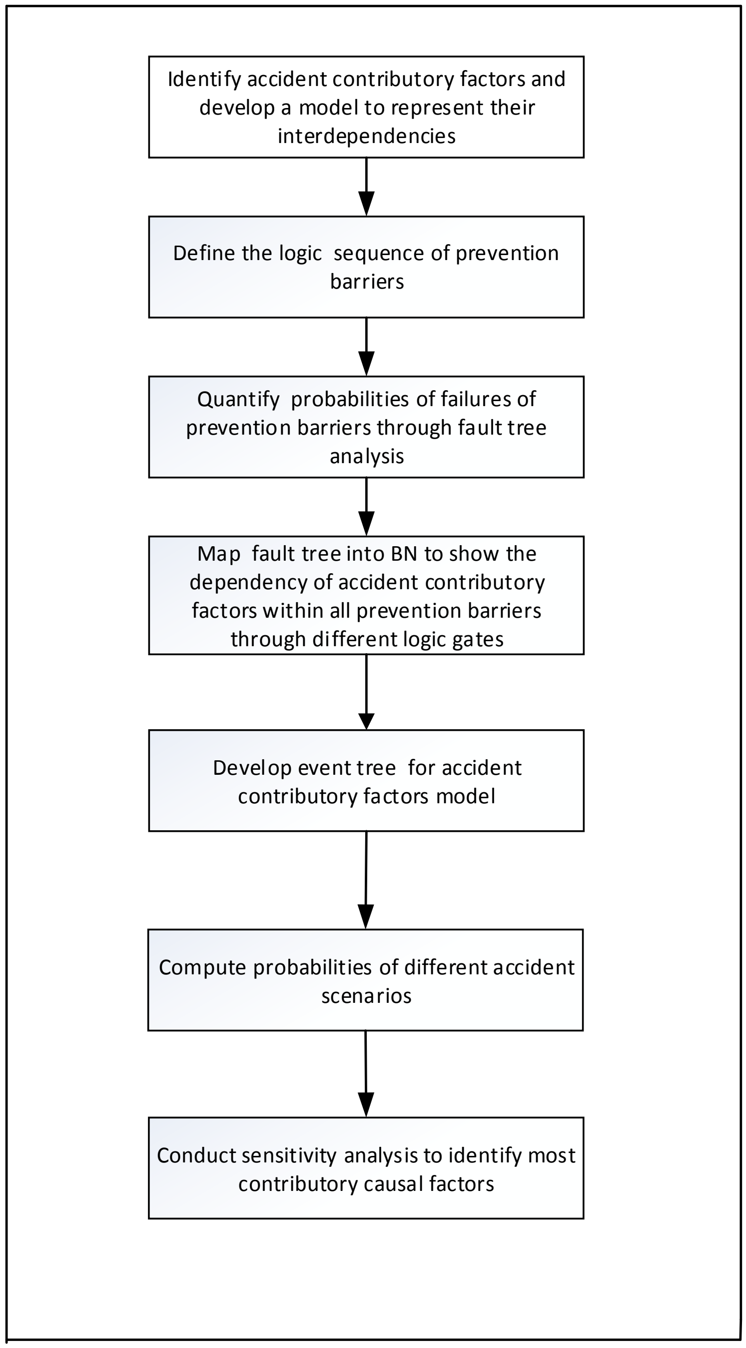



**Figure 4. 1. Flow chart for the proposed accident modeling steps.**

### 4.4.1. Accident contributory factors in the model

Generally, process accidents contributors can be categorised as human, management, technical (equipment), design and external factors (Kidam et al. 2014).

#### *4 .4.1.1. Design error*

One of the most common causes of accidents in process industries is design error. Evidence of this can be seen in different statistical studies of accident reports. Design error contributed 79% of accident cases that were analysed by Kidam & Hurme (2012). Normally, design error could be defined as features of a design which make it incapable of functioning according to its specification (Taylor 2007). A more applicable definition of design error for process accidents analysis is given by Kidam & Hurme (2012) : "… a design error is deemed to have occurred, if the design or operating procedures are changed after an incident has occurred.'' This definition encompasses both design and operating procedure changes after an accident has occurred.

#### *4. 4.1.2. Process equipment failure*

Process equipment failures are responsible for most process incidents. The deviations of process equipment from their original design objectives and normal operating conditions may result in catastrophic consequences (Mohammadfam et al. 2013; Khan & Abbasi 1999). Previous reviews of equipment failure related accidents show that the most common equipment to cause accidents in process industries are: reactors, storage tanks, pressure vessels, boilers and piping (Kidam & Hurme 2013).

#### *4.4.1. 3. Operational failure*

Operational failures are common in process industries. They entail all "disruptions and errors in materials, information and equipment'' (Adler-milstein et al. 2009) that originated from diverse causes including improper equipment maintenance, inspection, repair, and inadequate coordination among staff and management. The aftermath effect of an operational error could range from minor injuries to devastating catastrophes (Adler-milstein et al. 2009). Minor laxity of personnel during operation or maintenance could lead to an accident. For process accident analysis, a good working definition of operational failure is " any operational practice flaw that, if corrected, could have prevented the incident from occurring or would have significantly mitigated its consequences.'' (Bullemer & Laberge 2010) .

#### *4.4.1.4. Human failure*

Major accidents that have occurred in various process industries have been due to incorrect operation and maintenance errors. The Bhopal gas accident in 1984 and the Texas city refinery explosion are very good examples of major accidents that occurred as a result of human errors (Okoh & Haugen 2014).

Human error depends on several factors which are termed performance shaping factors (PSFs). These PSFs are classified into different categories: external, internal, psychological and physiological factors. External PSFs are factors related to the situation equipment characteristics and quality of the working conditions. Internal PSF factors are uniquely related to individual characteristics like skills, experience, motivation and mental capability. Psychological PSFs are factors which directly cause mental stress such as, task speed, task load and task type. Physiological factors are factors that cause physical stress such as hunger, extreme temperature, discomfort, thirst etc. (Abbassi et al. 2015).

#### *4.4.1.5. Organizational failure*

Organizational failure in most cases, is a result of organisational misalignment to realities (Sheppard & Chowdhury 2005). Organizational failure always contributes to accidents in process industries. The majority of technical failures can be traced back to organizational error (Rathnayaka et al. 2013).

#### *4.4.1.6. External factors*

External factors such as an earthquake, storm or lightning are potential sources of hazards for process industries. Hence there is a need for these factors to be adequately considered in process accident modelling ( Al-shanini et al. 2014). The interdependency of all these accident contributory factors is presented in Figure 4. 2.

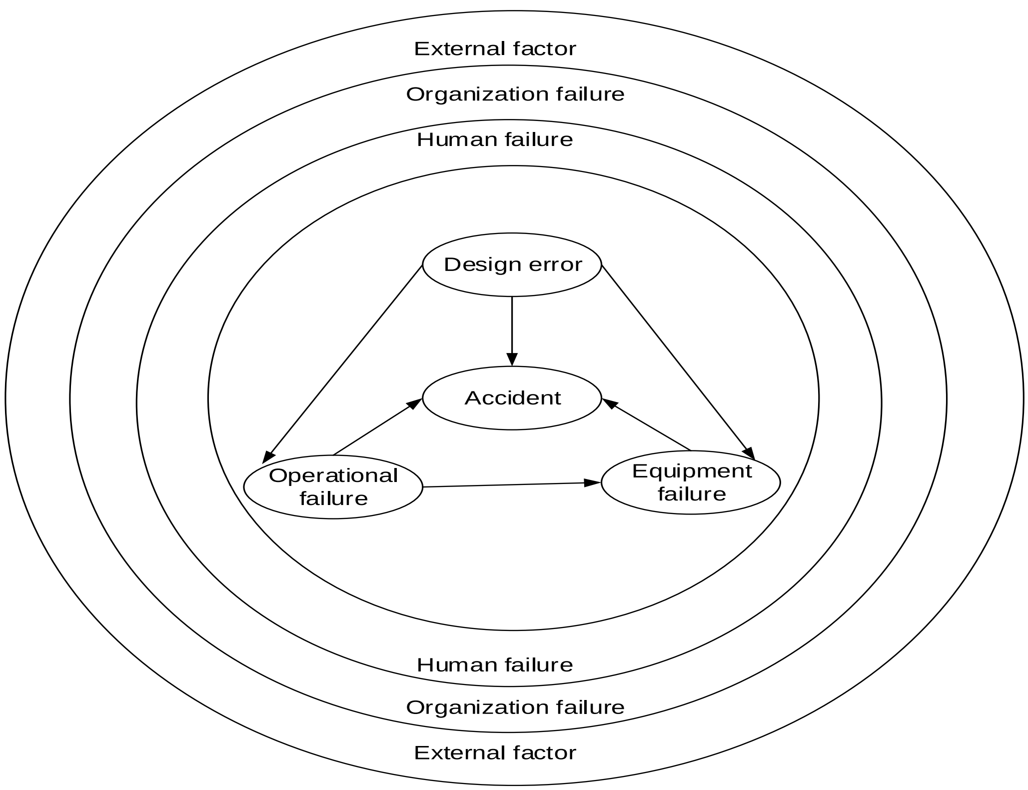



**Figure 4. 2. Interdependency of accident contributory factors in the main model.**

In this model, there are dependent relationships among design error, operational failure and equipment failure. Design error can lead to both operational and equipment failures. Operational failure can cause equipment failure. Similarly, design error, operational and equipment failures may result in a process accident independently. The other accident contributory factors in the model that could lead to an accident are: human failure, organisational failure and external factors.

### 4.4.2. Developing prevention barriers for the model

Prevention barriers play a significant role in mitigating the effect of any error that will result in an accident. The accident contributory factors are logically arranged into different prevention barriers along the accident pathway to avert or control the effects of a process accident. Fault tree and event tree analyses are used to represent the cause-consequence relationship for each of the prevention barriers.

### 4.4.3. Quantifying probabilities of failures of prevention barriers through fault tree analysis

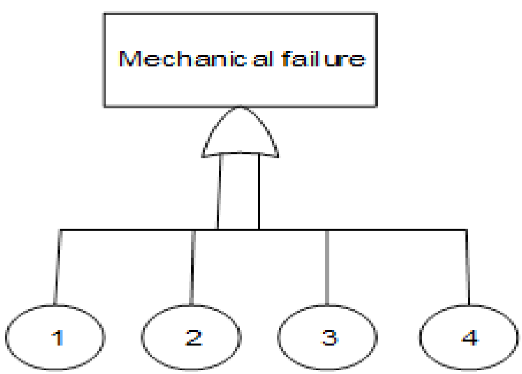
Fault tree analysis (FTA) is a well-developed tool for quantitative reliability and safety analysis of a complex system. FTA is a deductive graphical technique for detecting the possible causes of undesired events, popularly referred to as a top event. The top event normally denotes the main accident causing safety hazard (Khakzad et al. 2011).

Fault tree construction begins by placing the top event at the top of the tree. Every other possible way for this top event to occur is systematically constructed downwards until the primary events (root causes) causing the top event to occur are detected. The most commonly used gates are the AND-gate and OR-gate. FTA can be carried out in two basic ways: qualitatively and quantitatively. In qualitative analysis, a Boolean algebra expression of the top event is derived in terms of minimal cut-sets.

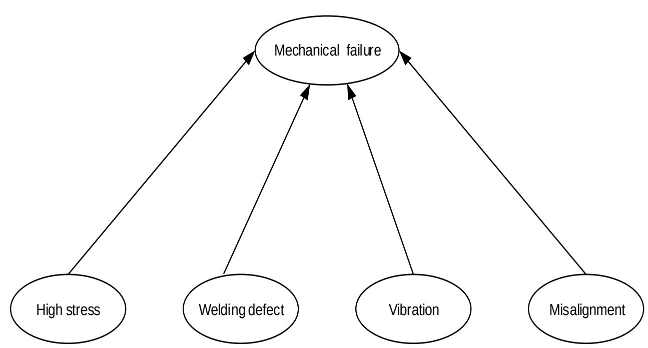
In quantitative analysis, the probability of occurrence of the top event in the fault tree is calculated based on the failure probabilities of basic events (Bobbio et al. 2001; Durga Rao et al. 2009; Khakzad et al. 2011).

Figure 4.3a gives an example of the fault tree that presents the causes of mechanical failure of a system. Based on failure probabilities given by Table 4.1, the failure probability of mechanical failure is calculated as follows:

$$1 - (1 - 0.001)(1- 0.025)(1- 0.003)(1- 0.1667) = 0.1907$$




(a)




(b)

**Figure 4. 3. Fault tree and Bayesian network for causes of mechanical failure.**

**Table 4. 1. Safe and failure probabilities of causes of mechanical failure.**

| No | Causes of mechanical failure | Failure probability | Safe probability |
|---|---|---|---|
| 1 | High stress | 0.001 | 0.999 |
| 2 | Welding defect | 0.025 | 0.975 |
| 3 | Vibration | 0.003 | 0.997 |
| 4 | Misalignment | 0.1667 | 0.8333 |

### 4.4.4 Mapping of the fault tree into the Bayesian network

The Fault tree (FT) can be mapped into the Bayesian network (BN) and the analysis performed with the FT can also be done with the BN, by using the inference in the BN. The mapping algorithm consists of both graphical and numerical tasks. In graphical mapping, the basic events, intermediate events and top event of the FT are converted respectively to, root nodes, intermediate nodes and the leaf node of the corresponding BN. The connections between the nodes in the BN remains the same as that of the corresponding FT. In numerical mapping, each root node is assigned the same prior probability as the corresponding basic event in the FT. The CPTs are formulated for both leaf and intermediate nodes. The CPTs need to be developed according to the logic gates used in the FT (Bobbio et al. 2001; Khakzad et al. 2011; Lampis & Andrews 2009). The purpose of mapping the fault tree into the BN is to allow the usage of different logic gates in the model: OR, Noisy-OR and Leaky Noisy-OR gates. The fault tree shown in Figure 4.3a was mapped into a Bayesian network given by Figure 4.3b.

It is common to use the OR-gate in BN models; therefore, its computational process is not discussed here. The following provides the procedure, based on which the probability of a top event (e.g., mechanical failure in Figure 4.3b) is computed when the Noisy-OR logic gate is applied in the BN model.

(1) Computing the safe (non-occurrence) probability of all parent nodes in the BN model

(2) Assigning the non-causation probability of all parent nodes in the CPT, based on expert opinion or data.

(3) Computing the conditional probabilities table.

(4) Using a conditional probability table with probabilities of the state of parent nodes (safe/false or failure/true) depending on the state involved, compute the probability of a top event.

Still following the previous example, based on step (1) to step (4), the probability of mechanical failure was calculated as shown in Table 4.2.

**Table 4. 2. Probability of mechanical failure for Noisy- OR gate.**

| State | High Stress | Welding defect | Vibration | Wrong specification | Causation probability of mechanical | Non causation probability of mechanical failure | Conditional probability of mechanical failue for different states |
|---|---|---|---|---|---|---|---|
| 1 | F | F | F | F | 0 | 1 | 0 *0.999 * 0.975 *0.997 * 0.8333 = 0 |
| 2 | F | F | F | T | 0.8 | 0.2 | 0.8 * 0.999 * 0.975 * 0.997 * 0.8333 = 1.29506E-01 |
| 3 | F | F | T | F | 0.7 | 0.3 | 0.7 * 0.999 * 0.975 * 0.003 * 0.8333 = 1.70447E-03 |
| 4 | F | F | T | T | 0.94 | 0.06 = 0.3*0.2 | 0.94 * 0.999 * 0.975 * 0.003 * 0.1667 = 4.57883E-04 |
| 5 | F | T | F | F | 0.75 | 0.25 | 0.75 * 0.999 * 0.025 * 0.997 * 0.8333 = 1.55619E-02 |
| 6 | F | T | F | T | 0.95 | 0.05 = 0.25*0.2 | 0.95 * 0.999 * 0.025 * 0.997 * 0.1667 = 3.9433E-03 |
| 7 | F | T | T | F | 0.925 | 0.075 = 0.25*0.3 | 0.925 * 0.999 * 0.025 * 0.003 * 0.8330 = 5.77524E-05 |
| 8 | F | T | T | T | 0.985 | 0.015 = 0.25*0.3*0.2 | 0.985 * 0.999 * 0.025 * 0.003 * 0.1667 = 1.23026E-05 |
| 9 | T | F | F | F | 0.65 | 0.35 | 0.65 * 0.001 * 0.975 * 0.997 * 0.8333 = 5.2652E- 04 |
| 10 | T | F | F | T | 0.93 | 0.07 = 0.35 *0.2 | 0.93 * 0.001 * 0.975 * 0.997 * 0.1667 = 1.50702E- 04 |
| 11 | T | F | T | F | 0.895 | 0.105 = 0.35 *0.2 | 0.895 * 0.001 * 0.975 * 0.003 * 0.8333 = 2.18148E-06 |
| 12 | T | F | T | T | 0.979 | 0.021 = 0.35*0.3*0.2 | 0.979 * 0.001 * 0.975 * 0.003 * 0.1667 = 4.77358 E-07 |
| 13 | T | T | F | F | 0.9215 | 0.0875 = 0.35*0.25 | 0.9215 * 0.001 * 0.025 * 0.997 * 0.8333 = 1.9139 6E-05 |
| 14 | T | T | F | T | 0.9825 | 0.0175 = 0.35*0.25*0.2 | 0.9825 * 0.001 * 0.025 * 0.997 * 0.1667 = 4.08229 E-06 |
| 15 | T | T | T | F | 0.9737 | 0.02625 = 0.35*0.25*0.3 | 0.9737 * 0.001 * 0.025 * 0.003 * 0.8333 = 6.08538 E-08 |
| 16 | T | T | T | T | 0.9947 | 5.25E-3 = 0.35*0.25*0.3*0.2 | 0.9947 * 0.001 * 0.025 *0.003 * 0.1667 = 1.24362 E-08 |
| | | | | | | | The probability of mechanical failure is the sum of all states = 0.15194 |

The following steps can be used to calculate the probability of top events when the Leaky Noisy-OR gate is used in a BN model.

(1) Assigning a leak probability.

(2) Computing the safe (non-occurrence) probability of all parent nodes in the BN.

(3) Assigning the non-causation probabilities of all parent nodes in the CPT, based on expert opinion or data.

(4) Computing the conditional probabilities table.

(5) Using the conditional probability table with probabilities of the state of parent nodes (safe/false or failure/true) depending on the state involved, compute the probability of a top event.

Still using the previous example for the purpose of illustration (the leak probability used here is 0.01). Based on the process described above, the probability of mechanical failure was calculated as shown in Table 4.3.

**Table 4 .3. Probability of mechanical failure for Leak Noisy- OR gate.**

| State | High stress | Welding defect | Vibration | Misalignment | Causation probability of mechanical failure | Non causation probabilty of mehanical failure | Conditonal Probability of mechanical failure for different states |
|---|---|---|---|---|---|---|---|
| 1 | F | F | F | F | 0.01 | 0.99 | 0 .01 * 0.999 * 0.975 * 0.997 * 0.8333 = 8.09220E-03 |
| 2 | F | F | F | **T** | 0.802 | 0.198 **= 0.2 ***0.99 | 0.802 * 0.999 * 0.975 * 0.997 * 0.1667 = 1.29830E-01 |
| 3 | F | F | **T** | F | 0.703 | 0.297 = **0.3** * 0.99 | 0.703 * 0.999 * 0.975 * 0.003 * 0.8333 = 1.71178E-03 |
| 4 | F | F | T | T | 0.9406 | 5.94 E-02 = 0.3*0.2*0.99 | 0.9406 * 0.999 * 0.975 * 0.003 * 0.1667 = 4.58175E-04 |
| 5 | F | **T** | F | F | 0.7525 | 0.2475 = **0.25** *0.99 | 0.7525 * 0.999 * 0.025 * 0.997 * 0.8333 = 1.56137E-02 |
| 6 | F | T | F | T | 0.9505 | 4.95E-02 = 0.25*0.2*0.99 | 0.9505 * 0.999 * 0.025 * 0.997 * 0.1667 = 3.94537E-03 |
| 7 | F | T | **T** | F | 0.9257 | 7.43E-02 = 0.25*0.3*0.99 | 0.9257 * 0.999 * 0.025 * 0.003 * 0.8333 = 5.77960E-05 |
| 8 | F | T | T | T | 0.9851 | 1.485E-02 = 0.25*0.3*0.2*0.99 | 0.9851 * 0.999 * 0.025 * 0.003 * 0.1667 = 1.23038E-05 |
| 9 | **T** | F | F | F | 0.6535 | 0.3465 = **0.35** * 0.99 | 0.6535 * 0.001 * 0.975 * 0.997 * 0.8333 = 5.29354E- 04 |
| 10 | T | F | F | T | 0.9307 | 0.0693 = 0.35*0.2*0.99 | 0.9307 * 0.001 * 0.975 * 0.997 * 0.1667 = 1.50815E- 04 |
| 11 | T | F | T | F | 0.89605 | 0.1039 = 0.35*0.3*0.99 | 0.89605 * 0.001 * 0.975 * 0.003 * 0.8333 = 2.18403E-06 |
| 12 | T | F | T | T | 0.9792 | 0.0207 = 0.35*0.3*0.2*0.99 | 0.9792 * 0.001 * 0.975 * 0.003 * 0.1667 = 4.77455 E-07 |
| 13 | T | T | F | F | 0.9133 | 0.0866 = 0.35*0.25*0.99 | 0.9133 * 0.001 * 0.025 * 0.997 * 0.8333 = 1.89692E-05 |
| 14 | T | T | F | T | 0.9826 | 0.0173 = 0.35*0.25*0.2*0.99 | 0.9826 * 0.001 * 0.025 * 0.997 * 0.1667 = 4.08270E-06 |
| 15 | T | T | T | F | 0.974 | 0.02598 = 0.35*0.25*0.3*0.99 | 0.9740 * 0.001 * 0.025 * 0.003 * 0.8333 = 6.08725 E-08 |
| 16 | T | T | T | T | 0.9948 | 5.1975E-03 = 0.35*0.25*0.3*0.2*0.99 | 0.9948 * 0.001 * 0.025 * 0.003 * 0.1667 = 1.24374 E-08 |
| | | | | | | | The probabiity of mechanical failure is sum of all states = 0.16042 |

### 4.4.5. Mapping a fault tree into a BN to represent the dependency of accident contributory factors within all prevention barriers

Any particular form of failure in any of the prevention barriers in the accident model is due to complex interactions of various factors. In most cases, some of these factors interact with one another in a complex manner. The dependency of these factors within each of the prevention barriers is represented by the BN in Figures 4. 4 - 4.9. Firstly, the failure probabilities of each factor (node) within each prevention barrier in the BN were evaluated without considering the dependencies within these factors using a fault tree, BN(Noisy-OR) and BN(Leaky noisy-OR ) logics.

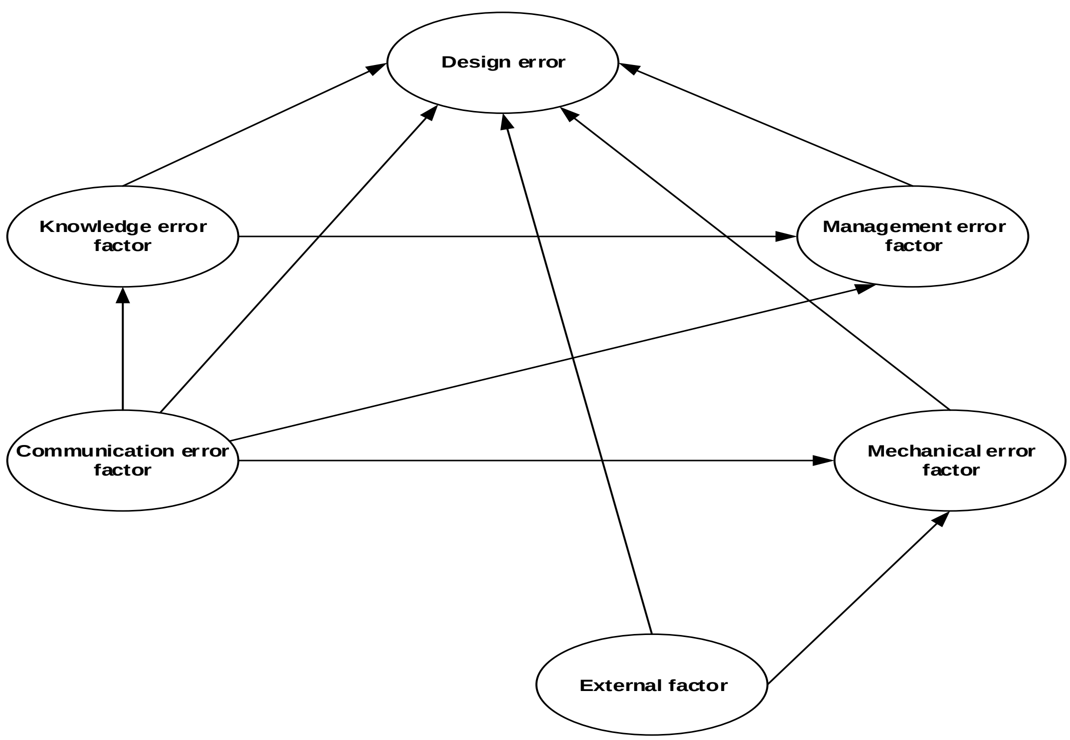



**Figure 4 .4. Interdependency of accident contributory factors in design preventive barrier.**

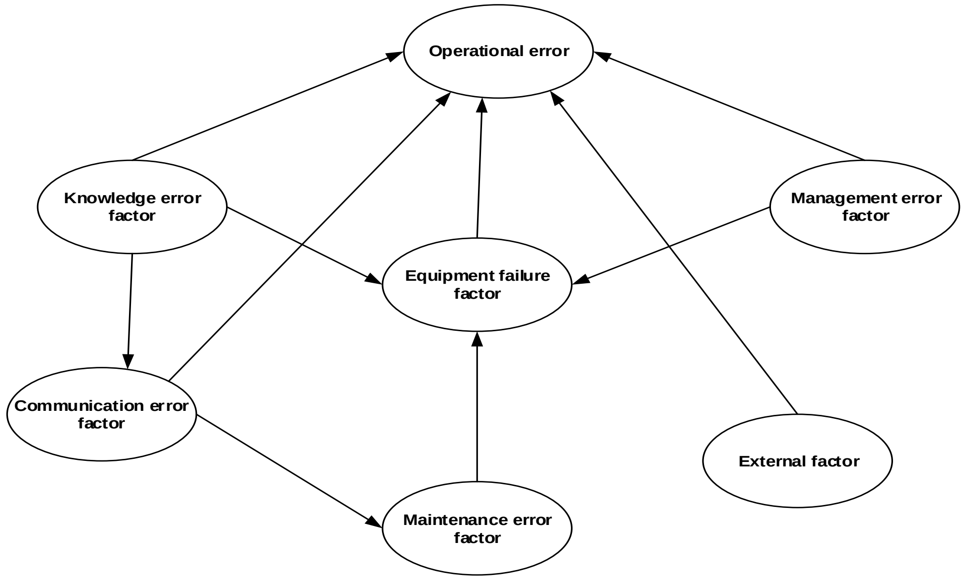



**Figure 4.5. Interdependency of accident contributory factors in operational prevention barrier.**

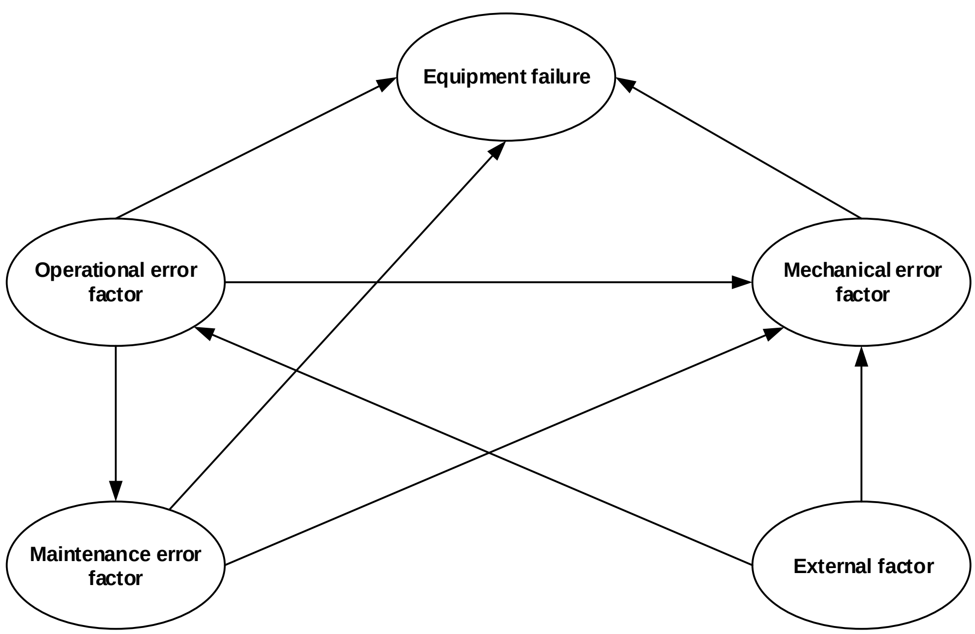



**Figure 4.6. Interdependency of accident contributory factors in Equipment failure prevention barrier.**

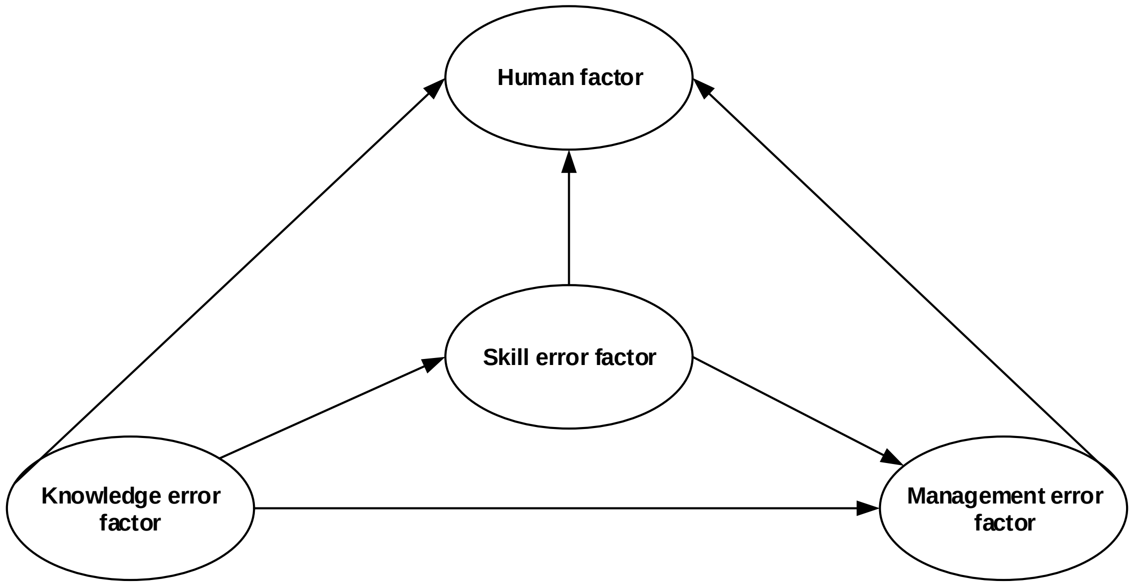



**Figure 4.7. Interdependency of accident contributory factors in human factor prevention barrier.**

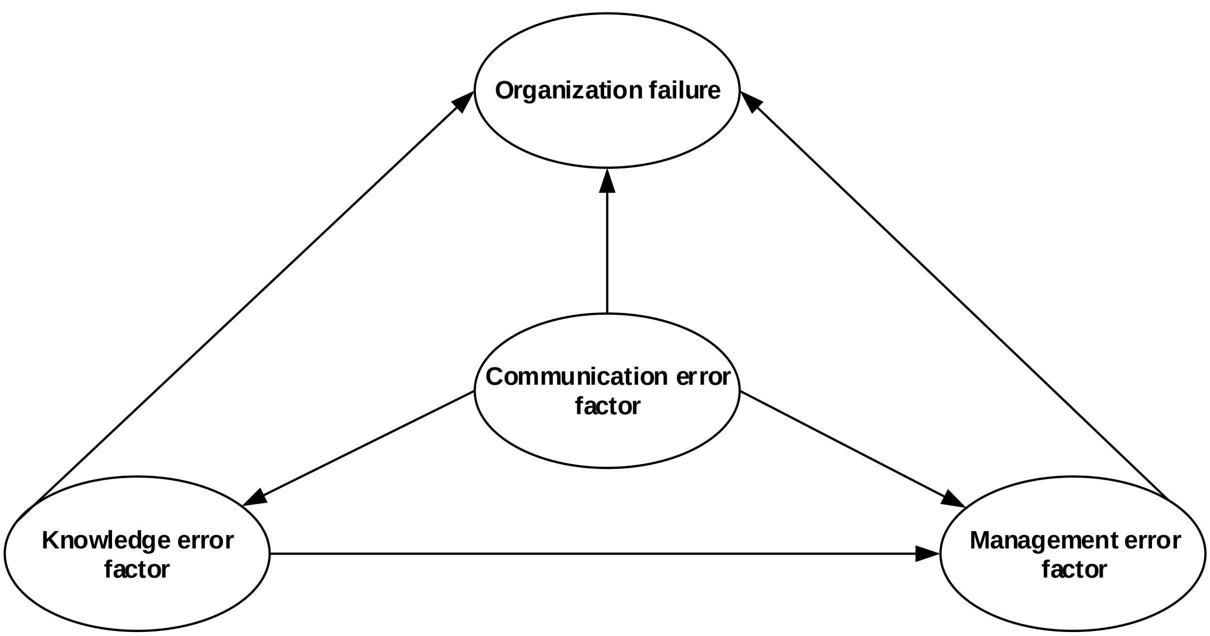



**Figure 4. 8. Interdependency of accident contributory factors in organisational failure prevention barrier.**

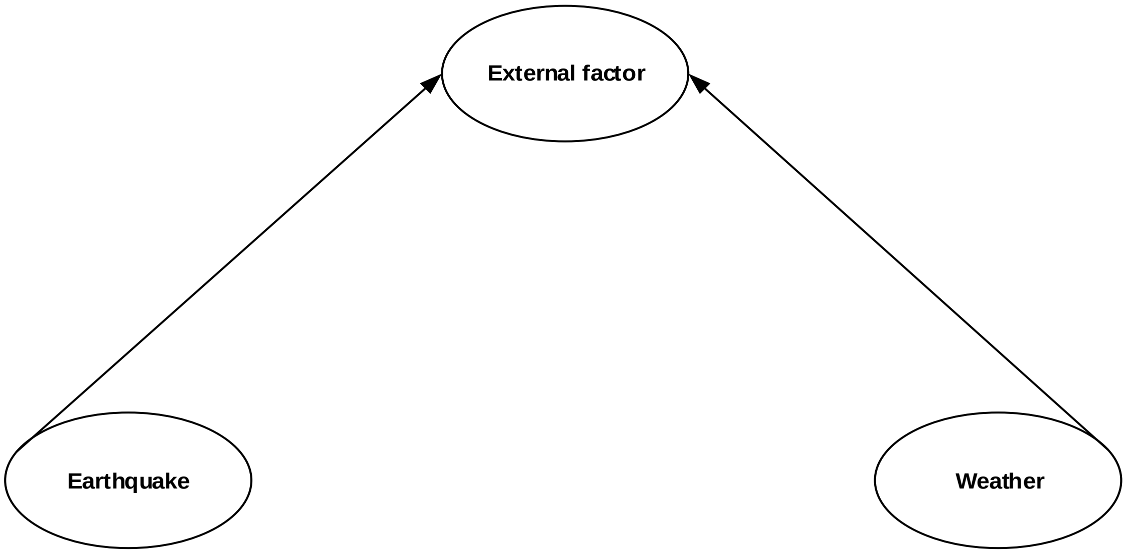



**Figure 4.9. Interdependency of accident contributory factors in external factor prevention barrier.**

Secondly, considering the dependencies of factors within each prevention barrier, different logic gates (OR, Noisy-OR, Leaky noisy-OR) are used in the BN to evaluate the failure probabilities of each prevention barrier using different failure probabilities of factors (nodes) within each prevention barrier calculated independently by different logics. The failure probabilities of each factor node (except root node) are used as a leaky probability when the leaky Noisy-OR gate is applied to the BN. The leaky probability of the top event in all barrier is 0.01. Failure probabilities of all factors (nodes) were used to evaluate the failure probability of prevention barriers when the leaky Noisy-OR gate was applied to the BN model. This overcomes the setback encountered in applying both OR and Noisy-OR logic in the BN, where only failure probabilities of root nodes are used to evaluate the failure probability of prevention barriers.

### 4.4.6. Developing an event tree for the accident contributory factor model

Event tree analysis is an inductive method that is mostly used in process industries accident analysis to represent the incident scenarios. Event tree analysis begins with a definite initiating event and ends up with all the possible consequences of the initiating event, normally referred to as the end state consequences of the event tree. An Event tree clearly displays the probabilities of success and failure of safety barriers and the progression of a specified initiating event to numerous potential scenarios (Nývlt & Rausand 2012). The various barriers in the event tree are represented by two distinct branches, one representing success and the other representing failure of that particular barrier.

The event tree model for the proposed accident model is shown in Figure 4.10.

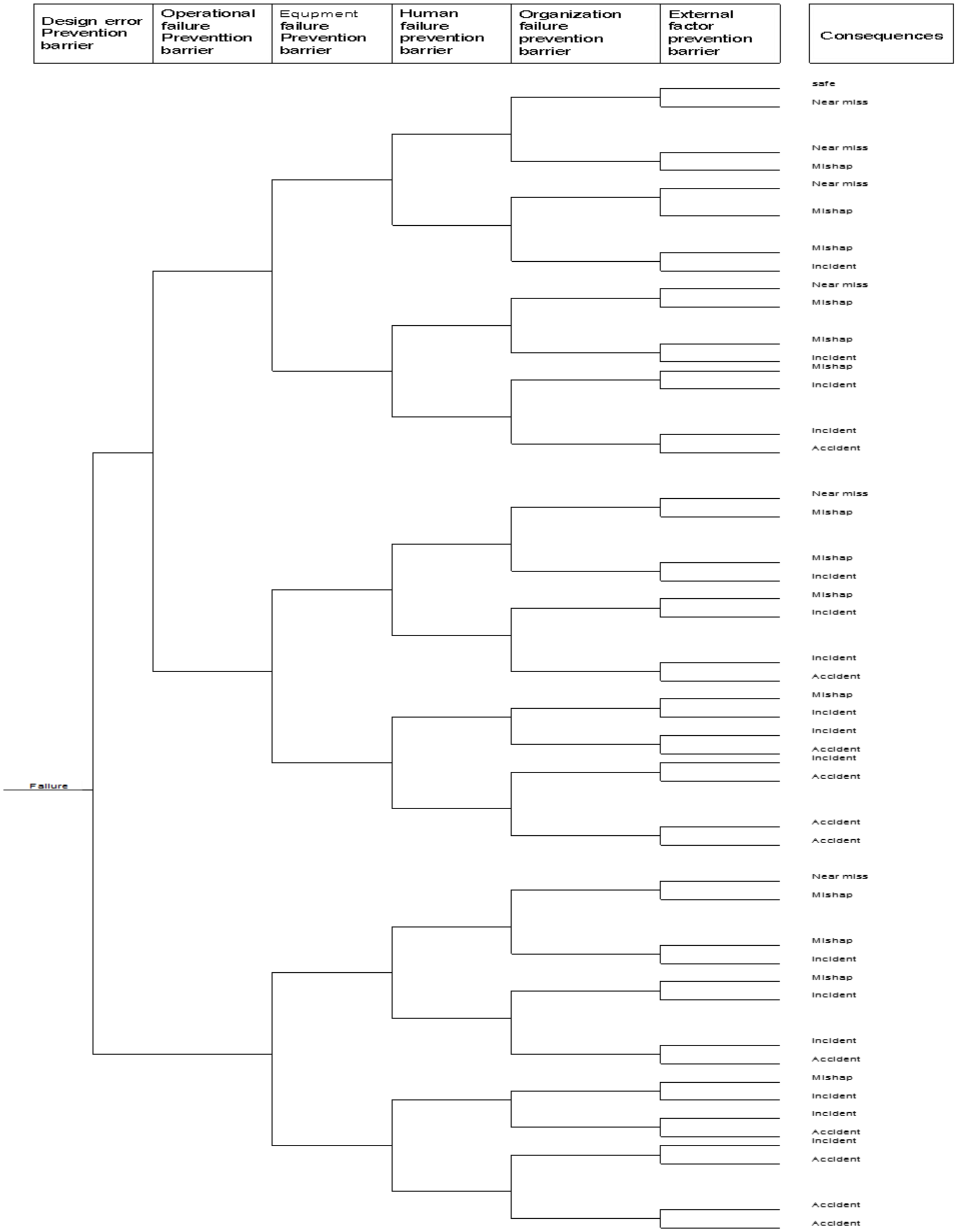



**Figure 4.10. Event tree for the accident model.**

After the design prevention barrier is triggered by a specified initiating event, the event sequence is propagated through success or failure branches of the operational, equipment, human, organisational and external factors prevention barriers. The event sequence leads to all possible end-state consequences in the event tree. Five major end state consequences have been identified: safe, near miss, mishap, incident and accident. Four of these end states are abnormal events and their description is given by Rathnayaka et al. (2011).

## 4.5. Case study

In this section, the Richmond refinery accident in the U.S is systematically analyzed using the proposed model. The model subsequently reveals how the accident could have been averted, supposing the relevant prevention barriers were kept safe.

The Chevron Richmond refinery in the United States of America experienced an appalling pipe rupture in the # 4 crude unit on August 6, 2012. The accident resulted from one of numerous process streams popularly known as a “4-sidecut” exiting the refinery’s C-1100 Crude Unit atmospheric column. The pipe rupture occurred on a 52- inch long component of the 4- sidecut 8-inch line.

The rupture pipe released flammable, high temperature light oil gas that was flowing through it at a rate of approximately 10800 barrels per day. The released process fluid ignited two minutes after the release (CSB 2015). The schematic diagram of the C-1100 Crude unit atmospheric column and upstream process equipment is shown in Figure 4.11.

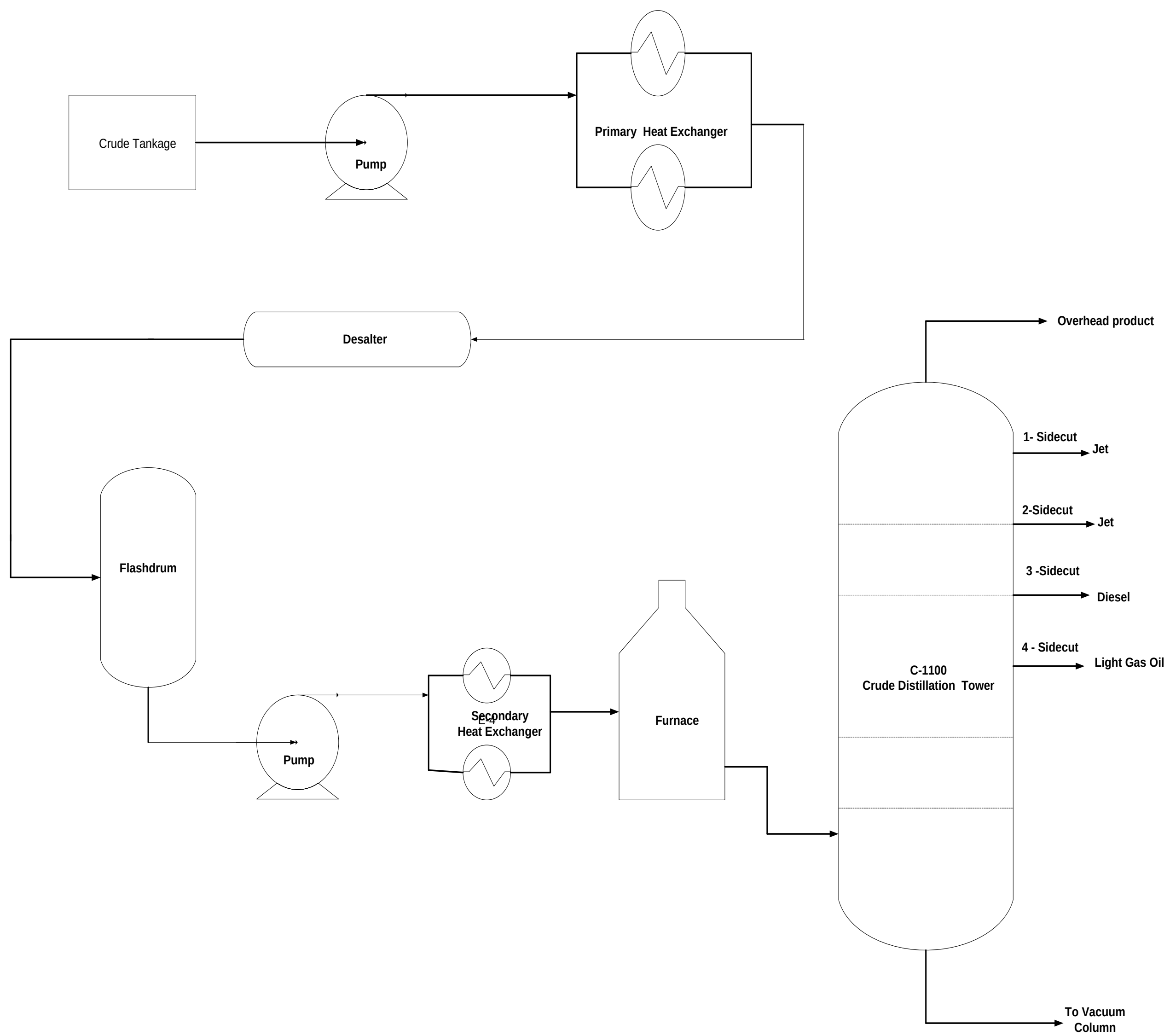



**Figure 4.11. Schematic of diagram of C-1100 Crude unit atmospheric column and upstream process equipment. (CSB 2015).**

The U.S Chemical Safety Board (CSB) carried out an extensive investigation of the incident. The CSB highlighted different technical issues and safety system deficiencies in their report. Based on information available in the CSB report, this accident has been modelled using the proposed methodology.

The design error prevention barrier was analysed first, in the proposed model. Various events that contributed to the failure of the design error prevention barrier were analysed with the fault tree. Subsequently, the operational failure prevention barrier, equipment failure prevention barrier, human failure prevention barrier, organisational failure prevention barrier and external factor prevention barrier were analyzed to identify the events that lead to failure of these barriers. Table 4.4 shows the event description and failure probabilities of the basic events for the design error prevention barrier in the model. The reliability data used in fault tree analysis were obtained from various journals ( Al-shanini et al. 2014; Rathnayaka et al. 2010; Abimbola et al. 2014; Rathnayaka et al. 2013; Rathnayaka et al. 2012) , Bercba's work( 1978) and expert opinions where the data is not available.

**Table 4. 4. Basic event failure probability for Design error prevention barrier.**

| Event | Event Description | Assigned probability |
|---|---|---|
| **1** | High operating temperature | $2.50 \times 10^{-2}$ |
| **2** | Wrong construction material for pipe (carbon) | $1.00 \times 10^{-2}$ |
| **3** | Poor design for Sulfidic corrosion | $1.00 \times 10^{-1}$ |
| **4** | Un even flow in the pipe | $1.00 \times 10^{-3}$ |
| **5** | Wrong pipe fittings | $1.00 \times 10^{-3}$ |
| **6** | Non explosion proof | $5.00 \times 10^{-3}$ |
| **7** | Inadequate management practice | $2.50 \times 10^{-2}$ |
| **8** | Insufficient funding | $3.00 \times 10^{-4}$ |
| **9** | Wrong work culture | $1.00 \times 10^{-3}$ |
| **10** | Welding defect on Pipe | $6.60 \times 10^{-2}$ |
| **11** | Pipe erosion and cracking | $1.00 \times 10^{-3}$ |
| **12** | Wrong chemical resistant specification | $1.00 \times 10^{-4}$ |
| **13** | Wrong piping thickness | $1.00 \times 10^{-3}$ |
| **14** | Operating conditions not specified | $5.00 \times 10^{-2}$ |
| **15** | Improper labelling | $5.00 \times 10^{-4}$ |
| **16** | Power failure | $1.5. \times 10^{-5}$ |
| **17** | Working conditions | $1.00 \times 10^{-4}$ |

The fault tree for the design error prevention barrier in the model is shown in Figure 4.12.

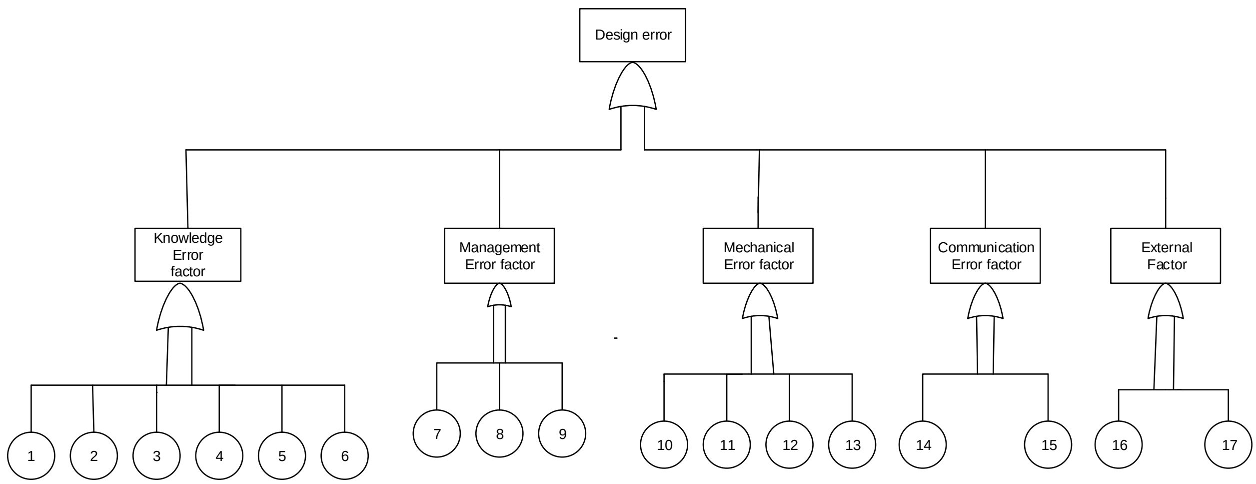



**Figure 4.12. Fault tree analysis of Design error prevention barrier.**

To reduce the number of figures and tables in this paper, reliability data of basic events and associated fault trees for other prevention barriers are not shown.

## 4.6. Results and discussion

The failure probabilities of all prevention barriers estimated using fault tree analysis without considering the dependency among the contributory factors that lead to failure of particular barriers are shown in Table 4. 5.

**Table 4. 5. Failure probability of prevention barriers from fault tree calculations.**

| Prevention barrier | Failure probability |
|---|---|
| Design error prevention barrier (DPB) | 0.2567 |
| Operational failure prevention barrier (OPB) | 0.2700 |
| Equipment failure prevention barrier (EPB) | 0.2628 |
| Human failure prevention barrier (HPB) | 0.2870 |
| Organisational failure prevention barrier (OrPB) | 0.2959 |
| External factor prevention barrier (ExPB) | 0.0171 |

From the analysis of failure probabilities of these prevention barriers, it can be deduced that it is the active failure of these relevant preventive barriers combined with latent conditions and concurrent management failures that led to the accident. In the Richmond refinery accident, all prevention barriers proposed in this model failed except the external factor barrier. The safety of these prevention barriers had been seriously compromised at the Richmond Refinery over time. It is the simultaneous failure of these barriers that ultimately led to the accident. If these prevention barriers had been kept intact by the management of Richmond refinery, the accident would have been prevented or mitigated. It was observed that organisational; human and operational failure prevention barriers have high failure probabilities respectively; the acute failure of these barriers contributed significantly to the accident.

The failure probabilities of different factors within the prevention barriers that contribute to the failure of these prevention barriers were estimated independently using different logic gates.

The computational process has been describd in Sections 4.3-4.5. Table 4.6 shows these failure proababities.

However, when dependency among various factors that lead to the failure of a particular barrier is modelled in the BN using different logic gates (OR, Noisy-OR and Leaky Noisy-OR), it is possible to determine both lower and upper failure probability bounds for each of the prevention barriers. The failure probabilities of different factors given in Table 4. 6 were used to compute the failure probabilities of the BN for different prevention barriers using different logic gates.

**Table 4.6. Failure probability of factors contributing to failure of all prevention barriers, using different logic dependency.**

| Prevention barrier | Logic Gate | | | | | | |
|---|---|---|---|---|---|---|---|
| DPB | | Knowledge error factor | Management error factor | Mechanical error factor | Communication error factor | External factor | |
| | Fault tree (OR) | 0.1373 | 0.0263 | 0.0679 | 0.0504 | 1.10E-04 | |
| | BN(Noisy-OR) | 0.0842 | 0.0145 | 0.0443 | 0.0327 | 7.97E-05 | |
| | BN(Leaky-Noisy-OR) | 0.0934 | 0.0243 | 0.0538 | 0.0422 | 0.0100 | |
| OPB | | Knowledge error factor | Equipment error factor | Maintenance error factor | External factor | Communication error factor | Management error factor |
| | Fault tree (OR) | 0.1069 | 0.1748 | 0.0017 | 0.0021 | 0.0023 | 0.0032 |
| | BN(Noisy-OR) | 0.0699 | 0.1248 | 0.0012 | 0.0012 | 0.0013 | 0.0020 |
| | BN(Leaky-Noisy-OR) | 0.0792 | 0.1336 | 0.0112 | 0.0112 | 0.0112 | 0.0120 |
| EPB | | Mechanical error factor | External factor | Operational error factor | Maintenance error factor | | |
| | Fault tree (OR) | 0.0810 | 0.0055 | 0.1126 | 0.0911 | | |
| | BN(Noisy-OR) | 0.0535 | 0.0030 | 0.0772 | 0.0682 | | |
| | BN(Leaky-Noisy-OR) | 0.0629 | 0.0130 | 0.0864 | 0.0776 | | |
| HPB | | Knowledge error factor | Management error factor | Skill error factor | | | |
| | Fault tree (OR) | 0.1630 | 0.1471 | 0.0012 | | | |
| | BN(Noisy-OR) | 0.1062 | 0.1047 | 0.0007 | | | |
| | BN(Leaky-Noisy-OR) | 0.1152 | 0.1137 | 0.0107 | | | |
| OrPB | | Knowledge error factor | Communication error factor | Management error factor | | | |
| | Fault tree (OR) | 0.0151 | 0.0037 | 0.2824 | | | |
| | BN(Noisy-OR) | 0.0093 | 0.0021 | 0.1955 | | | |
| | BN(Leaky-Noisy-OR) | 0.0192 | 0.0121 | 0.2035 | | | |
| ExPB | | External factor | | | | | |
| | Fault tree (OR) | 0.0169 | | | | | |
| | BN(Noisy-OR) | 0.0057 | | | | | |
| | BN(Leaky-Noisy-OR) | 0.0156 | | | | | |

From Table 4.7, it can be seen that the lower bound failure probability for all prevention barriers exists, when the Noisy-OR logic gate was applied in the BN.

**Table 4.7. BN dependency failure probabilities of prevention barriers using different logic dependency.**

| | Design error (DPB) | | | Operational failure (OPB) | | | Equipment failure (EPB) | | | Huaman failure (HPB) | | | Organisational failure(OrPB) | | | External factor(ExPB) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | OR | Noisy | Leaky | OR | Noisy | Leaky | OR | Noisy | Leaky | OR | Noisy | Leaky | OR | Noisy | Leaky | OR | Noisy |
| Fault tree (values) | 0.0501 | 0.0264 | 0.0921 | 0.1114 | 0.0790 | 0.1435 | 0.0050 | 0.0021 | 0.1152 | 0.1630 | 0.1177 | 0.1746 | 0.0040 | 0.0022 | 0.0790 | 0.0170 | 0.0120 |
| BN(Noisy OR values) | 0.0330 | 0.0174 | 0.0590 | 0.0727 | 0.0520 | 0.0992 | 0.0030 | 0.0012 | 0.0818 | 0.1060 | 0.0766 | 0.1195 | 0.0020 | 0.0011 | 0.0541 | 0.0059 | 0.0042 |
| BN(Leaky Noisy OR values) | 0.0518 | 0.0264 | 0.0746 | 0.1000 | 0.0663 | 0.1224 | 0.0130 | 0.0055 | 0.0957 | 0.1150 | 0.0831 | 0.1342 | 0.0120 | 0.0068 | 0.0658 | 0.0158 | 0.0111 |

*The probabilities shown in the column named OR, Noisy and Leaky were computed when the OR, Noisy-OR and leaky Noisy-OR gates were used in the BN. The row named fault tree values, BN (Noisy-OR), BN(Leaky Noisy-OR) are failure probabilities values given in table 4.6, which are substituted respectively based on the logic being considered in the BN.

This is because uncertainty of the conditional dependency between the linked nodes in the BN was considered so that lower conditional probabilities were assigned compared to the case in which the OR gate was used. The upper bound failure probability of all the prevention barrier exists when the Leaky Noisy-OR gate was applied in the BN.

The lower and upper bounds accident occurrence probability obtained from the event tree for the accident model are shown in Table 4. 8.

**Table 4. 8. Accident occurrence probability.**

| Logic | Accident occurrence probability | |
|---|---|---|
| | Lower bound | Upper bound |
| Fault tree | 6.93E-07 | 5.19E-04 |
| BN (Noisy-OR ) | 5.83E-08 | 1.00E-04 |
| BN (Leaky Noisy-OR ) | 1.37E-06 | 2.45E-04 |

*The probabilities shown in the column named lower and upper bounds were computed using the probabilities shown in the column named Noisy and leaky condition in Table 4.7.

To compute both the lower and upper bounds of accident probabilities, the lower and upper bounds of prevention barrier failure probabilities were used in the event tree to compute the consequences. Table 4.8 shows that a narrower span of probability estimation was obtained when dependency among causal factors was considered. This indicates the importance of modeling of these dependencies in an accident modelling approach and highlights the advantage of using BN instead of the fault tree. Table 4.8 also shows that the estimated range of accident probability changed significantly when the leaky Noisy-OR gate was used instead of the Noisy-OR. This demonstrates the importance of choosing an appropriate logic dependency to represent the actual causal relationship between two factors.

This would become even more important when insufficient or scarce data or information is available for the accident analysis. The accident occurrence probability of an accident in each case is the summation of all accident occurrence probabilities in the event tree. To evaluate accident occurrence probability, when the leaky noisy gate was applied among the causal factors of each barrier. The failure probability of OR gate values was used for the external barrier due to the lack of dependency relationships among the causal factors in this barrier and so the leaky Noisy-OR gate could not be applied to this barrier.

The sensitivity analysis for the design error prevention barrier in the model is shown in Figure 4.13.

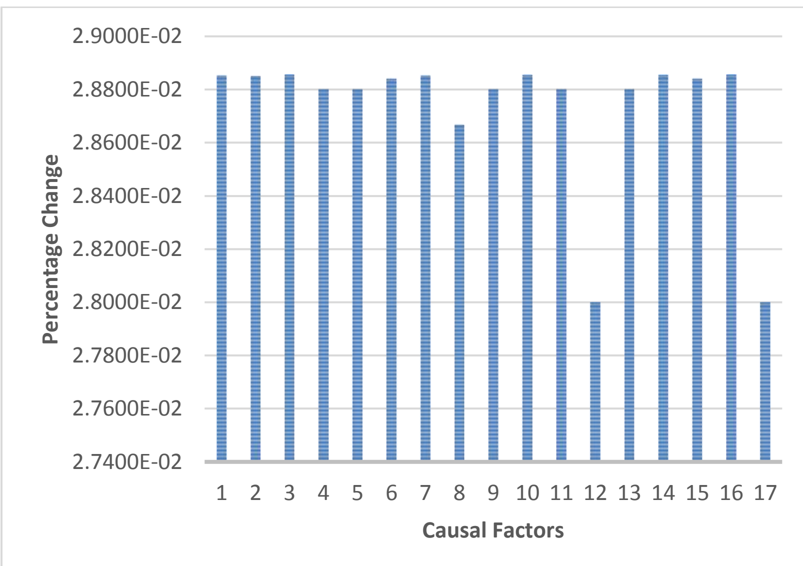



**Figure 4. 13. Sensitivity analysis of design error prevention barrier (Causal factors are defined as shown in Table 4.4).**

The causal factors are represented with numbers as defined in Table 4.4. It is observed that virtually all the causal factors in the DPB except two factors (i.e., wrong chemical resistant specification and wrong working conditions) are significant contributors to the failure of design error prevention barrier. Preventive actions must be placed on all these causal factors in order to eliminate or reduce drastically the negative influence of all these causal factors on these barriers. Systematic and timely analysis of these causal factors is of paramount importance to enhance the reliability of these prevention barriers and subsequently prevent the re- occurrence of the accident.

This analysis has demonstrated that BN is an effective technique for estimating the contribution of different conditional dependencies and nonlinear interaction among accident contributory factors within safety barriers. Conditional dependencies can be accurately modelled in the BN by means of a direct causal arc among various dependent variables with various relaxation strategies. The modelling of nonlinear interactions and dependencies with the BN using relaxation strategies has provided the opportunity to estimate accident probability considering the uncertainty dependency among the contributory factors of a prevention barrier. The predicted accident occurrence probabilities based on this model will provide valuable information in the development of accident prevention strategies based on an interval estimation of the risk.

## 4.7. Conclusions

This paper presented a new process accident model with emphasis on interdependency of contributory factors that lead to the failure of a particular prevention barrier. Six barriers were defined to prevent process accidents before they escalate into catastrophic events. The effectiveness of the proposed model was partially validated through the application of the model to the Richmond refinery accident. The BN model is capable of modelling the dependencies among these accident contributory factors. The application of Noisy-OR and leaky Noisy-OR gates helps to represent the uncertainties of the probabilities that are used in the CPTs of the BN model. Consequently, the proposed model is able to provide the lower and upper boundary of the failure probability of a process accident. The accident model provides a mechanism for predicting a process accident based on the interdependency and nonlinear interaction of contributory factors.

Process monitoring data is needed to effectively implement this accident model; with process monitoring data, the model can quantitatively estimate the dynamic risk profile that will greatly guide dynamic decision making. The use of predicted accident probabilities based on this model will help to take early corrective actions to prevent process accidents and developed effective process safety management plan.

## Acknowledgement

The authors thankfully acknowledge the financial support provided by, Natural Science and Engineering Council of Canada and Canada Research Chair (CRC) Program.

## References

Abbassi, R. et al., 2015. An integrated method for human error probability assessment during the maintenance of offshore facilities. *Process Safety and Environmental Protection*, 94(i), pp.172–179.

Abimbola, M., Khan, F. & Khakzad, N., 2014. Dynamic safety risk analysis of offshore drilling. *Journal of Loss Prevention in the Process Industries*, 30, pp.74–85. A

Adler-milstein, J., Singer, S.J. & Toffel, M.W., 2009. Operational Failures and Problem Solving : An Empirical Study of Incident Reporting Operational Failures and Problem Solving : *Public Health*.

Al-shanini, a., Ahmad, A. & Khan, F., 2014. Accident modelling and safety measure design of a hydrogen station. *International Journal of Hydrogen Energy*, 39(35), pp.20362–20370.

Al-shanini, A., Ahamad, A. & Khan, F., 2014. Accident modelling and analysis in process industries. *Journal of Loss Prevention in the Process Industries*, 32, pp.319–334.

Bercba, F.G., 1978. *Title Probabilities of blowouts in canadian Arctic waters*, Fisheries and enviroment Canada. Canada:Environmental Protection Service. Economic and Technical Review, Report EPS 3-EC-78-12.

Bobbio, a. et al., 2001. Improving the analysis of dependable systems by mapping Fault Trees into Bayesian Networks. *Reliability Engineering and System Safety*, 71, pp.249–260.

Bullemer, P. & Laberge, J.C., 2010. Common operations failure modes in the process industries. *Journal of Loss Prevention in the Process Industries*, 23(6), pp.928–935.

Crowl, D.. A. & Louvar, J.F., 2001. *Title Chemical process safety:fundamentals with appplication* Third edit., Pearson Education.

CSB, 2015. *Final Investigation Report: Cheveron Richmond Refinery Pipe rupture and Fire*, Washington, D.C.

Diez, F.. & Druzdzel, M.., 2007. *Canonical probabilistic models for Knowledge Engineering. Technical Report CISIAD-06-01.*,

Durga Rao, K. et al., 2009. Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. *Reliability Engineering and System Safety*, 94, pp.872–883.

Heckerman, D. & Breese, J.S., 1996. Causal independence for probability assessment and inference using Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, 26(6), pp.826–831.

Khakzad, N., Khan, F. & Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety*, 96(8), pp.925–932.

Khan, F.I. & Abbasi, S. a., 1999. Major accidents in process industries and an analysis of causes and consequences. *Journal of Loss Prevention in the Process Industries*, 12, pp.361–378.

Kidam, K. et al., 2014. Accident prevention approach throughout process design life cycle. *Process Safety and Environmental Protection*, 92(5), pp.412–422..

Kidam, K. & Hurme, M., 2013. Analysis of equipment failures as contributors to chemical process accidents. *Process Safety and Environmental Protection*, 91(1-2), pp.61–78..

Kidam, K. & Hurme, M., 2012. Design as a contributor to chemical process accidents. *Journal of Loss Prevention in the Process Industries*, 25(4), pp.655–666.

Kraaijeveld, P. & Druzdzel, M.J., 2005. GeNIeRate: An interactive generator of diagnostic Bayesian network models. *Proc. 16th Int. Workshop on Principles of Diagnosis*, pp.175–180. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.5119&rep=rep1&type=pdf.

Kujath, M.F., Amyotte, P.R. & Khan, F.I., 2010. A conceptual offshore oil and gas process accident model. *Journal of Loss Prevention in the Process Industries*, 23(2), pp.323–330.

Lampis, M. & Andrews, J.., 2009. Title Bayyesian Belief Networks for System fault Diagnostics. *Quality and Reliability Engineering International*, 25, pp.409–426.

Leveson, N., 2004. A new accident model for engineering safer systems. *Safety Science*, 42, pp.237–270.

Mohammadfam, I. et al., 2013. Developing a process equipment failure knowledge base using ontology approach for process equipment related incident investigations. *Journal of Loss Prevention in the Process Industries*, 26(6), pp.1300–1307.

Nývlt, O. & Rausand, M., 2012. Dependencies in event trees analyzed by Petri nets. *Reliability Engineering and System Safety*, 104, pp.45–57.

Okoh, P. & Haugen, S., 2014. Application of inherent safety to maintenance-related major accident prevention on offshore installations. *Chemical Engineering Transactions*, 36, pp.175–180.

Oniśko, A., Druzdzel, M.J. & Wasyluk, H., 2001. Learning Bayesian network parameters from small data sets: Application of Noisy-OR gates. *International Journal of Approximate Reasoning*, 27, pp.165–182.

Qureshi, Z.H., 2008. A Review of Accident Modelling Approaches for Complex Critical Sociotechnical Systems. , 86, p.72.

Rathnayaka, S., Khan, F. & Amayotte, P., 2013. Accident modeling and risk assessment framework for safety critical decision-making: application to deepwater drilling operation. *Journal of Risk and Reliability*, 227(1), pp.86–105.

Rathnayaka, S., Khan, F. & Amyotte, P., 2012. Accident modeling approach for safety assessment in an LNG processing facility. *Journal of Loss Prevention in the Process Industries*, 25(2), pp.414–423.

Rathnayaka, S., Khan, F. & Amyotte, P., 2010. SHIPP methodology : Predictive accident modeling approach . Part II . Validation with case study. *Process Safety and Environmental Protection*, 89(2), pp.75–88.

Rathnayaka, S., Khan, F. & Amyotte, P., 2011. SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection*, 89(3), pp.151–164.

Sheppard, J.P. & Chowdhury, S.D., 2005. Riding the wrong wave: Organizational failure as a failed turnaround. *Long Range Planning*, 38, pp.239–260.

Tan, Q. et al., 2013. Dynamic accident modeling for high-sulfur natural gas gathering station. *Process Safety and Environmental Protection*, 92(6), pp.565–576.

Taylor, J.R., 2007. Statistics of design error in the process industries. *Safety Science*, 45, pp.61–73.

Wasyluk, H., 2001. Learning Bayesian network parameters from small data sets : application of. , 27, pp.165–182.

Zagorecki, A. & Druzdzel, M., 2004. An Empirical Study of Probability Elicitation Under Noisy-OR Assumption. *FLAIRS Conference*, pp.880–885.

# Chapter 5

## 5.0 Dynamic safety analysis of process system using nonlinear and non – sequential accident model

### Preface

*A version of this chapter has been published in the **Journal of Chemical Engineering Research and Design 2016; 111:169-183**. I am the primary author. Co-author Faisal Khan provided fundamental understanding, assisted in developing the conceptual model and subsequently translated this to the numerical model. Co-author Ming Yang provided much needed support in implementing the concept and testing the model. I carried out most of the data collection and analysis. I prepared the first draft of the manuscript and subsequently revised the manuscript, based on the feedback from co-authors and also a peer review process. The two co-authors assisted in developing the concept and testing the model, reviewed and corrected the model and results. They also contributed to the review and revision of the manuscript.*

### Abstract

Analysis of the safety and reliability of complex engineering systems is becoming challenging and highly demanding. In complex engineering systems, accident causation is a function of nonlinear interactions of several accident contributory factors. Traditional accident models normally use a fault and event trees sequential approach to predict cause-consequence relationships, which unable to capture real interaction thus have limited predictability of accident.

This paper presents a new non-sequential barrier-based process accident model. The conditional dependencies among accident contributory factors within prevention barriers are modelled using the Bayesian network with various relaxation strategies, and non-sequential failure of prevention (safety) barriers. The modelling of non-linear interactions in the model led to significant

improvement of the predicted probability of an accident when compared with that of sequential technique. This renders valuable information for process safety management. The proposed accident model is tested on a real life case study from the U.S Chemical Safety Board.

**Keywords:** Sequential accident model, Non-sequential accident model, Accident prediction, Bayesian network analysis, Leaky Noisy-OR Gate

## 5.1. Introduction

Due to the complexity of modern industrial technological systems, the risk posed by accidents in such systems is becoming more worrisome. Operating modern process plants demands high levels of safety and reliability. Process accidents occur due to chains of abnormal events instigated by human error, technical failure, external causes and deviation from process parameters. The majority of devastating accidents in process plants such as toxic release, fire and explosion are initiated by process hazards. A vital part of the safety system is to identify hazards associated with a process system and estimate the probability of occurrence and the subsequent consequences involved. Comprehensive analysis of how an accident process evolves from initiation to the termination stage is of paramount importance in designing safety into process systems to avoid accidents (Rathnayaka et al. 2010; Tan et al. 2013).

Accident models are theoretical frameworks which typically show the relationship between causes and consequences and vividly explain why and how accidents occurred. Accident models are used as techniques for risk assessment during the system development stage and for subsequent use as post hoc accident investigation to analyse the root causes of an accident (Qureshi 2008). Quite a number of accident models and numerous methods of accident modelling (FMEA, ETA, FTA and MORT) have been developed in the last few decades (Katsakiori et al. 2009). Detailed descriptions of these accident models and their applications can be can be found in literature (Attwood et al.

2006; Rathnayaka et al. 2011; Rasmussen 1997; Heinrich 1941; Kujath et al. 2010; Qureshi 2008). A detailed review of accident models that have been developed for the chemical process industry with a detailed analysis of the class of accident model known as the dynamic sequential accident model can be found in Al-shanini et al. (2014). Existing accident models have their own strengths and weakness and they depend mainly on the areas of their application, purpose and focus. The majority of the existing accident models are sequential accident models where an accident processes from initiation to termination and are considered as a chain of independent events that occurred in a definite particular order. The severity of effects is presumed to progress through the sequential failure of independent events. These traditional models use a fault and event trees sequential approach to predict the cause-consequence relationship, which provides a sequential explanatory mechanism of accident propagation. However, in a real life situation, this need not be true.

Also, existing models are not capable of modelling multiple risk factors in process systems where interactions among sub systems are nonlinear and extremely complex, and they are not capable of using accident precursor data to evaluate risk and develop accident prevention strategies (Tan et al. 2013; Rathnayaka et al. 2011). Recently, Baksh et al. (2015) allowed random failure of safety barriers in their predictive accident modelling but nonlinear interactions of accident contributory factors within the safety barriers were ignored in their model.

Due to the complexity of process's operations and the high level of interaction among sub-systems, accident causation is a function of nonlinear interaction of various factors. A thorough review of existing accident models reveals that the majority of the models belong to the class of sequential accident models, where the accident process is described as a chain of independent events that take place sequentially.

The modelling flexibility of the BN structure can accommodate various kinds of conditional dependencies that cannot be readily included in FTA and ETA. Application of Bayesian network to accident model analysis is of great advantage and hence non-linear interaction among accident contributory factors can be easily modelled and predicted with various relaxation strategies.

The major objective of this study is to model the nonlinear interaction of accident contributory factors within the safety barriers under relaxation strategies and subsequently allow the non-sequential failure of safety barriers to cause adverse events randomly. The accident processes from initiation to termination are viewed as non-sequential processes and modeled using BN in this study. The non-sequential accident model is a relatively new concept.

The remaining parts of this paper are organised as follows. Section 5.2 presents a brief description of failure analysis techniques. Section 5.3 presents the comparison of sequential and non-sequential techniques with a case study. Section 5.4 presents the results and discussion. Finally, Section 5.5 provides the conclusion.

## 5.2. Failure Analysis Techniques

Quite a number of methodologies have been developed for accident analysis; the most widely used technique is fault tree analysis. Recently BNs have gained much attention because they can accommodate different kinds of statistical dependencies that cannot be easily included in other accident analysis techniques.

### 5.2.1. Fault tree

A fault tree is a deductive, graphic methodology used to determine failure probability of a complex system. The top event in the fault tree represents a major accident initiating hazard. The top event is placed at the top of the fault tree and the fault tree is graphically modelled downward to allow the visualization of all possible combinations of malfunctions and wrong actions that could initiate the top event. Fault trees are usually constructed from events and logic gates (Khakzad et al. 2011). The underlying technical failures that lead to accidents are usually represented by basic events. The logic gates in the fault tree represent numerous ways by which machines and human error interact to cause the accident. AND and OR gates are the commonly used logic gates in the fault tree. Analysis in the fault tree can proceed both qualitatively and quantitatively (Nivolianitou et al. 2004). In AND gate, process components interact in parallel structure and process failure requires the simultaneous failure of all components in parallel. The failure probability of the top event in parallel structure (AND gate) is calculated by equation 5.1. Also in OR gate, process components interact in series structure and failure of any single components in series leads to failure of the process. The failure probability of the top event in series structure (OR gate) is calculated by equation 5.2.

$$P = \prod_{i=1}^{n} P_i \tag{5.1}$$

$$P = 1 - \prod_{i=1}^{n} (1 - P_i) \tag{5.2}$$

### 5.2.2. Event tree

Event tree analysis is a systematic method for studying an accident scenario in complex systems. It is an inductive/forward looking analysis that is used to analyse event sequence, after an initiating event. It normally starts with a specific initiating event and usually ends with the possible consequences referred to as end states. The safety barriers or functions in an event tree are usually arranged in chronological order, meaning the events are considered in the same sequence they are expected to happen during an accident sequence. It is a proactive risk analysis methodology used to identify and illustrate a potential event sequence to obtain both qualitative and quantitative representation and assessment (Sklet 2004; Nivolianitou et al. 2004).

The occurrence probabilities of end state consequences $P(C_k)$ in the event tree is calculated by equation 5.3.

$$P(C_K) = \prod_{j \in SB_k} x_i^{\theta_{i,k}} (1 - x_i)^{1-\theta_{i,k}} \qquad (5.3)$$

Where $SB_k$ represents prevention barrier related to level k; and $\theta_{i,k} = 1$ when level k failure passes through the failure branch of prevention barrier i; $\theta_{i,k} = 0$ when level k failure passes through the success branch of prevention barrier i. $x_i$ is the failure probabilities of prevention barriers.

### 5.2.3. Bayesian network

The Bayesian network is a graphical technique; it provides a robust probabilistic technique of reasoning under uncertainty. BN techniques have been extensively used in risk and safety analysis based on probabilistic and uncertain knowledge. BN (also known as a probabilistic dependence graph) is a direct acyclic graph with numerous nodes representing variables and arcs signifying direct causal relationships among the linked nodes. A conditional probability table (CPT) is assigned to the various nodes to denote conditional dependencies among the linked nodes (Bobbio et al. 2001; Khakzad et al. 2013). Based on both conditional independence and the chain rule, the BN represents the joint probability distribution P(U) of a set of discrete random variables U= $\{A_1 \ldots, A_n\}$, incorporated in the network as:

where Pa ($A_i$) is the parent of variable $A_i$ and P (U) is the joint probability distribution of variables (Pearl 1998; Jensen & Nielsen 2007).

$$P(U) = \prod_{i=1}^{n} P(A_i \mid P_{a(A_i}))  \tag{5.4}$$

The BN makes use of Bayes theorem to update prior occurrence probability of events to give consequence probability (posterior) provided new information called evidence is given. The following equation is used to estimate posterior probability.

$$P(U|E) = \frac{P(U,E)}{P(E)} = \frac{P(U,E)}{\sum_U P(U,E)}  \tag{5.5}$$

Canonical probabilistic models are of paramount importance because they make the construction of a probabilistic model very easy and drastically reduce the computation time required. Canonical models are increasingly used in probabilistic systems and different canonical models could coexist in any probabilistic network. One way of minimizing the complexity of elicitation of numerical probabilities is to depend on canonical probabilistic models which provide the opportunity of building probability distribution from a small number of parameters (Diez & Druzdzel 2007). The Common type of canonical interactions famously used in the Bayesian network are known as Noisy-OR and Leaky Noisy-OR gates. The application of canonical interactions in BN provides an effective technique for modeling various kinds of statistical dependencies and nonlinear interactions.

#### *5.2.3.1. Noisy-OR Gate*

The Noisy-OR gate canonical model assumes that causes and effects are binary with two unique states: true and false. A Noisy-OR model is normally used to describe various interactions between n causes $X_1$, $X_2$… $X_n$ and their common effect Y. The model assumes that causes $X_i$ influence Y independently from each other and that causes $X_i$ have a probability $P_i$ strong enough to produce effect Y if all other causes are false. These assumptions offer the opportunity to completely specify the conditional probability distribution with n parameters $p_1,..., p_n$. $P_i$ denotes the probability that effect Y will be true if the cause $X_i$ is true and all other causes $X_j$, $j \neq i$ is false (Ońiśko et al. 2001). This can be mathematically expressed as:

$$P_i = \Pr(y|\bar{x}_1, \bar{x}_2, \dots, x_i, \dots, \bar{x}_{n-1}, \bar{x}_n) \tag{5.6}$$

Therefore, the probability of y given a subset of Xp of the $X_i$s that are true is given by the following formula:

$$\Pr(Y|X_P) = 1 - \prod_{i:X_i \in X_p} (1 - P_i) \tag{5.7}$$

The Noisy-OR model adopts an independent mechanism, and application of the model results in a substantial reduction in the number of probabilities needed to quantify the cause-effect interaction. The model requires n probabilities while the unrestricted model needs $2^n$ probabilities (Heckerman & Breese 1996).

#### *5.2.3.2. Leaky Noisy-OR*

Leaky Noisy-OR is an extension of the binary Noisy-OR gate for the situation where a subsystem could fail though all of its components are functional. The Leaky Noisy-OR gate is normally applicable to a situation where a model does not capture all potential causes of effect Y (Bobbio et al. 2001). Invariably virtually all situations encountered in practice fit this class. In this model, the combined effect of all unmolded causes of effect Y is called the leak probability *l*. The leak probability (*l*) is the probability that effect Y will happen spontaneously (True) though all its causes are absent (False) (Zagorecki & Druzdzel 2004; Oniśko et al. 2001). The leaky Noisy-OR formula that can be used to estimate the probability of effect Y given the subset $X_p$ of $X_i$ which are true is given by:

$$\Pr(Y|X_P) = 1 - \left[\!\left[(1-l)\prod_{i:X_i \in X_p}(1-P_i)\right]\!\right] \tag{5.8}$$

## 5. 3. Safety analysis

### 5.3.1. Case study: The Tesoro Anacortes Refinery accident.

The Tesoro Anacortes Refinery in United States of America experienced an appalling rupture of a heat exchanger in the catalytic Reformer/ Naphtha Hydro treater unit on April 2, 2010. The rupture occurred on the E-6600E heat exchanger as a result of a high temperature hydrogen attack (HTHA). The ruptured heat exchanger released highly flammable hydrogen and naphtha at more than 500$^{o}$F. The flammable hydrogen and naphtha were ignited, causing an explosion and severe fire that lasted for more than three hours. Till now, this is the largest devastating incident at a US petroleum refinery after the BP Texas city accident in March 2005 (U.S. Chemical Safety and Hazard Investigation Board 2014).

The U.S Chemical Safety Board (CSB) investigated the incident extensively. CSB highlighted various safety and technical laxities in their report. Based on the information made available in the CSB report, this accident has been modelled using the proposed methodology. The accident contributory factors highlighted are systematically arranged into seven prevention barriers along the accident pathway to prevent the effects of the accident. A brief description of various prevention barriers in this model is given below.

Release prevention barrier (RPB): The release of material is mainly responsible for the loss of containment that initiates the accident process. It has been identified that operational error, inspection error, maintenance error and design error are the major factors that influence the failure of the release prevention barrier.

Dispersion prevention barrier (DPB): The primary function of this barrier is to minimize the extent of a hazardous event to prevent the further spreading of material and energy. The major factors that are responsible for the failure of this barrier are safety system failure, operational error and communication error.

Ignition prevention barrier (IPB): Ignition prevention is of paramount importance in process facilities that handle and process diverse flammable materials. To prevent an outburst of fire and explosion in process facilities, a safety barrier must be installed to focus on all potential sources of ignition in the process facilities to prevent the outburst of fire and explosion. The major factors that influence the failure of this barrier are: hot work failure, heat exchanger failure and operator's error.

Escalation prevention barrier (EPB): The primary function of this barrier is to isolate the surroundings to avoid domino accident scenarios once ignition has occurred in process facilities. It minimizes the extent and duration of ignition. The major factors that are responsible for the failure of this barrier are: fire detection system failure, operator error and emergency shutdown failure.

Emergency management failure prevention barrier (EMFPB): This barrier is installed to control the extent of hazardous events as much as possible, or to drastically reduce their consequences. The main reason why this barrier is installed is to prevent fatalities. Factors that mainly influence the failure of this barrier are: evacuation error, communication error and emergency preparedness failure.

Human factor prevention barrier (HFPB): Recently, human error has contributed significantly to the major accidents that have occurred in various process industries. Good examples of accidents that occurred due to human error are the Bhopal gas accident and the Texas City refinery accident (Okoh & Haugen 2014). Factors that are mostly responsible for human error are: knowledge error, skill error and management error.

Organization failure prevention barrier (OrFPB): Organisation failure contributes significantly to accidents in process industries. In most cases, the underlying causes for accidents are management and organizational factors (Rathnayaka et al. 2013).

### 5.3.2. Sequential accident analysis (sequential cause-consequence relationship)

#### *5.3.2.1. Fault tree analysis*

Fault tree analysis (FTA) which is deterministic is used to quantify the failure probabilities of all prevention barriers. The prevention barriers in this accident model were systematically analysed with FTA to establish a sequential causal relationship. The top event in the fault tree represents the failure of the prevention barrier. The second layer of the fault tree represents all accident contributory factors for each prevention barrier; their failure will induce the failure of the top event. The third layer represents the causes of accident contributory factors. A combination of AND and

OR logics was used to evaluate the failure probability of the top event. Table 5.1 presents the event description and failure probabilities of the basic events for RPB for the case study in the model.

**Table 5.1. Basic event failure probability for Release prevention barrier.**

| Event | Event Description | Assigned probability |
|---|---|---|
| 1 | High temperature hydrogen attack | $2.50 \times 10^{-2}$ |
| 2 | Difficulty with valve operation during start up | $1.50 \times 10^{-2}$ |
| 3 | Leaks from heat exchanger during start up not reported | $5.00 \times 10^{-2}$ |
| 4 | Hydrogen induced cold cracking | $1.00 \times 10^{-3}$ |
| 5 | Inexperience | $1.00 \times 10^{-2}$ |
| 6 | Job carried out without permit to work | $1.00 \times 10^{-2}$ |
| 7 | External supervision failure | $8.30 \times 10^{-2}$ |
| 8 | Wrong procedure | $5.00 \times 10^{-3}$ |
| 9 | Poor construction material for NHT heat exchanger | $1.00 \times 10^{-2}$ |
| 10 | High mechanical stress | $1.00 \times 10^{-2}$ |
| 11 | Insufficient instrumentation to measure process conditions | $1.00 \times 10^{-3}$ |
| 12 | Long delay in inspection schedule | $5.00 \times 10^{-2}$ |
| 13 | Inadequate methods for detecting HTHA | $9.00 \times 10^{-2}$ |
| 14 | Inadequate training of the inspectors to detect HTHA easily | $2.50 \times 10^{-2}$ |
| 15 | Failure of HTHA inspection on heat exchanger | $5.50 \times 10^{-2}$ |
| 16 | Failure to detect leaks from heat exchanger flanges | $5.00 \times 10^{-2}$ |
| 17 | Failed to detect minor release | $5.00 \times 10^{-2}$ |
| 18 | Wrong maintenance procedure (Nelson curve Methodology | $5.00 \times 10^{-3}$ |
| 19 | Delay maintenance operations | $5.00 \times 10^{-2}$ |
| 20 | HTHA degradation monitoring performed but failed to detect | $6.60 \times 10^{-2}$ |
| 21 | HTHA degradation monitoring specified but not performed | $5.00 \times 10^{-2}$ |

The reliability data used in fault tree analysis were sourced from several journals (Rathnayaka et al. 2010; Rathnayaka et al. 2012; Rathnayaka et al. 2013; Tan et al. 2013; A. Al-shanini et al. 2014) and used expert judgement where the data was not readily available . The fault tree analysis for RPB in the model is shown in Figure 5. 1.

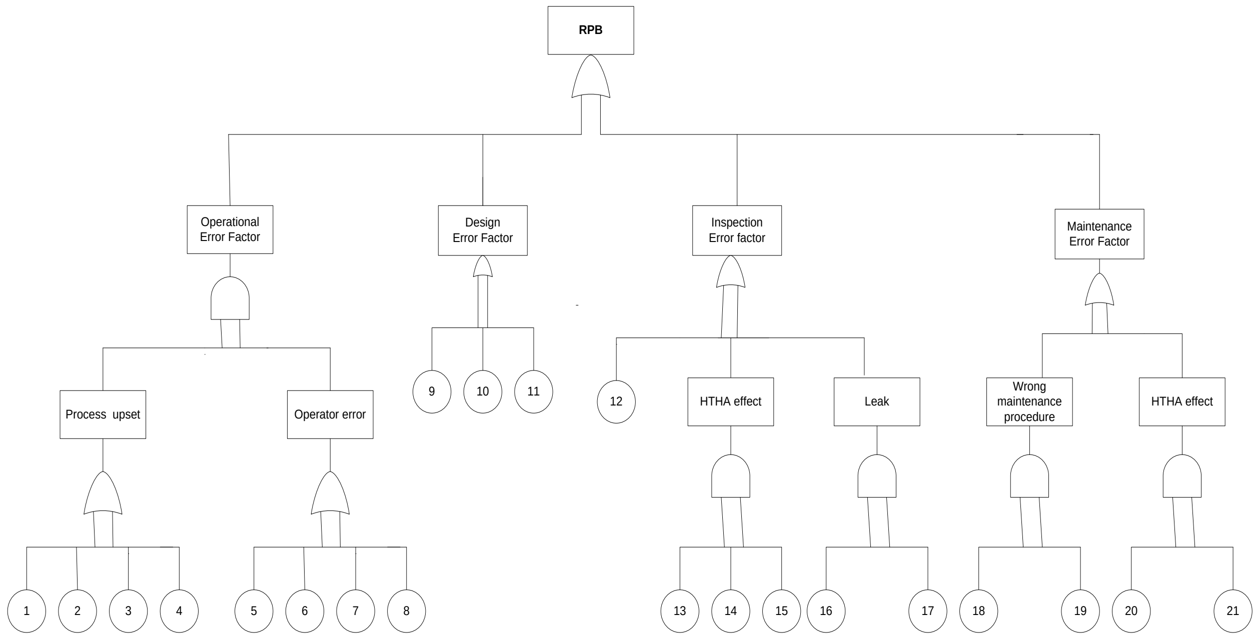



**Figure 5.1. Fault tree analysis of Release prevention barrier.**

The top event in the fault tree of RPB represents the failure of this prevention barrier. The second layer of the RPB fault tree represents all accident contributory factors; their failure will induce the failure of the RPB. The third layer denoted by circles represents the basic events (causes) of accident contributory factors.

It is common to use AND and OR logics in fault tree analysis; therefore, its computational procedure is not discussed here. In order to minimize the number of figures and tables in this paper, the reliability data of basic events and accompanying fault trees for other prevention barriers are not shown. The failure probabilities of all prevention barriers through FT analysis is given in Table 5.6.

#### *5.3.2.2. Sequential estimation of consequences occurrence probabilities*

The occurrence probability of consequences for sequential cause-consequence relationships is estimated with the combination of fault tree and event tree analysis. The occurrence probability of consequences is estimated by propagating failure probabilities (obtained from FTA) through successive success or failure branches of an event tree. The event sequence leads to all potential consequences in the event tree. Six major consequences have been identified: safe, near miss, mishap, incident, accident and catastrophe. Figure 5.2 shows the event tree in the model.

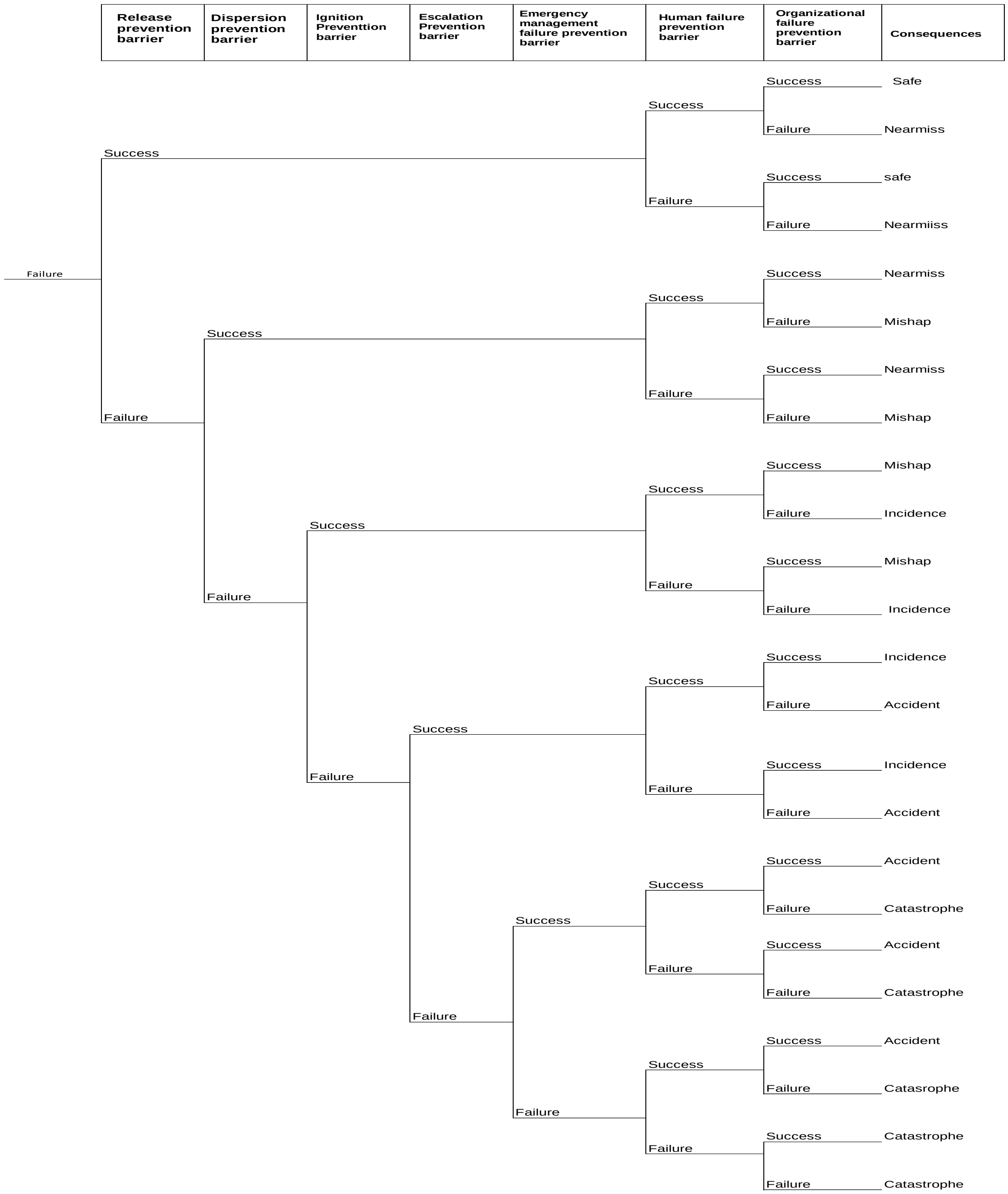



**Figure 5. 2. Event tree for the accident model.**

### 5. 3.3. Non sequential accident analysis (Non sequential cause-sequential consequence relationship*)*

#### *5.3.3.1. Quantifying failure probabilities of all accident contributory factors within safety barriers using BN*

The failure probabilities of all accident contributory factors within each prevention barrier are quantified with BN using different logic gates (Noisy-OR and Leaky Noisy-OR gates). The failure probabilities estimated here will subsequently be used to estimate the failure probabilities of all prevention barriers when the non-linear interaction of accident contributory factors is considered, using BN in Figures 5.4-5.10. The following procedures are followed in quantifying failure probabilities of all accident contributory factors using BN, when Noisy-OR and Leaky Noisy-OR gates are applied to the BN respectively.

(1) Calculating the safe (non-occurrence) probability of all parent nodes in the BN

(2) Assigning the non-causation probability of all parent nodes in the conditional probability table (CTP) based on expert judgement or data. When the Leaky Noisy-OR gate is considered in the BN, leak probability will be assigned prior to this step.

(3) Computing the CTP

(4) Applying the CTP with probabilities of the state of the parent nodes (safe/false or failure/true) conditional to the state involved, compute the probability of a top event.

The BN in Figure 5.3 shows causes of mechanical failure in a process system.

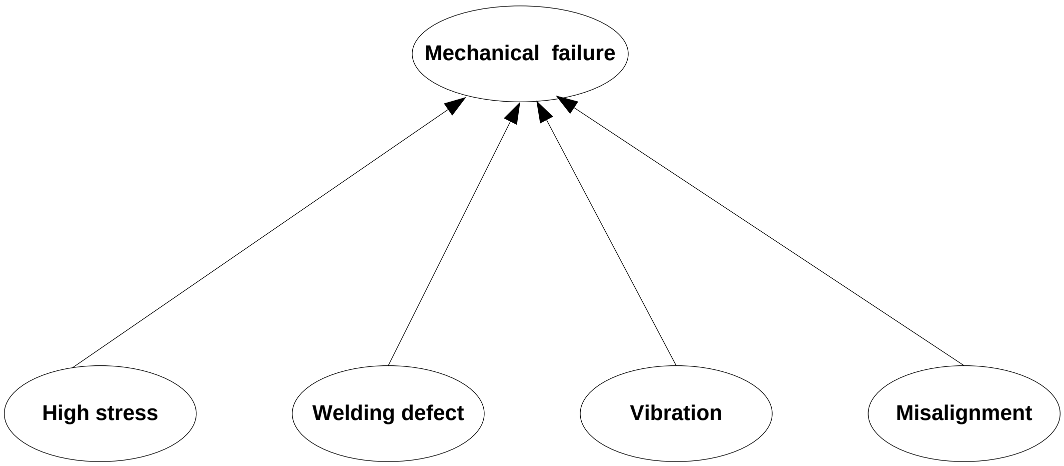



**Figure 5.3. Bayesian network for causes of mechanical failure.**

The probability of mechanical failure, using the Noisy-OR gate in the BN is calculated as follows. Using Step (1), based on the failure probabilities given in Table 5. 2, the safe probabilities of causes of mechanical failure are computed as shown in Table 5.2.

**Table 5.2. Failure and safe probabilities of causes of mechanical failure in process system.**

| No | Causes of mechanical failure | Failure probability | Safe probability |
|---|---|---|---|
| 1 | High stress | 0.001 | 0.999 |
| 2 | Welding defect | 0.025 | 0.975 |
| 3 | Vibration | 0.003 | 0.997 |
| 4 | Misalignment | 0.1667 | 0.8333 |

Based on step (2) to step (4), the probability of mechanical failure is computed as shown in Table 5.3.

**Table 5.3. Probability of mechanical failure for Noisy-OR gate.**

| State | High stress | Welding defect | Misalignment | Causation probability of | Non causation probability of mehanical failure | Conditonal Probability of mechanical failure for different states |
|---|---|---|---|---|---|---|
| 1 | F | F | F | 0 | 1 | 0 * 0.999 * 0.975 * 0.997 * 0.8333 = 0 |
| 2 | F | F | **T** | 0.75 | **0.25** | 0.75 * 0.999 * 0.975 * 0.997 * 0.1667 = 0.121412 |
| 3 | F | F | F | 0.65 | **0.35** | 0.65 * 0.999 * 0.975 * 0.003 * 0.8333 = 1.582E-03 |
| 4 | F | F | T | 0.9125 | 8.75 E-02 = 0.35*0.25 | 0.9125 * 0.999 * 0.975 * 0.003 * 0.1667 = 4.58175E-04 |
| 5 | F | **T** | F | 0.7 | **0.3** | 0.7 * 0.999 * 0.025 * 0.997 * 0.8333 = 1.45244E-02 |
| 6 | F | T | T | 0.925 | 7.5E-02 = 0.3*0.25 | 0.925 * 0.999 * 0.025 * 0.997 * 0.1667 = 3.839529E-03 |
| 7 | F | T | F | 0.895 | 1.05E-01 = 0.3*0.35 | 0.895 * 0.999 * 0.025 * 0.003 * 0.8333 = 5.58592E-05 |
| 8 | F | T | T | 0.97375 | 2.625E-02 = 0.3*0.35*0.25 | 0.97375 * 0.999 * 0.025 * 0.003 * 0.1667 = 1.21621E-05 |
| 9 | **T** | F | F | 0.6 | **0.4** | 0.6 * 0.001 * 0.975 * 0.997 * 0.8333 = 4.86E- 04 |
| 10 | T | F | T | 0.9 | 0.1 = 0.4*0.25 | 0.9 * 0.001 * 0.975 * 0.997 * 0.1667 = 1.4584E-04 |
| 11 | T | F | F | 0.86 | 0.14= 0.4*0.35 | 0.86 * 0.001 * 0.975 * 0.003 * 0.8333 = 2.096E-06 |
| 12 | T | F | T | 0.965 | 3.5E-2 = 0.4*0.35*0.25 | 0.965 * 0.001 * 0.975 * 0.003 * 0.1667 = 4.7053E-07 |
| 13 | T | T | F | 0.88 | 0.12 = 0.4*0.3 | 0.88 * 0.001 * 0.025 * 0.997 * 0.8333 = 1.8277E-05 |
| 14 | T | T | T | 0.97 | 3.0E-2 = 0.4*0.3*0.25 | 0.97 * 0.001 * 0.025 * 0.997 * 0.1667 = 4.0303E-06 |
| 15 | T | T | F | 0.958 | 4.2E-2= 0.4*0.3*0.35 | 0.958 * 0.001 * 0.025 * 0.003 * 0.8333 = 5.98726E-08 |
| 16 | T | T | T | 0.9895 | 1.05E-2 = 0.4*0.3*0.35*0.25 | 0.9895 * 0.001 * 0.025 * 0.003 * 0.1667 = 1.23712E-08 |
| | | | | | | The probabiity of mechanical failure is sum of all states = 0.1425 |

Still following the previous example, based on step (1) to step (4), the probability of mechanical failure using the Leaky Noisy-OR gate was computed as shown in Table 5.4.

**Table 5. 4. Probability of mechanical failure for Leaky Noisy-OR gate.**

| State | High stress | Welding defect | Vibration | Misalignment | Causation probability of mechanical failure | Non causation probabilty of mehanical failure | Conditonal Probability of mechanical failure for different states |
|---|---|---|---|---|---|---|---|
| 1 | F | F | F | F | 0.01 | 0.99 | 0 .01 * 0.999 * 0.975 * 0.997 * 0.8333 = 8.0922E-03 |
| 2 | F | F | F | **T** | 0.7525 | 0.2475 = **0.25** *0.99 | 0.7525 * 0.999 * 0.975 * 0.997 * 0.1667 = 1.2182E-01 |
| 3 | F | F | **T** | F | 0.6535 | 0.3465 = **0.35** * 0.99 | 0.6535 * 0.999 * 0.975 * 0.003 * 0.8333 = 1.5912E-03 |
| 4 | F | F | T | T | 0.913375 | 8.6625 E-02 = 0.35*0.25*0.99 | 0.913375 * 0.999 * 0.975 * 0.003 * 0.1667 = 4.4491E-04 |
| 5 | F | **T** | F | F | 0.703 | 0.297 = **0.3** *0.99 | 0.703 * 0.999 * 0.025 * 0.997 * 0.8333 = 1.4587E-02 |
| 6 | F | T | F | T | 0.92575 | 7.425E-02 = 0.3*0.25*0.99 | 0.92575 * 0.999 * 0.025 * 0.997 * 0.1667 = 3.8426E-03 |
| 7 | F | T | **T** | F | 0.89605 | 1.0395E-01 = 0.3*0.35*0.99 | 0.89605 * 0.999 * 0.025 * 0.003 * 0.8333 = 5.5925E-05 |
| 8 | F | T | T | T | 0.9740125 | 2.59875E-02 = 0.3*0.35*0.25*0.99 | 0.9740125 * 0.999 * 0.025 * 0.003 * 0.1667 = 1.2165E-05 |
| 9 | **T** | F | F | F | 0.604 | 0.396 = **0.4** * 0.99 | 0.604 * 0.001 * 0.975 * 0.997 * 0.8333 = 4.8926E- 04 |
| 10 | T | F | F | T | 0.901 | 0.099= 0.4*0.25*0.99 | 0.901 * 0.001 * 0.975 * 0.997 * 0.1667 = 1.4600E- 04 |
| 11 | T | F | T | F | 0.8614 | 0.1386 = 0.4*0.35*0.99 | 0.8614 * 0.001 * 0.975 * 0.003 * 0.8333 = 2.0996E-06 |
| 12 | T | F | T | T | 0.9635 | 0.0 3465 0.4*0.35*0.25*0.99 | 0.9635 * 0.001 * 0.975 * 0.003 * 0.1667 = 4.6980E-07 |
| 13 | T | T | F | F | 0.8812 | 0.1188 = 0.4*0.3*0.99 | 0.8812 * 0.001 * 0.025 * 0.997 * 0.8333 = 1.8303E-05 |
| 14 | T | T | F | T | 0.9703 | 0.0297 = 0.4*0.3*0.25*0.99 | 0.9703 * 0.001 * 0.025 * 0.997 * 0.1667 = 4.0316E-06 |
| 15 | T | T | T | F | 0.95842 | 0.04158 = 0.4*0.3*0.35*0.99 | 0.95842 * 0.001 * 0.025 * 0.003 * 0.8333 = 5.9899 E-08 |
| 16 | T | T | T | T | 0.989605 | 0.010395 = 0.4*0.3*0.35*0.25*0.99 | 0.989605 * 0.001 * 0.025 * 0.003 * 0.1667 = 1.24373 E-08 |
| | | | | | | | The probabiity of mechanical failure is sum of all states = 0.15110 |

Following this procedure, the failure probability of factors contributing to the failure of all prevention barriers in the model is given by Table 5.5.

**Table 5.5. Failure probability of factors contributing to failure of all prevention barrier, using different logic gates.**

| Prevention barrier | Logic Gate | Operational error factor | Inspection error factor | Maintenance error factor | Design error factor |
|---|---|---|---|---|---|
| RPB | BN(Noisy-OR) | 0.061 | 0.084 | 0.047 | 0.006 |
| | BN(Leaky-Noisy-OR) | 0.071 | 0.093 | 0.057 | 0.016 |
| DPB | | **Communication error factor** | **Safety system failure factor** | **Operator error factor** | |
| | BN(Noisy-OR) | 0.083 | 0.133 | 0.116 | |
| | BN(Leaky-Noisy-OR | 0.093 | 0.141 | 0.125 | |
| IPB | | **Heat exchnger failure factor** | **Operator error factor** | **Hot work failure factor** | |
| | BN(Noisy-OR) | 0.036 | 0.116 | 0.081 | |
| | BN(Leaky-Noisy-OR) | 0.046 | 0.125 | 0.090 | |
| EPB | | **Fire detection system failure factor** | **operator error factor** | **Emergency detection system failure factor** | |
| | BN(Noisy-OR) | 0.097 | 0.044 | 0.078 | |
| | BN(Leaky-Noisy-OR) | 0.106 | 0.053 | 0.088 | |
| EMPB | | **Communication error factor** | **Evacuation error factor** | **Emergency response failure factor** | |
| | BN(Noisy-OR) | 0.130 | 0.109 | 0.110 | |
| | BN(Leaky-Noisy-OR) | 0.141 | 0.118 | 0.119 | |
| HFPB | | **Knowledge error factor** | **Skill error factor** | **Management error factor** | |
| | BN(Noisy-OR) | 0.047 | 0.012 | 0.044 | |
| | BN(Leaky-Noisy-OR) | 0.057 | 0.022 | 0.054 | |
| OrFPB | | **Knowledge error factor** | **Communication error factor** | **Management error factor** | |
| | BN(Noisy-OR) | 0.052 | 0.066 | 0.068 | |
| | BN(Leaky-Noisy-OR) | 0.062 | 0.077 | 0.076 | |

In a complex engineering system, accident causation is a function of nonlinear interactions of various factors and accident progression could be viewed as complex interactions of diverse factors. The dependency of accident contributory factors and nonlinear interaction within each of the prevention barriers is represented by the BNs in Figures 5.4- 5.10.

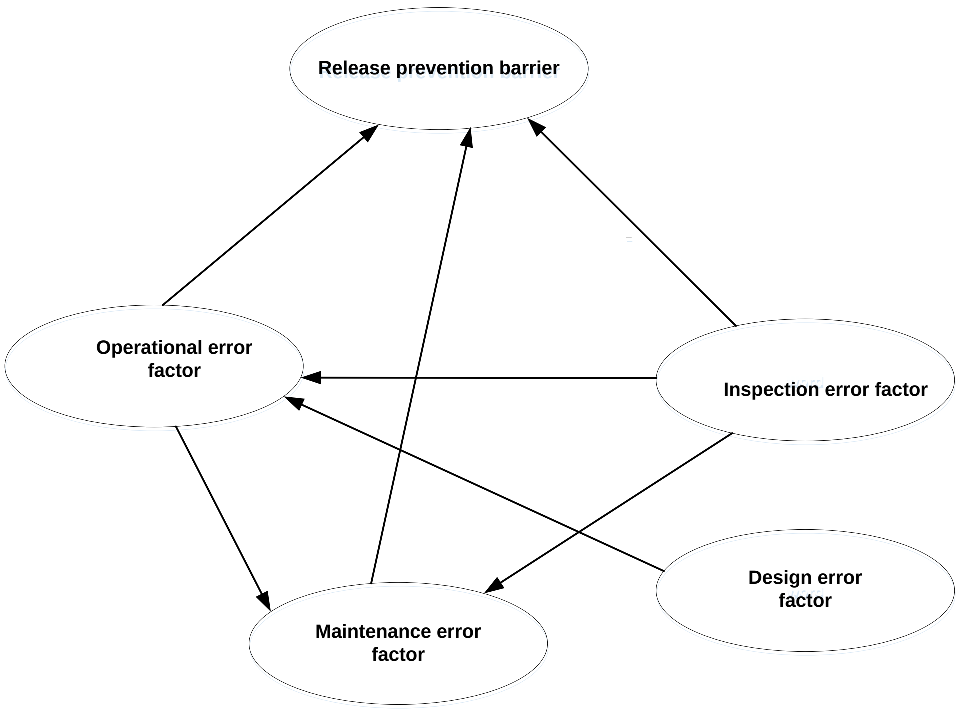



**Figure 5.4. Interdependency of accident contributory factors in Release prevention barrier.**

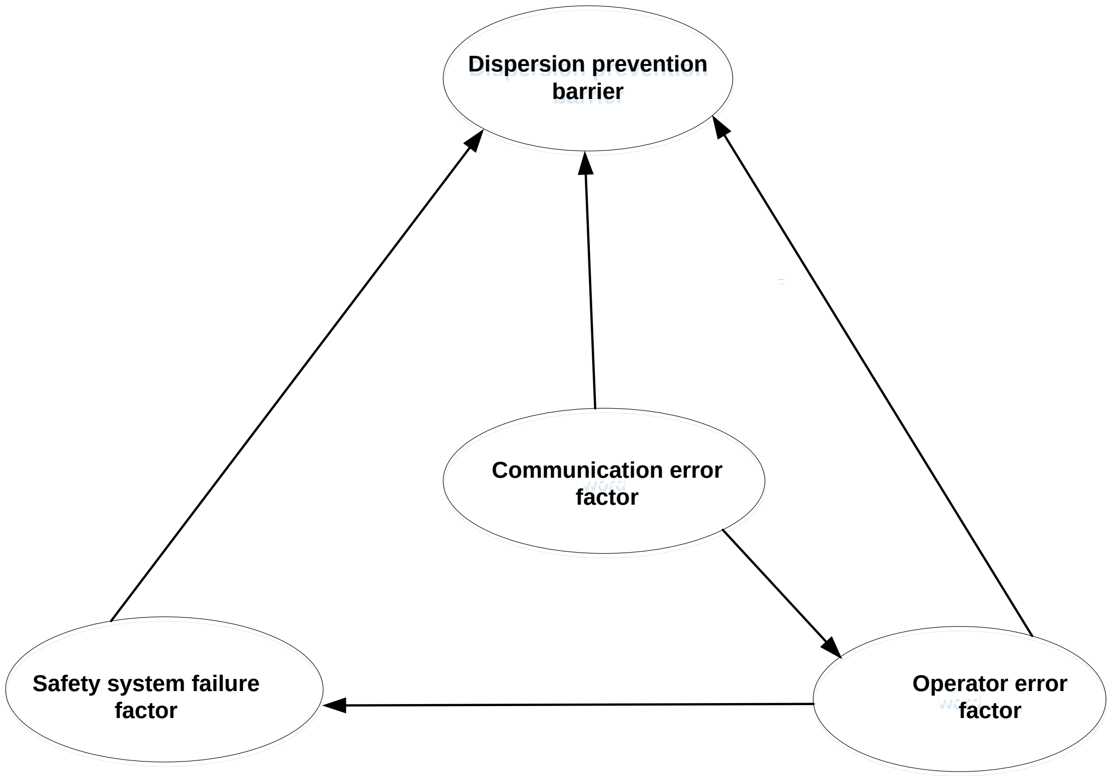



**Figure 5.5. Interdependency of accident contributory factors in Dispersion prevention barrier.**

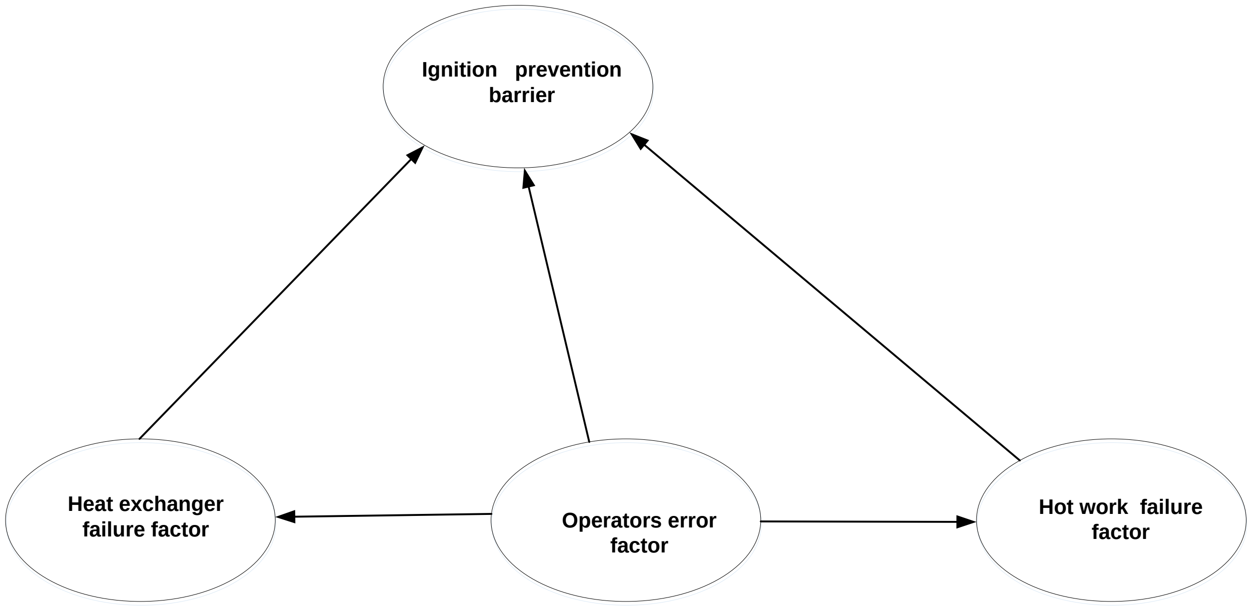



**Figure 5.6 . Interdependency of accident contributory factors in Ignition prevention barrier.**

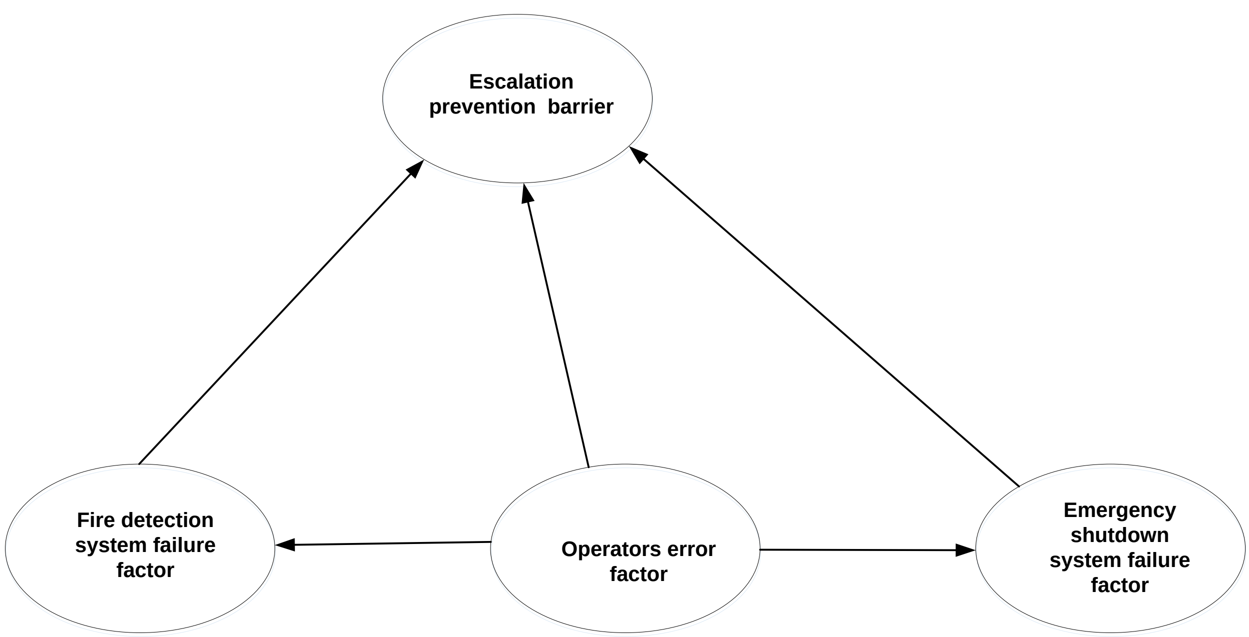



**Figure 5.7. Interdependency of accident contributory factors in Escalation prevention barrier.**

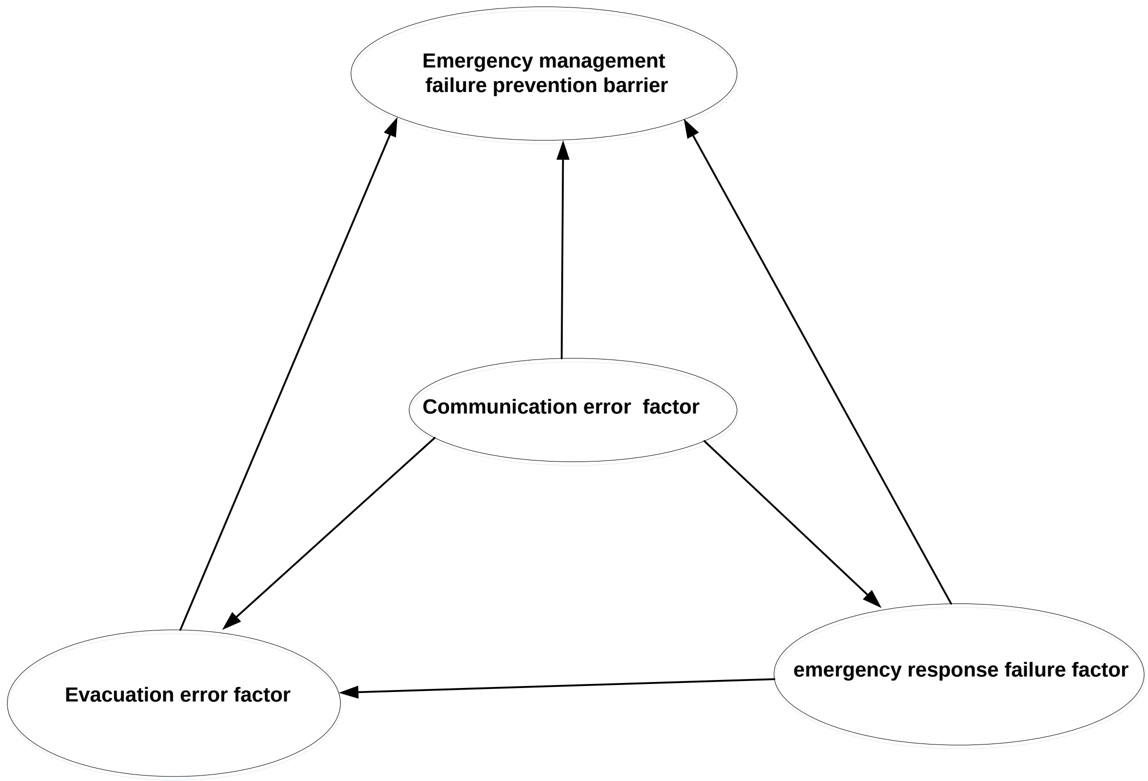



**Figure 5. 8. Interdependency of accident contributory factors in Emergency management failure prevention barrier.**

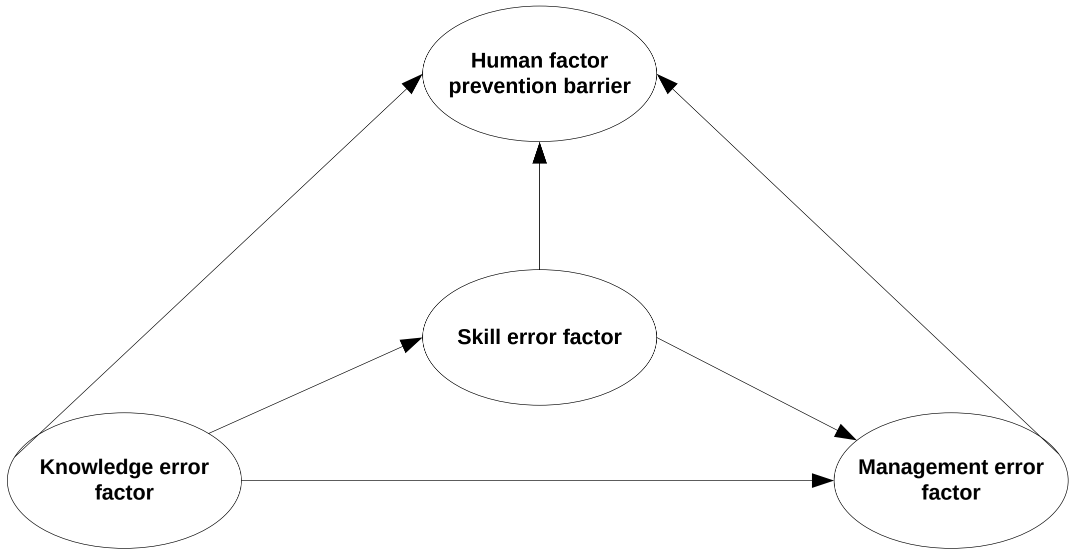



**Figure 5.9. Interdependency of accident contributory factors in Human factor prevention barrier.**

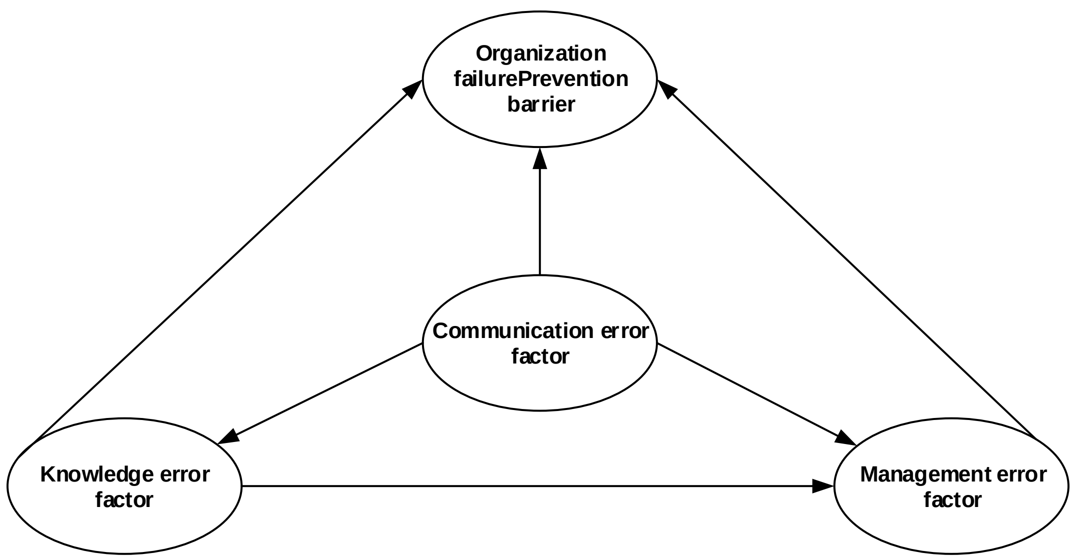



**Figure 5.10. Interdependency of accident contributory factors in Organization failure prevention barrier.**

#### *5.3.3.3. Quantifying failure probabilities of all prevention barriers via BN*

The BN approach was used to quantify failure probabilities of prevention barriers. The BN approach captured the dependency and nonlinear interaction of various factors that lead to the failure of a particular barrier. This approach overcomes the weakness of sequential causal relationships predicted in fault tree analysis. Two different logic gates (Noisy-OR and Leaky Noisy-OR) are used respectively in the BN of Figures 5.4-5.10 to evaluate failure probabilities of each prevention barrier.

Firstly, the failure probabilities of each factor (node) within each prevention barrier evaluated independently using Noisy-OR logic in section 5.3.3.1 were used to evaluate the failure probabilities of all prevention barriers when Noisy-OR and Leaky Noisy -OR logic were applied to the BNs (barriers) respectively. Secondly, the failure probabilities of each factor (node) within each prevention barrier evaluated independently using Leaky Noisy-OR logic in section 5.3.3.1 were used to evaluate the failure probabilities of all prevention barriers when Noisy-OR and Leaky Noisy-OR logic were applied to the BNs (barriers) respectively. For instance, to evaluate failure probability of RPB when Noisy-OR is applied to the BN (barrier), the following failure probabilities of different accident contributory factors (0.061, 0.084, 0.047 and 0.006) from Table 5.5 are substituted in the BN of Figure 5. 4 and the failure probability of the BN (RPB) is evaluated using Noisy-OR logic. The procedure is repeated by evaluating the failure probability of the BN (RPB) using Leaky Noisy-OR logic with the same failure probabilities of different factors (0.061, 0.084, 0.04 and 0.006). The two step procedure above is repeated with failure probabilities of accident contributory factors in RPB quantified with Leaky Noisy-OR logic given by Table 5.5.

The failure probability of each factor node (except the root node) is used as a leaky probability when the Leaky Noisy-OR gate is applied to the BN (barrier). The leaky probability of a top event in all barriers is 0.01. Following the procedures in this section, the failure probabilities of all prevention barriers using BN analysis is given in Table 5. 6.

#### 5.3.3.4. Estimation of consequences occurrence probabilities

The occurrence probability of non-sequential cause-sequential consequences is predicted by propagating the failure probabilities of prevention barriers obtained through BN analysis given in Table 5.6, through successive success or failure branches of the event tree of Figure 5. 2.

### 5.3.4. Non Sequential accident analysis (Non sequential cause- consequence relationship)

#### 5.3.4.1. Non sequential estimation of consequences occurrence probabilities

An event tree normally models an accident as a sequence of events with the underlying belief that the severity of the adverse events increases only through sequential failure of prevention barriers considered. However, this need not be true in a real life situation. To mitigate this weakness in the event tree, the BN is used to allow non sequential failure of prevention barriers to cause adverse events randomly. The failure probabilities of prevention barriers obtained through BN analysis given by Table 5.6 were substituted into the BN of Figure 5.11 and subsequently, occurrence probabilities of consequences were evaluated.

**Table 5. 6. Failure probabilities of prevention barrier obtained through fault tree and BN analysis.**

| Prevention barriers | Fault tree analysis | BN Analysis | | | |
|---|---|---|---|---|---|
| | | Noisy -OR values | | Leaky Noisy-OR values | |
| | | Noisy-OR | Leaky Noisy-OR | Noisy-OR | Leaky Noisy-OR |
| **RPB** | 0.0842 | 0.052 | 0.098 | 0.063 | 0.114 |
| **DPB** | 0.0025 | 0.041 | 0.121 | 0.046 | 0.13 |
| **IPB** | 0.026 | 0.07 | 0.113 | 0.075 | 0.124 |
| **EPB** | 0.0286 | 0.027 | 0.082 | 0.032 | 0.092 |
| **EMFPB** | 0.0229 | 0.071 | 0.145 | 0.077 | 0.155 |
| **HFPB** | 0.00145 | 0.024 | 0.046 | 0.029 | 0.057 |
| **OrFPB** | 0.0069 | 0.029 | 0.069 | 0.034 | 0.079 |

*Noisy-OR values are failure probabilities of different factors that contributed to the failure of prevention barriers obtained independently by using Noisy-OR logic. Leaky Noisy-OR values are failure probabilities of different factors that contributed to the failure of prevention barriers obtained independently by using Leaky Noisy-OR logics. Noisy-OR and Leaky Noisy-OR are logic gates used in the BN (barriers) respectively

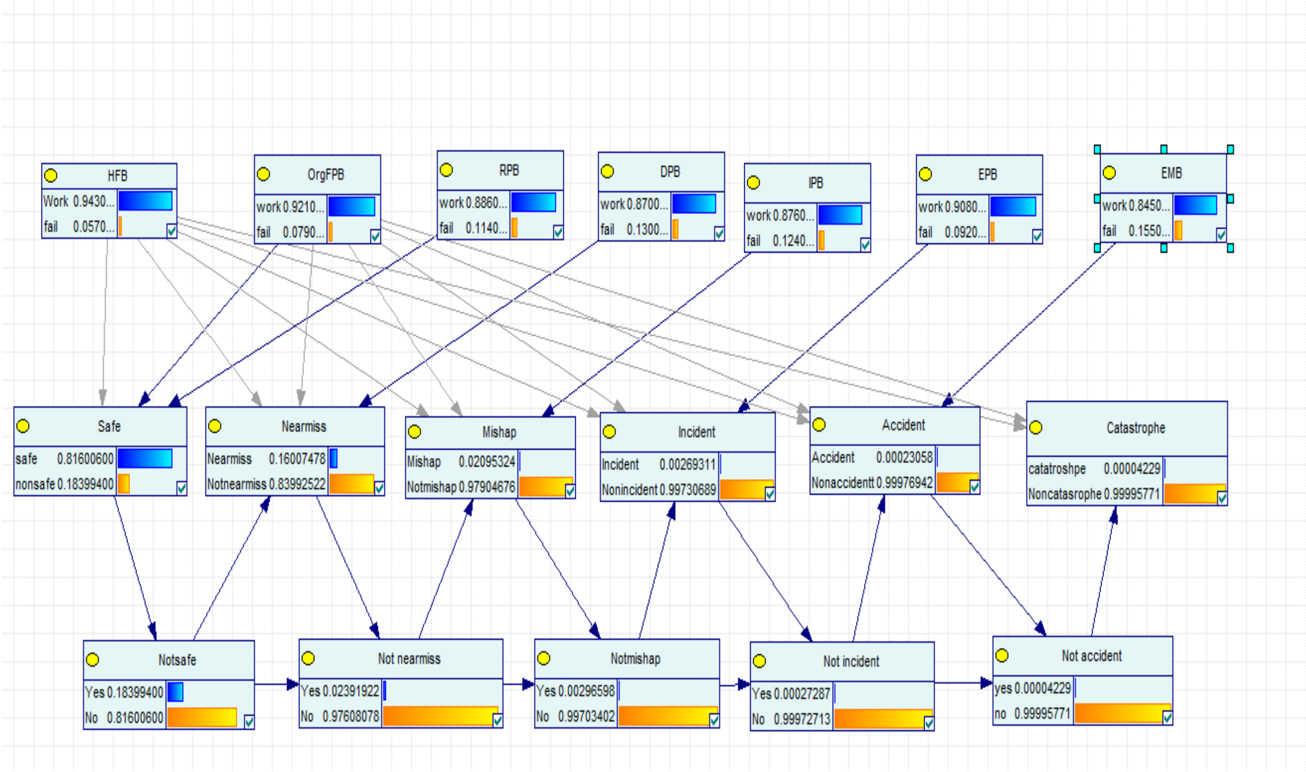



**Figure 5.11. GeNIe output of conditional event occurrence probability for non-sequential cause consequence relationship in the model.**

Table 5.7 provides the conditional probability table that is used in the BN of Figure 5.11.

**Table 5.7. Conditional probability table.**

| Consequences | Safety barrier | | | | | | |
|---|---|---|---|---|---|---|---|
| | HFPB | 0rFPB | RPB | DPB | IPB | EPB | EMB |
| Safe | Success | Success | Success | - | - | - | - |
| | Fail | Success | Success | - | - | - | - |
| Nearmiss | Success | Success | Fail | Success | - | - | - |
| | Fail | Success | Fail | Success | - | - | - |
| | Success | Fail | Fail | Success | - | - | - |
| | Fail | Fail | Fail | Success | - | - | - |
| Mishap | Success | Success | Fail | Fail | Success | - | - |
| | Fail | Success | Fail | Fail | Success | - | - |
| | Success | Fail | Fail | Fail | Success | - | - |
| | Fail | Fail | Fail | Fail | Success | - | - |
| Incident | Success | Success | Fail | Fail | Fail | Success | - |
| | Fail | Success | Fail | Fail | Fail | Success | - |
| | Success | Fail | Fail | Fail | Fail | Success | - |
| | Fail | Fail | Fail | Fail | Fail | Success | - |
| Accident | Success | Success | Fail | Fail | Fail | Fail | Success |
| | Fail | Success | Fail | Fail | Fail | Fail | Success |
| | Success | Fail | Fail | Fail | Fail | Fail | Success |
| | Fail | Fail | Fail | Fail | Fail | Fail | Success |
| Catastrophe | Success | Success | Fail | Fail | Fail | Fail | Fail |
| | Fail | Success | Fail | Fail | Fail | Fail | Fail |
| | Success | Fail | Fail | Fail | Fail | Fail | Fail |
| | Fail | Fail | Fail | Fail | Fail | Fail | Fail |

Following the entire procedure in section 5.3., Table 5.8 shows the occurrence probability of consequences as the accident modelling scenario moves gradually from a sequential to a non-sequential approach.

## 5. 4. Result and discussion

The failure probabilities of prevention barriers obtained through fault tree and BN analysis are presented in Table 5.6.

From fault tree analysis of the prevention barrier, it can be deduced that release, escalation, ignition, and dispersion prevention barriers have relatively high failure probabilities; it is the active failure of these prevention barriers that greatly contributed to the accident. The management of the Tesoro Anacortes Refinery had seriously compromised the safety of these prevention barriers, especially the release prevention barrier. If the management had used standard inspection and maintenance procedures, preventive measures could have been applied to prevent the release and further escalation.

The occurrence probability of the consequences for the accident model is shown in Table 5.8.

**Table 5. 8. Occurrence probability of different level of consequences.**

| Consequences | Fault tree with event tree (Sequential cause-consequence relationship ) | BN with event tree (Non sequential cause-sequential consequence relationship) | | | | BN withh BN (Non sequential cause- consequence relation ship) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Noisy-OR values | | Leaky Noisy-OR values | | Noisy-OR values | | Leaky Noisy-OR values | |
| | | Noisy-OR | Leaky-Noisy-OR | Noisy-OR | Leaky-Noisy-OR | Noisy-OR | Leaky Noisy-OR | Noisy-OR | Leaky-Noisy-OR |
| **Safe** | 9.09E-01 | 9.21E-01 | 8.40E-01 | 9.05E-01 | 8.16E-01 | 9.21E-01 | 8.40E-01 | 9.05E-01 | 8.16E-01 |
| **Nearmiss** | 8.97E-02 | 7.59E-02 | 1.42E-01 | 8.99E-02 | 1.61E-01 | 7.62E-02 | 1.41E-01 | 9.05E-02 | 1.60E-01 |
| **Mishap** | 7.83E-04 | 3.37E-03 | 1.57E-02 | 4.63E-03 | 1.98E-02 | 3.03E-03 | 1.72E-02 | 4.00E-03 | 2.10E-02 |
| **Incident** | 6.69E-06 | 1.98E-04 | 1.87E-03 | 2.94E-04 | 2.56E-03 | 2.21E-04 | 2.00E-03 | 3.16E-04 | 2.69E-03 |
| **Accident** | 1.89E-07 | 7.85E-06 | 1.72E-04 | 1.34E-05 | 2.65E-04 | 5.72E-06 | 1.54E-04 | 9.67E-06 | 2.30E-04 |
| **Catastrophe** | 3.14E-11 | 1.76E-08 | 2.08E-06 | 3.95E-08 | 4.13E-06 | 4.44E-07 | 2.61E-05 | 8.06E-07 | 4.23E-05 |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

The numbering below each column in Table 5.8 (1 - 9) is used to define relationships that exist in the accident modelling scenario described in Table 5.9.

**Table 5.9. Relationships that exist in accident modelling scenario of Table 5.8.**

| **Accident modeling scenar**io | **Relationship** | **Data** | **Dependency** |
|---|---|---|---|
| 1 | Sequential | Deterministic | Independent barrier<br>Independent causal factor |
| 2 | Sequential | Probabilistic (Noisy-OR) | Independent barrier<br>Conditional dependent causal factor (Noisy-OR) |
| 3 | Sequential | Probabilistic (Noisy-OR) | Independent barrier<br>Conditional dependent causal factor ( Leaky-Noisy-OR) |
| 4 | Sequential | Probabilistic (Leaky Noisy-OR) | Independent barrier<br>Conditional dependency causal factor (Noisy-OR) |
| 5 | Sequential | Probabilistic (Leaky Noisy-OR) | Independent barrier<br>Conditional dependent causal factor (Leaky Noisy-OR) |
| 6 | Non-Sequential | Probabilistic (Noisy-OR) | Dependent barrier<br>Conditional dependent causal factor (Leaky Noisy-OR) |
| 7 | Non-Sequential | Probabilistic (Noisy-OR) | Dependent barrier<br>Conditional dependent causal factor (Leaky Noisy-OR) |
| 8 | Non-Sequential | Probabilistic ( Leaky Noisy-OR) | Dependent barrier<br>Conditional dependent causal factor (Noisy-OR) |
| 9 | Non-Sequential | Probabilistic ( Leaky Noisy-OR) | Dependent barrier<br>Conditional dependent causal factor (Leaky Noisy-OR) |

The plots of the data in Table 5.8 are shown in Figure 5.12.

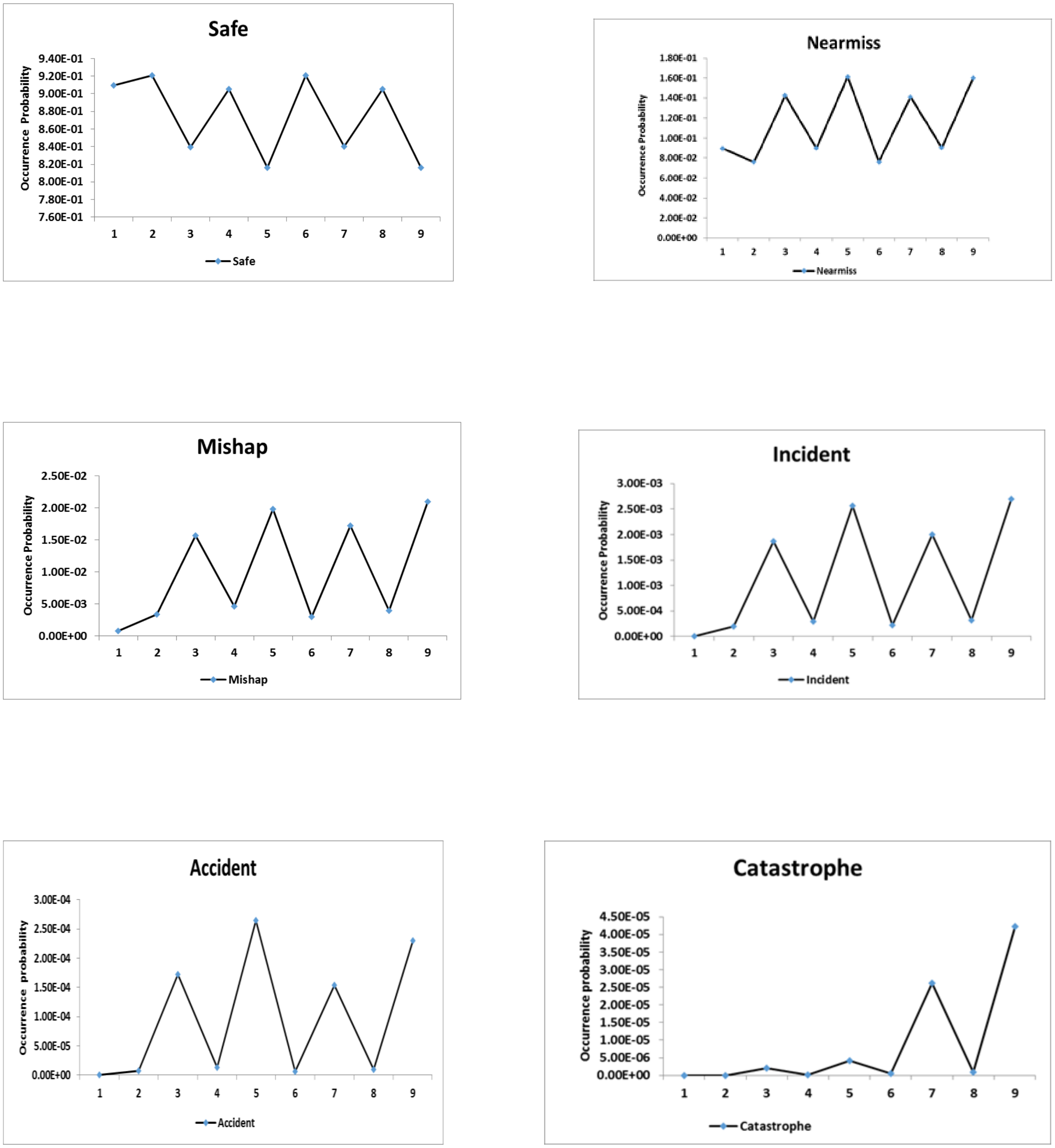



**Figure 5. 12. Plot of Consequence occurrence probabilities versus accident modelling scenario.**

In Figure 5.12, it is worth observing that, as the accident modeling scenario changes gradually from a sequential approach to a non-sequential approach, there is about a 10% change in occurrence probability of "safe" to that of a sequential approach. Also, the estimated range of occurrence probability of catastrophe changes significantly up to the order of $10^5$ which significantly increases the risk involved. This indicates the importance of modeling non sequential cause-consequence relationships in the accident modelling process and highlights the importance of using BN to evaluate non sequential cause-consequence relationships in the accident modeling process. The following are some observations derived from Figure 5.12.

- There is a significant difference between the results of accident modeling scenario #1 and #9. This highlights the importance of considering dependency among causal factors and the non-sequential nature of an accident process.
- There is a significant difference between the results of accident modeling scenario #8 and #9. This indicates the importance of defining a proper dependency among causal factors.
- There is little difference between the results of accident modeling scenario #7 and #9. This shows that the quality of data may not have a significant impact on the model outcomes given that the dependency among causal factors is considered and the model is developed based on a non-sequential structure.

This present study has illustrated that BN is an effective technique for modeling various nonlinear interactions within prevention barriers and non-sequential failure of safety barriers to cause adverse events. Relaxation strategies can be accurately used to model conditional dependencies in the BN.

## 5. 5. Conclusions

This present study has demonstrated the use of BN in modelling conditional dependencies among accident contributory factors within safety barriers and non-sequential failure of safety barriers to cause adverse events. In general, the modelling flexibility of the BN structure can accommodate various kind of conditional dependencies that cannot be readily included in FT structure. The effectiveness of the proposed non-sequential causes-consequence barrier-based process accident model was partially validated through the application of the model to the Tesoro Anacortes Refinery accident. The main source of uncertainty in accident models (such as FT, ET) is the ignorance of interdependency among accident contributory factors and the assumption of linear event sequence. Additionally, the uncertainty is also caused by the inappropriate use of logic gates in the accident models. This paper attempts to highlight the importance of modelling the interdependency of accident contributory factors, nonlinear event sequence, and the selection of appropriate logic gates to the reduction of the above-mentioned uncertainty. BN is more appropriate to represent complex dependencies among prevention barriers and to include uncertainty in modelling. BN has high capability for abductive reasoning and the ability to handle uncertainty makes it a more appropriate technique for analysing accidents. This accident model provides methodology for predicting a process accident based on nonlinear interactions within prevention barriers and non-sequential failure of prevention barriers to cause adverse events. Application of this in models in predicting accident occurrence probability will help to take early remedial actions to prevent process accidents and this consequently provides additional valuable information for process safety management.

## Acknowledgement

The authors thankfully acknowledge the financial support provided by, Natural Science and Engineering Council of Canada and Canada Research Chair (CRC) Program

## References

Al-shanini, A., Ahmad, A. & Khan, F., 2014. Accident modelling and safety measure design of a hydrogen station. *International Journal of Hydrogen Energy*, 39(35), pp.20362–20370.

Al-shanini, A., Ahamad, A. & Khan, F., 2014. Accident modelling and analysis in process industries. *Journal of Loss Prevention in the Process Industries*, 32, pp.319–334.

Attwood, D., Khan, F. & Veitch, B., 2006. Occupational accident models-Where have we been and where are we going? *Journal of Loss Prevention in the Process Industries*, 19, pp.664–682.

Baksh, A. et al., 2015. Network based approach for predictive accident modelling. *Safety Science*, 80, pp.274–287.

Bobbio, a. et al., 2001. Improving the analysis of dependable systems by mapping Fault Trees into Bayesian Networks. *Reliability Engineering and System Safety*, 71, pp.249–260.

Diez, F.. & Druzdzel, M.., 2007. *Canonical probabilistic models for Knowledge Engineering. Technical Report CISIAD-06-01.*,

Heckerman, D. & Breese, J.S., 1996. Causal independence for probability assessment and inference using Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, 26(6), pp.826–831.

Heinrich, W.., 1941. *Industrial Accident Prevention,* McGraw-Hill, NY.

Jensen, F.. & Nielsen, T.., 2007. *Bayesian Networks and Decision Graphs* 2 edition., Springer, New York.

Katsakiori, P., Sakellaropoulos, G. & Manatakis, E., 2009. Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. *Safety Science*, 47(7), pp.1007–1015.

Khakzad, N., Khan, F. & Amyotte, P., 2013. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection*, 91(1-2), pp.46–53.

Khakzad, N., Khan, F. & Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety*, 96(8), pp.925–932.

Kujath, M.F., Amyotte, P.R. & Khan, F.I., 2010. A conceptual offshore oil and gas process accident model. *Journal of Loss Prevention in the Process Industries*, 23(2), pp.323–330.

Nivolianitou, Z.S., Leopoulos, V.N. & Konstantinidou, M., 2004. Comparison of techniques for accident scenario analysis in hazardous systems. *Journal of Loss Prevention in the Process Industries*, 17, pp.467–475.

Okoh, P. & Haugen, S., 2014. Application of inherent safety to maintenance-related major accident prevention on offshore installations. *Chemical Engineering Transactions*, 36, pp.175–180.

Oniśko, A., Druzdzel, M.J. & Wasyluk, H., 2001. Learning Bayesian network parameters from small data sets: Application of Noisy-OR gates. *International Journal of Approximate Reasoning*, 27, pp.165–182.

Pearl, J., 1998. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmannn, San Francisco, California.

Qureshi, Z.H., 2008. A Review of Accident Modelling Approaches for Complex Critical Sociotechnical Systems. , 86, p.72.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. *Safety Science*, 27(2-3), pp.183–213.

Rathnayaka, S., Khan, F. & Amayotte, P., 2013. Accident modeling and risk assessment framework for safety critical decision-making: application to deepwater drilling operation. *Journal of Risk and Reliability*, 227(1), pp.86–105.

Rathnayaka, S., Khan, F. & Amyotte, P., 2012. Accident modeling approach for safety assessment in an LNG processing facility. *Journal of Loss Prevention in the Process Industries*, 25(2), pp.414–423.

Rathnayaka, S., Khan, F. & Amyotte, P., 2010. SHIPP methodology : Predictive accident modeling approach . Part II . Validation with case study. *Process Safety and Environmental Protection*, 89(2), pp.75–88.

Rathnayaka, S., Khan, F. & Amyotte, P., 2011. SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection*, 89(3), pp.151–164.

Reason & J, 1990. *Human error,* University press,Cambridge.

Sklet, S., 2004. Comparison of some selected methods for accident investigation. *Journal of Hazardous Materials*, 111(April), pp.29–37.

Tan, Q. et al., 2013. Dynamic accident modeling for high-sulfur natural gas gathering station. *Process Safety and Environmental Protection*, 92(6), pp.565–576.

U.S. Chemical Safety and Hazard Investigation Board, 2014. Catastrophic Rucpture of Heat Exchanger Tesoro Anacortes Refinery.

Zagorecki, A. & Druzdzel, M., 2004. An Empirical Study of Probability Elicitation Under Noisy-OR Assumption. *FLAIRS Conference*, pp.880–885.

# Chapter 6

## 6.0 Dynamic Failure Analysis of Process Systems Using Principal Component Analysis and Bayesian Network

### Preface

*A version of this chapter has been published in the **Journal of Industrial &Engineering Chemistry Research 2017; 56:2094-2106.** I am the primary author. Co-author Faisal Khan provided fundamental understanding, assisted in developing the conceptual model and subsequently translated this to the numerical model. Co-author Ming Yang provided much needed support in implementing the concept and testing the model. I carried out most of the data collection and analysis. I prepared the first draft of the manuscript and subsequently revised the manuscript, based on the feedback from co-authors and also a peer review process. The two co-authors assisted in developing the concept and testing the model, reviewed and corrected the model and results. They also contributed to the review and revision of the manuscript.*

### Abstract

Modern industrial processes are highly instrumented with more frequent recording of data. This provides abundant data for safety analysis; however, these data resources have not been well used. This paper presents an integrated dynamic failure prediction analysis approach using principal component analysis (PCA) and the Bayesian network (BN). The key process variables that contribute the most to process performance variations are detected with PCA; while the Bayesian network is adopted to model the interactions among these variables to detect faults and predict the time-dependent probability of system failure. The proposed integrated approach uses big data analysis. The structure of BN is learned using past historical data. The developed BN is used to detect faults and estimate system failure risk. The risk is updated subsequently as new process

information is collected. The updated risk is used as a decision-making parameter. The proposed approach is validated through a case of a crude oil distillation unit operation**.**

**Keywords:** Principal components analysis, Bayesian network, Process safety, Accident probability estimation, Dynamic failure prediction

## 6.1. Introduction

The sudden surge in complexity of modern process systems greatly improves the versatility and productivity of systems and also poses a challenge in ensuring safe operation of these process systems. This sudden surge in complexity is proportionally connected to the enormous number of variables on which the system depends. To guarantee safe and optimal operation of the system, it is of paramount importance that the states of these variables be monitored in real time. Real time process variables monitoring gives rise to a generation of enormously high dimensional data vector. Due to the increase in dimensionality, the relationship among system variables becomes extremely complex and non-linear. To guarantee the safety of the system, identification of these non-linear relationships that exist among the monitored variables is of paramount importance. Faults (abnormal behaviours) arises as a result of disturbances in the relationships among system variables. The progression of the fault decreases the safety of the process systems (Yu et al. 2015a; Yu et al. 2015b).

A dimensionality reduction technique is normally used in the monitoring of complex systems. Generally, Process parameters that vividly described the maximum variances of the system are collected to create a new set of variables and usually parameters denoting least variance are often ignored. The system performance is monitored using the variables that possess fewer dimensionality (Yu et al. 2015c). The transformed historical process data can be used as "priori knowledge to a diagnostic system" in various ways. This process is referred to as feature extraction

(Venkatasubramanian et al. 2003). Feature extraction can be done by qualitative and quantitative approaches. The two main techniques that extract information qualitatively from process history are the trend and expert modelling techniques (Venkatasubramanian et al. 2003). Also, the two Principal techniques that quantitatively extract information are: statistical and non-statistical techniques. The most widely used statistical feature extraction techniques for process monitoring are partial least square (PLS) and Principal Component analysis (PCA). Good examples of non-statistical classifiers are neutral networks (Venkatasubramanian et al. 2003). Principal component analysis (PCA) is used as an appropriate technique for data compression and information extraction in process monitoring data (Li & Qin 2001; Wise & Gallagher 1996).

Recently, PCA is one of the famously used data driven techniques to detect instant fault in chemical process industries it has been used in wide range of applications (He et al. 2006; Yu et al. 2015c). Zadakbar et al. (2012) used PCA in dynamic risk assessment of distillation column and dissolution tank. A robust PCA technique is used for fault detection of the Tennessee Eastman chemical process (Pan et al. 2016). A new process monitoring technique that detects fault automatically in a process system using dynamic weighted principal component analysis was recently proposed. The approach was demonstrated on the Tennessee Eastman chemical process (Fei & Liu 2016). PCA has been applied to propose a novel sensor selection technique for monitoring the performance of wind turbines (Wang et al. 2016). Similarly, Recursive Kernel principal components analysis (RKPCA) has been applied to monitor time –varying system. The approach is validated on Penicillin fermentation process (Zhang et al. 2012).

Databases are increasing rapidly in many modern industrial processes. There are many potential opportunities for using databases to assist the construction of probabilistic networks. The probabilistic network structure generated from databases will provide a precise and concise representation of probabilistic dependencies that exist among variables (Cooper & Herskovits 1991). The subjectivity in the development of the Bayesian network is eliminated, if the structure is learn from databases.

One of the Bayesian score techniques that has been widely used to learn structure of Bayesian network is Tree Augmented Naïve Bayes (TAN). This technique has been widely used in different disciplines to construct probabilistic network from data. TAN Bayes classifier has been used to construct Bayesian network structure from data and subsequently the probability of rock burst is predicted from the Bayesian network (Li et al. 2017). The technique of structure learning has been implemented to develop early warning system for chemical process operations (Wang et al. 2015).

Probabilistic Risk assessment (PRA) has been widely used for failure prediction of the process system (Khan et al. 2015) ,Probabilistic Risk assessment (PRA) is the numerical study of the risk. It assesses the most significant risk contributors to the risk of the process system. It gives an accurate numerical assessment for good understanding of the system. PRA is used as a decision making parameter. In this approach, probabilities of series of events leading to hazards are estimated and the corresponding consequences are predicted. The risk (expected loss) is measured as the product of frequency and it consequences (Khan et al. 2015; Mohammad 2006).

This paper aim to propose a new methodology that integrates PCA and the Bayesian network to detect faults and predict the probability of failure using real-time process data. The methodology developed analyzes an accident and its associated consequences in real time. This paper is organized as follows. Section 6. 2 presents a brief description of PCA and its application in process data analysis. Section 6.3 describes how a Bayesian network structure can be derived from the data. The proposed methodology is presented in Section 6.4. In Section 6.5, a case study of a crude distillation unit is used to show the application of the proposed approach. Finally, Section 6.7 provides the conclusion.

## 6.2. The Theory of Principal Component Analysis (PCA)

PCA has been extensively applied as a technique for finding faults. PCA is used in processing extremely correlated process variables data with high dimensionality. PCA lessens the dimensionality of the raw data set by subjecting the data set unto a subspace of smaller dimensionality including specifying a series of new variables to retain or protect the main original data information (He et al. 2006; Yu et al. 2014). PCA theorem is based on singular value decomposition (SVD) of the covariance matrix of the process variables in the direction that describes the highest variation of data (Venkatasubramanian et al. 2003). PCA provides a mechanism to reduce a complex data set to a smaller dimension to show the sometimes concealed basic structures that often under lie it (Shlens 2014).

Lately, original process variables data can be vividly described using fewer factors than original process data and vital information will still be retained. Consequently, the data overload that is normally encountered in industrial process monitoring is resolved. Similarly, PCA gives linear set of process variables that give detail performance description of the process. The integrations of these process variables are used as strong indicator of process performance than any individual process variable (Wise & Gallagher 1996). PCA gives a superior performance compared to other statistical process monitoring techniques for analyzing raw historical plant data (Joe Qin 2003; MacGregor & Kourti 1995).

Consider a process data matrix $\mathbf{M}^{n \times p}$. Let p represent the total number of process variable being monitored and n the represent the total number of samples. All the columns of process data matrix

M is a mean centered and at the same time scaled process variable with a covariance matrix of ∑. The rows in the **matrix M, $m_1$, $m_2$,..Mn** are p vectors equivalent to samples. Likewise, the columns are n vectors equivalent to variable. Applying Singular value decomposition to ∑, the covariance matrix is decomposed to a diagonal matrix **K** through a definite orthonormal p x p matrix **U,** i.e., $\sum = \mathbf{UKU^T}$ . The column of matrix U are usually referred to as "principal component loading vectors" The diagonal elements of K are named the eigenvalue of covariance matrix.

The scores **T** are expressed as

$$\mathbf{T} = \mathbf{MU} \qquad (\mathbf{6.1})$$

Equally, PCA decomposed **M** as:

$$\mathbf{M} = \mathbf{TU^T} \qquad (\mathbf{6.2})$$

The n x p matrix **T** = ($\theta_1$, $\theta_2$… $\theta_p$) contains the scores of principal components (PCs) that define all n observations. In n process monitoring, PCs are monitored instead of monitoring individual variables (Venkatasubramanian et al. 2003; Zadakbar et al. 2012).

## 6.3. Structure-learning of Bayesian network

A Bayesian network (BN) is an extensively used graphical structure that encodes probabilistic dependencies among a collection of variables of interest. The BN has been used extensively to represent accident scenarios in offshore and maritime systems, beginning from initiating factors and ending with potential consequences (Baksh et al. 2015; Khakzad et al. 2013). The BNs are acyclic graphs, where the nodes signifying the variables are linked to each other by arcs that indicate causal or dependent interactions among the connected nodes. If a causative probabilistic relationship exists among variables of interest, then the nodes are linked together by a direct arc

(Baksh et al. 2015). A conditional probability table (CPT) is allotted to all nodes to signify conditional interactions among the nodes connected. (Bobbio et al. 2001; Khakzad et al. 2013) The joint probability distribution P (U) of a collection of discrete random variables U= {$A_1$…,$A_n$}, integrated in as:

$$\mathbf{P(U) = \prod_{i=1}^{n} P(A_i \mid P_{a(A_i)})} \qquad (6.3)$$

where Pa ($A_i$) is the parent of variable $A_i$ and P (U) is the joint probability distribution of variables (Adedigba et al. 2016; Jensen & Nielsen 2007; Pearl 1998).

Formerly, the conditional probabilities table (CTP) in BN were assessed by expert judgement and the direct acyclic graphs (DAG) in BN were usually hand-constructed by a domain expert. Eliciting BN from a domain expert can be an extremely difficult task for large networks (Neapolitan 2004). Consequently, researchers have developed methods that can both learn DAG (structure) and CPT (parameter) from directly observed data. The methods for learning the structure of Bayesian networks are: constraint-based and Bayesian score- based methods (Dash & Druzdzel 1999; Jensen & Nielsen 2007).

### 6.3.1. Constraint-based Learning Methods

The constraint based methods scrutinise the data for a set of conditional independence relations. The collection of the conditional independence relations are used to infer the Markov equivalent class of an underlying graph (Dash & Druzdzel 1999).

The constraint based approaches are beneficial because they are moderately fast and possess a good capability to handle latent variables. However, the weaknesses of these methods are:

(1) They use random significance level to determine independencies.

(2) They are unstable: error generated in the search process could result in a significantly different graph (Dash & Druzdzel 1999).

Two famous constraint based algorithms are the PC and FCI algorithms. The PC algorithm is based on the assumption that no hidden variables exist and the FCI algorithm has the capability of learning the underlying relationships, by assuming that latent variables are present in the data (Spirtes et al. 1993; Dash & Druzdzel 1999) .

### 6.3.2. Bayesian score- based methods

The Bayesian score methods apply a search and score technique to search the space of DAG, to produce a series of candidate Bayesian networks. They use the posterior density as the scoring function to find a candidate with the highest score (Jensen & Nielsen 2007; Dash & Druzdzel 1999). The score reflects the likelihood of using the structure to generate the data at hand(Jensen & Nielsen 2007), These methods exhibit the following advantages:

(1) They are applicable with very small data, where conditional independence tests might not hold.

(2) They have capability of handling incomplete data in the database (Dash & Druzdzel 1999).

Tree Augmented Naïve Bayes (TAN) is principally a modification of the Naïve Bayes classifier (Dash & Druzdzel 1999; Dash & Druzdzel 2002). The Naïve Bayes classifier has performed extremely well on both small and relatively large data. The fundamental assumption made by Naïve Bayes is that all the essential features in the dataset are conditionally independent as long as the value of the class is known (Cerquides & Lopez De Antaras 2003).This strong underlying assumption is very likely not to be fulfilled; nevertheless, the Naïve Bayes classifier performs excellently in practice when strong dependencies hold in the dataset. TAN relaxes this assumption made by Naïve Bayes and keeps the same way of reasoning. TAN has demonstrated excellent performance irrespective of its simplicity and strong independence assumptions. TAN is more coherent and has a better performance than Naïve Bayes (Cerquides & Lopez De Antaras 2003) .

In TAN model, class nodes are openly linked to all the attribute nodes with the directed edges. Each attribute node has maximum of one parent node from other attribute nodes. Directed cycle is not permitted among the attribute nodes (Jiang et al. 2012). Figure 6.1 shows a TAN Bayes model.

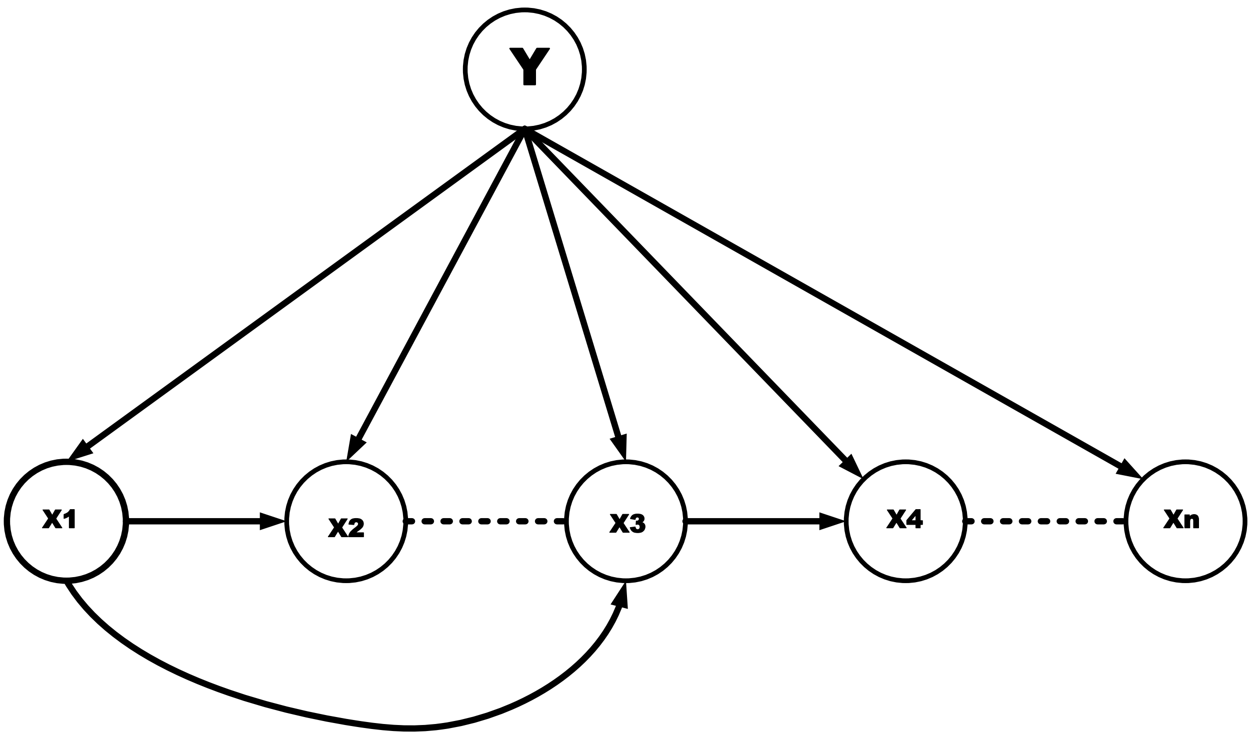



**Figure 6.1. An example of Tree Augmented Naïve (TAN) Bayes model(BenaditP & FrancisF 2015).**

Let$X_1$,$X_2$,…,$X_n$ denotes n attributes and Y represents the class variable. The Conditional mutual information ($I_p$) is used in in building TAN model. The conditional mutual information is calculated using the formula below (BenaditP & FrancisF 2015).

$$I_{P}(X,Y|Z)= \sum_{x,y,z} P\,(x,y,z)log\left(\frac{P(x,y|z)}{P(x|z)P(y|z)}\right) \quad (6.4)$$

The TAN algorithm procedure consists of the following steps (Friedman et al. 1997; Jensen & Nielsen 2007; Jiang et al. 2012; BenaditP & FrancisF 2015). These steps are:

- Input the training data set Z
- Calculate the conditional mutual information $I_p(X_i, X_j|C)$ between each pair of attributes, i ≠ j.
- Construct the complete undirected graphs in which the vertices are the attributes n variables. The edges are weighted based on the pairwise mutual information, $X_i\ toX_j$ by $I_p(X_i, X_j|C)$.
- Construct the highest weight spanning tree.
- Change the undirected graph to a direct graph by selecting the class variable as the root node and setting the direction of the links to be outwards from it.
- Build a TAN model by drawing an arc from the class variable to all other variables.

## 6. 4. The proposed methodology

### 6.4.1. Projecting historical process data into PCA Space

Historical analysis of past real time data is important because it gives detailed information about past plant performance. Detailed analysis of plant history information can be used as prior information to predict how the plant will operate in its current state. PCA is a commonly used statistical process control technique for monitoring a huge number of process variables in process industries. PCA has the capability of compressing the data into low dimensional spaces which retain most of the vital information in the data set (He et al. 2006; Yu et al. 2014). The goal of projecting past real time data into PCA space is to find the principal components of the numerous process variables data monitored. PCA gives the principal components (PCs) that fully describe the highest variation of the data (Venkatasubramanian et al. 2003). Operating personnel can use the PCs to monitor plant performance instead of using all the monitored process variables.

### 6.4.2 Construction of probabilistic network structure from analysis of historical process database

In spite of the enormous amount of process data stored in databases for most modern industrial processes, little analysis and interpretation of these data are taking place (Kresta et al. 1991; Wise & Gallagher 1996). Historical real time data of the plant could be systematically analysed and critical points set for each of the process variables in the database to obtain their states. In the present study, the states of historical real time data are used to construct a probabilistic network using the Tree Augmented Naïve Bayes algorithm. The probabilistic network structure constructed from databases will provide an exact depiction of probabilistic dependencies that is present among variables (Cooper & Herskovits 1991).

The probabilistic network structure generated based on historical real time data is used as a model to predict the real time probability of failure based on the current operating state of the operation. Each of the process variable nodes in the probabilistic network is linked to a failure node using the logic gate AND. The AND logic is used because faults in principal components (variables) will be propagated to all other components (process variables) based on the conditional probabilistic dependencies among the process variables in the network. Hence, real time probability of failure is estimated based on the resultant conditional probabilistic dependency effect of all process variables in the probabilistic networks structure. The essence of determining the principal components using PCA is that, PCA gives the principal components (PCs) that fully describe the highest variation of the data (Venkatasubramanian et al. 2003). Operating personnel can use the PCs to monitor plant performance instead of using all the monitored process variables.

The principal components (PCs) determined in the probabilistic network through PCA analysis are continuously monitored at each time interval. The state of the PCs and their probability are continuously evaluated. The current probability of the PCs' state obtained is fed back into the probabilistic network to update it and subsequently real time probability of failure is predicted.

The following gives the procedure used to predict the real time probability of failure.

(1) Constructing the probabilistic network structure from historical data using TAN algorithm. TAN algorithm give both structure and the conditional dependencies among process variables.
(2) Computing the principal components (PCs) from historical data using principal component analysis (PCA).
(3) Computing the failure probability of PCs from process monitoring data.
(4) Computing the safe probabilities of the PCs.
(5) Linking all the nodes in the probabilistic network structure to failure node.
(6) Assigning an appropriate logic gate to failure node. AND logic gate is assigned in this case study.

(7) Computing the Conditional probability table (CPT) for different states of the failure node.

(8) Plugging back the computed probabilities of PCs computed in steps 3 and 4 into the probabilistic network structure.

(9) Evaluating the probabilistic network structure using both the CPTs and the probabilities of state of the failure node (True/false and Success/ failure) based on the state under consideration.

The real time probability of failure (R) is obtained using the equation below.

$$\boldsymbol{R} = \sum_{i=1}^{i=n} \boldsymbol{C_i} \tag{6.5}$$

Where R is the real time probability, i is the state of the failure node, and $C_i$ is the conditional probability for different states of failure node.

For illustration purpose, Figure 6.2 provides an example of probabilistic network structure generated from the historical data.

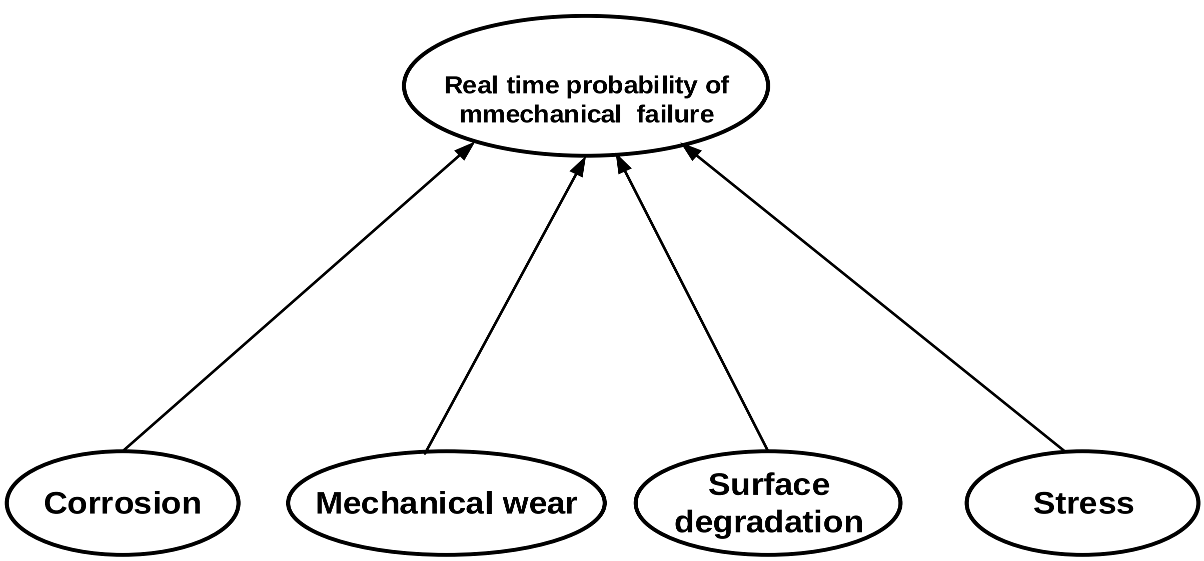



**Figure 6. 2. Probabilistic networks structure of causes of mechanical failure generated from historical data.**

The probabilistic networks structure represents the causes of mechanical failure of process equipment. Table 6.1 gives the failure probabilities of the mechanical failure. Based on the steps given above, the real time failure probability of mechanical failure is computed as shown in Table 6.2

**Table 6. 1. Failure and Safe probabilities of causes of mechanical failure.**

| No | **Causes of mechanical failure** | **Failure probability** | **Safe probability** |
|---|---|---|---|
| 1 | Corrosion | 0.2550 | 0.7450 |
| 2 | Mechanical wear | 0.0270 | 0.9730 |
| 3 | Surface degradation | 0.1250 | 0.8750 |
| 4 | Stress | 0.1330 | 0.8670 |

**Table 6.2 - Real time probability of mechanical failure.**

| State (i) | Corrosion | Mechanical wear | Surface degradation | Stress | Success | Failure | Conditonal Probability of mechanical failure for different states ( C ) |
|---|---|---|---|---|---|---|---|
| 1 | F | F | F | F | 0 | 1 | 1 * 0.2550 * 0.0270 * 0.1250 * 0.1330 = 1.14463E-04 |
| 2 | F | F | F | T | 1 | 0 | 0 * 0.2550 * 0.0270 * 0.1250 * 0.8670 = 0 |
| 3 | F | F | T | F | 1 | 0 | 0 * 0.2550 * 0.0270 * 0.8750 * 0.1330 = 0 |
| 4 | F | F | T | T | 1 | 0 | 0 * 0.2550 * 0.0270 * 0.8750 * 0.8670 = 0 |
| 5 | F | T | F | F | 1 | 0 | 0 * 0.2550 * 0.9730 * 0.1250 * 0.1330 = 0 |
| 6 | F | T | F | T | 1 | 0 | 0 * 0.2550 * 0.9730 * 0.1250 * 0.8670 = 0 |
| 7 | F | T | T | F | 1 | 0 | 0 * 0.2550 * 0.9730 * 0.8750 * 0.1330 = 0 |
| 8 | F | T | T | T | 1 | 0 | 0 * 0.2550 * 0.9730 * 0.8750 * 0.8670 = 0 |
| 9 | T | F | F | F | 1 | 0 | 0 * 0.7450 * 0.0270 * 0.1250 * 0.1330 = 0 |
| 10 | T | F | F | T | 1 | 0 | 0 * 0.7450 * 0.0270 * 0.1250 * 0.8670 = 0 |
| 11 | T | F | T | F | 1 | 0 | 0 * 0.7450 * 0.0270 * 0.8750 * 0.1330 = 0 |
| 12 | T | F | T | T | 1 | 0 | 0 * 0.7450 * 0.0270 * 0.8750 * 0.8670 = 0 |
| 13 | T | T | F | F | 1 | 0 | 0 * 0.7450 * 0.9730 * 0.1250 * 0.1330 = 0 |
| 14 | T | T | F | T | 1 | 0 | 0 * 0.7450 * 0.9730 * 0.1250 * 0.8670 = 0 |
| 15 | T | T | T | F | 1 | 0 | 0 * 0.7450 * 0.9730 * 0.8750 * 0.1330 = 0 |
| 16 | T | T | T | T | 1 | 0 | 0 * 0.7450 * 0.9730 * 0.8750 * 0.8670 = 0 |
| | | | | | | | The real time probabiity of mechanical failure is sum of all states = 1.14463E-04 |

*** F, mean False, T mean True, Success means no mechanical failure, Failure, means mechanical failure.**

### 6.4.3. Hazard identification and analysis

The primary objective of this stage is to find the likely process hazards and subsequently analyse how the process hazards will occur. A detailed review of these techniques with their strengths and weaknesses is given by Khan et al . (1998) .When the process hazards and their underlying factors have been known, it is of paramount importance to assess accident pathway (sequences) and their corresponding consequences. The prevention barriers identified is placed alongside the accident path (sequence) to avoid and mitigate the effect of the accident. Accidents occur because of failure of significant prevention barriers (Rathnayaka et al. 2011). A detailed hazard analysis and thorough identification of relevant prevention barriers for a crude oil distillation unit (CDU) has been done in previous work. Comprehensive description of the prevention barriers can be found in Adedigba et al. (2016).

#### *6.4.3.1. Failure probability assessment of prevention (safety) barriers*

Fault tree analysis (FTA) is a reliable tool that is widely used to predict the likelihood of a hazard, as a result of failure events. Fault trees are both graphical and logical vivid explanation of several categories of failure events. To draw fault trees, hazards are foremost determined and the series of events causing the hazard are identified. The topmost event in the fault tree denotes a main accident instigating hazard. FTA depend on both Boolean algebra and probability theory (Khakzad et al. 2011; Rajakarunakaran et al. 2015). A fault tree analysis could be quantitative, qualitative or a combination of both ( Adedigba et al. 2016; Khakzad et al. 2011).The prevention barriers identified in the model are systematically analyzed with fault tree to represent a causative relationship and subsequently there failure probabilities estimated.

### *6.4.3. 2. Event tree construction*

Event tree is an inductive systematic technique that starts with a specified accident initiating event and terminates with all the feasible consequences normally called the "end state consequences" of the event tree. Event tree techniques are widely used to denote incident scenarios. It represents a probable sequence related with an accident initiating events that transit through successive prevention barriers and terminating with ultimate consequences (Nývlt & Rausand 2012).

The likelihoods (probabilities) of end state consequences $P(C_k)$ are quantify by equation 6.6.

$$\mathbf{P(C_K)} = \prod_{\mathbf{j\,\epsilon\,SB_k}} \mathbf{x_i^{\theta_{i,k}}\,(1 - x_i)^{1-\theta_{i,k}}} \qquad (6.6)$$

where $SB_k$ represents the prevention barrier related to level k; and $\theta_{i,k} = 1$ whenever a level k failure transits through the failure branch of safety ( prevention) barrier i; $\theta_{i,k} = 0$ whenever a level k failure transits through the success branch of safety( prevention) barrier i. $x_i$ is the failure probability of prevention (safety) barriers ( Adedigba et al. 2016; Rathnayaka et al. 2010).

The proposed methodology incorporates BN structure from dataset, fault and event trees analysis to denote causes and consequences. The initiating event in the model is the real time probability of failure predicted from the probabilistic network structure. The proposed methodology is represented by the flowchart given in Figure 6. 3. The event tree for the model is presented in Figure 6.4.

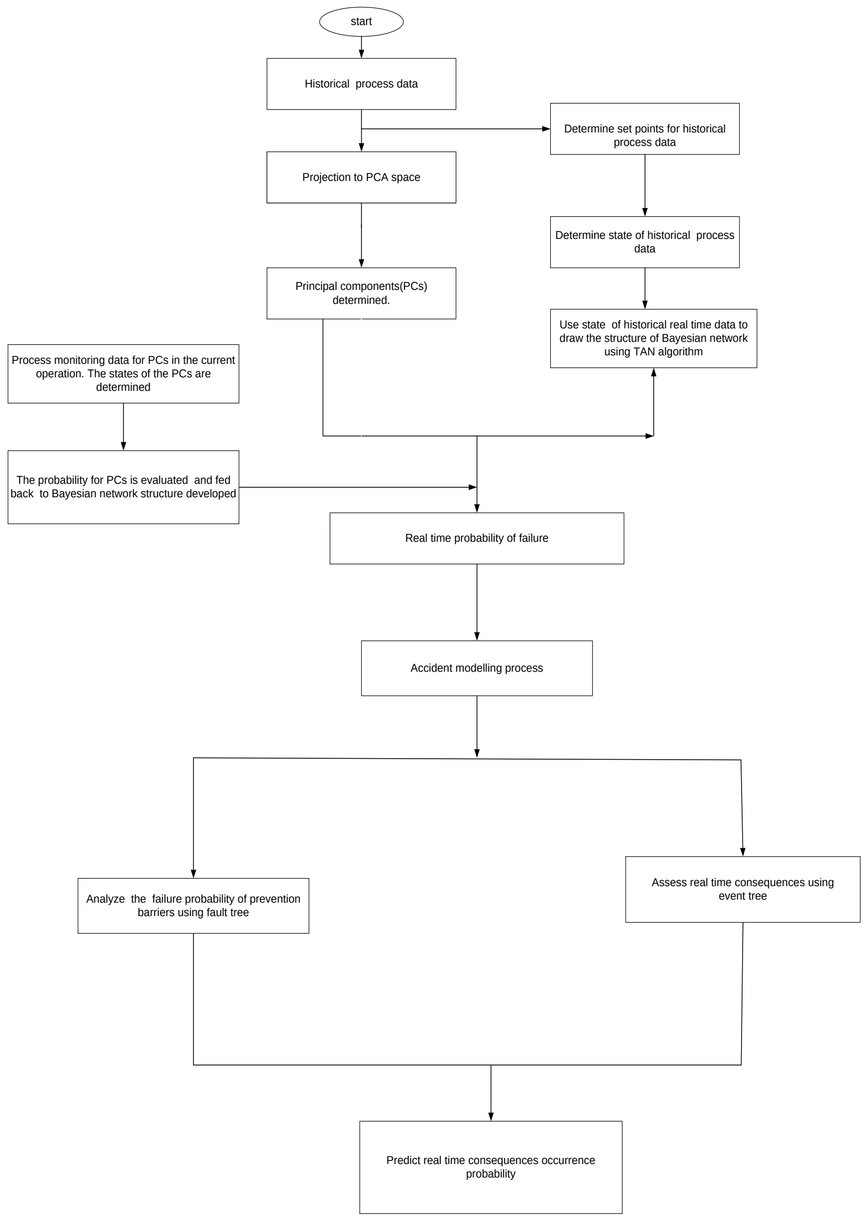



**Figure 6. 3. Proposed accident modelling methodology flowchart.**

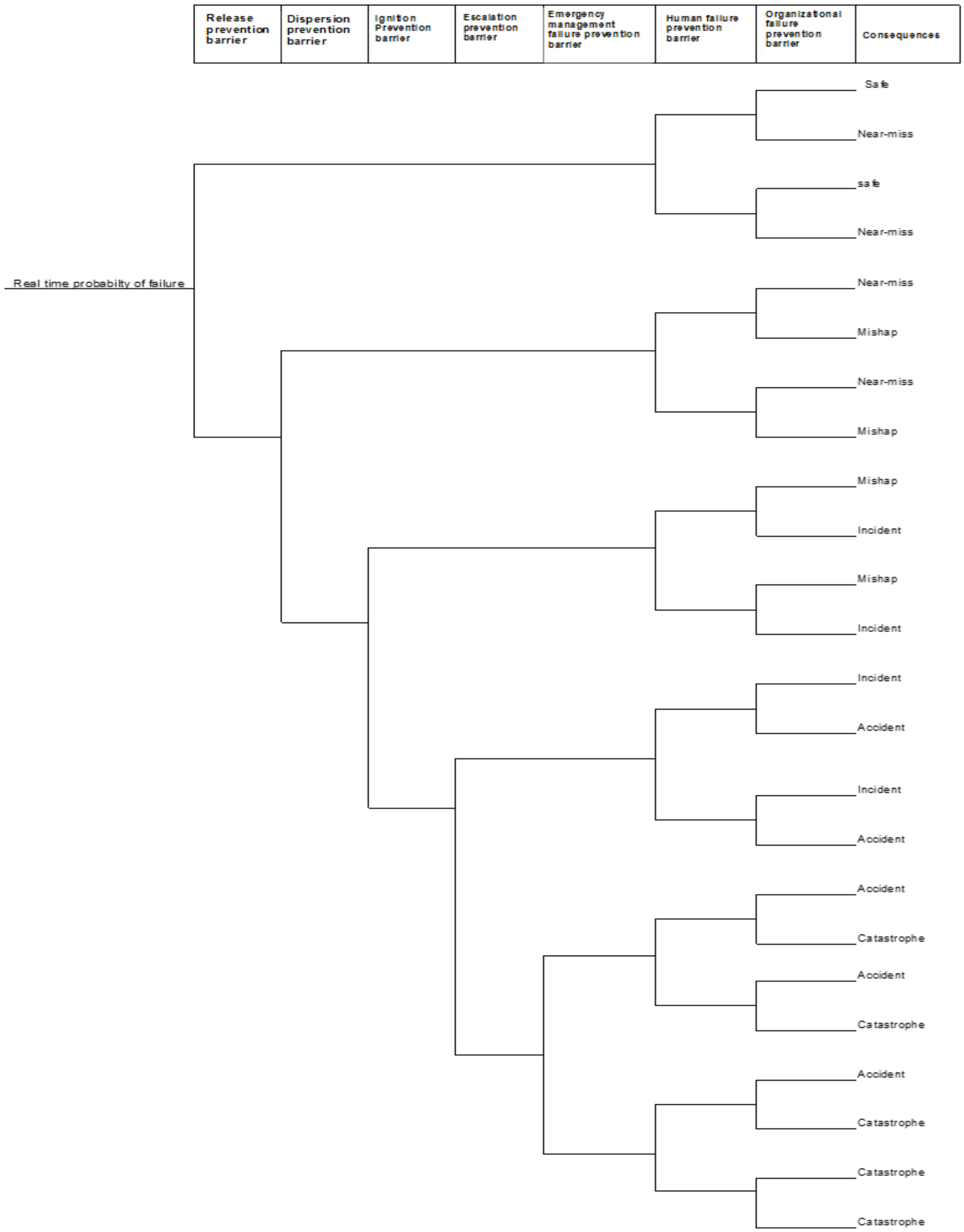



**Figure 6. 4. Revised Event tree for the accident modelling methodology (Adedigba et al. 2016).**

#### *6.4.3.3. Real time updating of occurrence probabilities of consequences*

In the last phase, the real time updating of consequence occurrence probability is executed. The real time probability of failure obtained in section 6.4.2 on a time interval basis are fed back into the process accident model frame work and subsequently, the real time occurrence probability of specific consequences is obtained.

## 6.5. Case study

### 6.5.1. Description of a crude distillation unit and related safety issues

The Crude distillation unit (CDU) is the principal fractionation unit and one of the most significant processes in the refinery. Among the various units in the refinery, the crude distillation unit is of primary concern because it defines the possible quantities of products directly (Al-Mayyahi et al. 2014; Yang & Barton 2015). The size of the refinery is measured by the capacity of the crude distillation unit (Wolf 2009). Hence, safe operation of the CDU is of paramount concern. The crude oil is processed in two units in most distillation plants. The first is the atmospheric distillation unit; it separates light hydrocarbons. The second is the vacuum distillation unit; it separates heavy hydrocarbons (Waheed & Oni 2015; Waheed et al. 2014) .

Figure 6.5 shows the schematic process flow diagram of a typical atmospheric distillation unit.

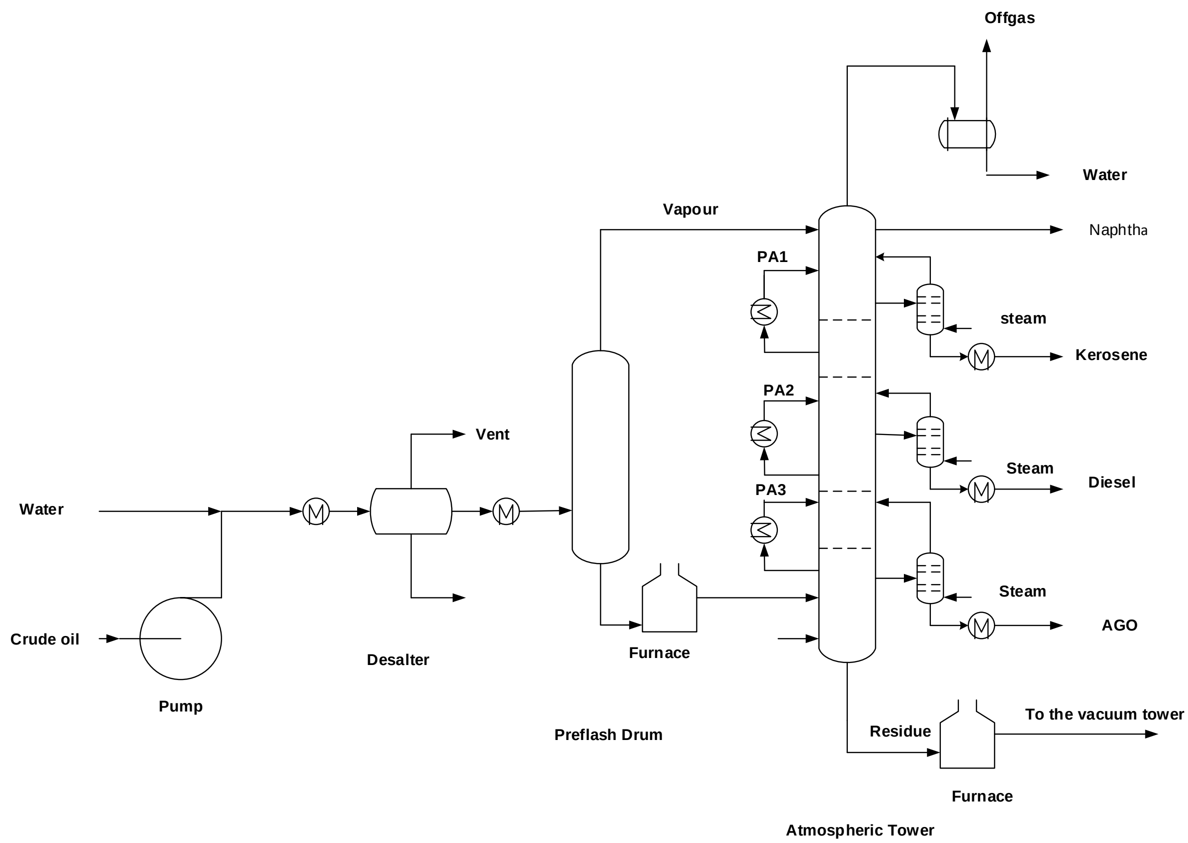



**Figure 6. 5. Process flow diagram of the crude distillation unit. (Al-Mayyahi et al. 2014).**

The crude oil is initially heated up by heat exchangers using the hot column products prior to their entering the desalting unit. The desalter unit installed in the preheat train reduces the crude salt content drastically by means of electrical desalination mechanism. The temperature of the crude is raised to 120 $^{0}$C and water is mixed with it prior to being routed to the Desalter. The temperature of the resulting crude oil is raised to 200 $^{0}$C. The function of the flash-vessel is to separate vapour content from the liquid content (Al-Mayyahi et al. 2014). The temperature of the resulting liquid is raised to range of 200-280 $^{0}$C by train of heat exchangers. The fire heater raised the temperature of the resulting liquid to about 400 $^{0}$C. Light hydrocarbon are separated as a result of abrupt change in the column volume.

Light vapours move to the column top while liquid hydrocarbons fall to the bottom of the column. Hydrocarbons fractions are withdrawn from the column based on the specific boiling temperatures. Naphtha exists as vapour and is subsequently condensed by the overhead condenser. Other hydrocarbon products are collected as side-streams (Al-Mayyahi et al. 2014).

The different hydrocarbon products are subsequently processed in downstream units to market requirement and the atmospheric residue is directed to the vacuum distillation unit, where separation occurs under a vacuum at lower temperatures into distinct cuts (Al-Mayyahi et al. 2014; Waheed & Oni 2015).

Quite a number of CDU failures have been reported in various accident databases. A good recent example is the CDU failure of the Chevron Richmond refinery that occurred on August 6, 2012. (CSB 2015) Safety of a CDU is critical and extremely sensitive because loss of control of a CDU can lead to devastating and cascading consequences beyond the plant boundaries. Therefore, it is crucial that refinery plants have a high level of safety and reliability. One of the challenges confronting the refinery industry today is the safety of its operations (Shaluf et al. 2003; Bertolini et al. 2009). The world has witnessed numerous accidents in refineries as a result of leakages, fire and explosion. The compendium of accident data in refineries from 1972 – 2011 with the associated insured losses in each case is given by Thomson (2013).

Since the real process data is not accessible, the CDU operation was simulated using Aspen HYSYS (8.8). The process variables including the feed temperature, feed pressure, condenser temperature and reboiler temperature were monitored and the data obtained were systematically analysed using the proposed methodology. The computational procedure of the methodology has been thoroughly explained in section 6.4.

## 6.6. Results and discussions

The simulated historical data of a CDU was projected into a PCA space. Table 6.3 shows percentage variance of the process variables; hence, the principal components among the process variables that explained the highest variation of the data were determined.

**Table 6.3. Principal Component Analysis of historical data.**

| No | Process Variables | % variance captured | % variance captured total |
|---|---|---|---|
| 1 | Feed Temperature | 53.30 | 53.30 |
| 2 | Pressure | 30.15 | 83.45 |
| 3 | Condenser Temperature | 11.36 | 94.81 |
| 4 | Reboiler Temperature | 5.19 | 100.00 |

The principal components are feed temperature and pressure. Table 6.4 shows the state of simulated historical process variables.

**Table 6. 4. The state of the historical process data.**

| Observation | Feed Temperature | Pressure | Condenser Temperature | Reboiler Temperature |
|---|---|---|---|---|
| 1 | Absent | Absent | Absent | Absent |
| 2 | Absent | Absent | Absent | Absent |
| 3 | Absent | Absent | Absent | Absent |
| 4 | Absent | Present | Present | Present |
| 5 | Present | Absent | Absent | Absent |
| 6 | Absent | Present | Absent | Absent |
| 7 | Present | Absent | Present | Absent |
| 8 | Absent | Present | Absent | Present |
| 9 | Absent | Absent | Absent | Absent |
| 10 | Absent | Absent | Absent | Absent |
| 11 | Absent | Absent | Absent | Absent |
| 12 | Absent | Absent | Present | Absent |
| 13 | Present | Present | Absent | Absent |
| 14 | Present | Absent | Absent | Absent |
| 15 | Absent | Absent | Absent | Absent |
| 16 | Absent | Absent | Absent | Absent |
| 17 | Absent | Absent | Absent | Absent |
| 18 | Absent | Absent | Absent | Present |
| 19 | Absent | Absent | Absent | Absent |
| 20 | Absent | Absent | Absent | Absent |
| 21 | Absent | Absent | Absent | Absent |
| 22 | Absent | Absent | Present | Absent |
| 23 | Absent | Present | Absent | Absent |
| 24 | Absent | Absent | Absent | Absent |
| 25 | Absent | Present | Absent | Absent |
| . | . | . | . | . |
| . | . | . | . | . |
| . | . | . | . | . |
| 1000 | Present | Present | Absent | Present |

*** Present signifies process variable set point is exceeded**

*** Absent indicates process variable set point is not exceeded**

The states of historical real time data were used to construct probabilistic networks using TAN Bayes algorithm. The Bayesian network structure (Figure 6. 6) generated from these data provides a concise depiction of probabilistic dependencies that is present among process variables.

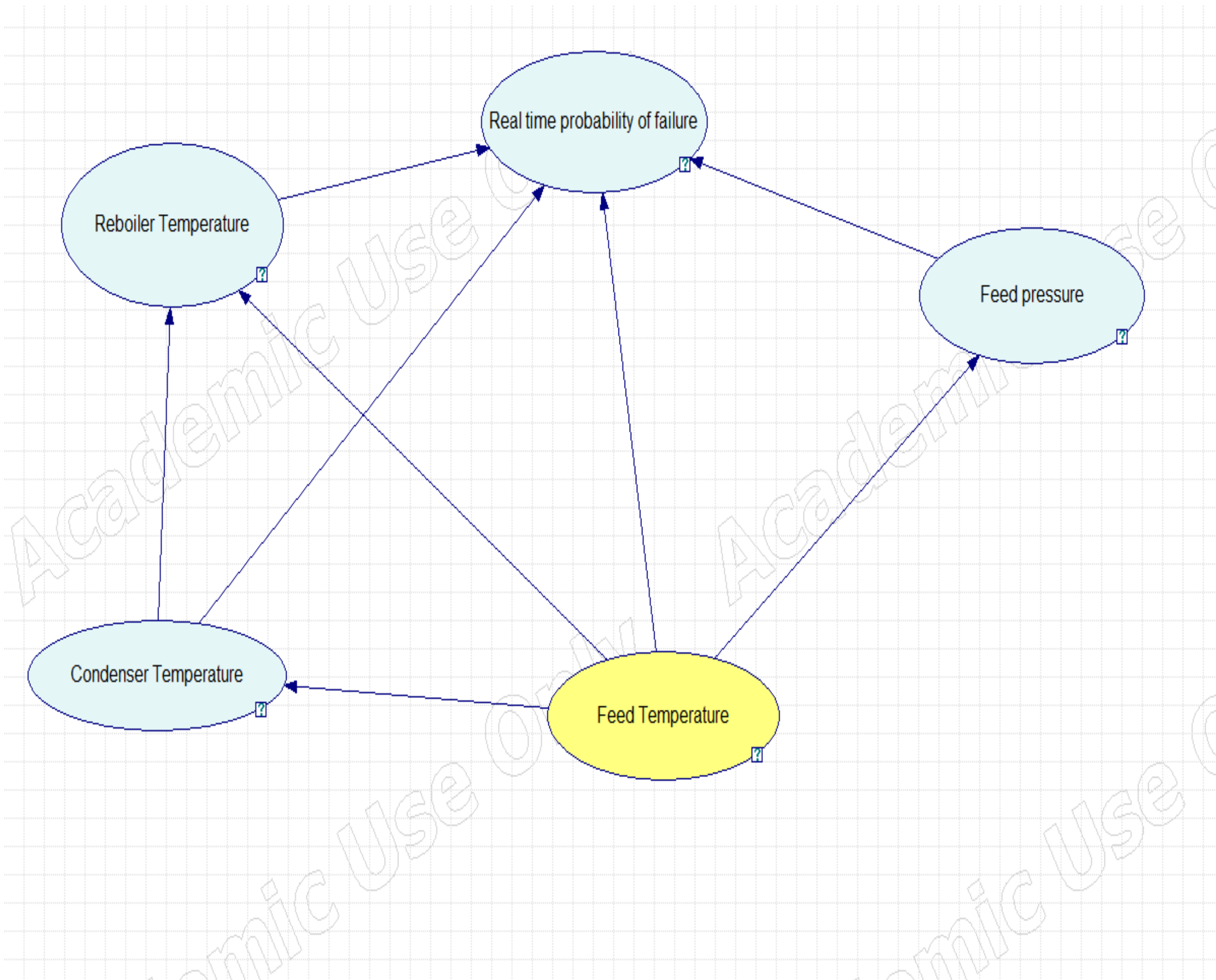



**Figure 6. 6. Probabilistic networks structure generated from historical data.**

The dynamic failure probability of the CDU was estimated using the generated BN in the following way. The probability of a fault, given that the temperature exceeded the set point, was evaluated and fed back into the Bayesian network. On every occasion pressure exceeded the set point, the evidence was set in the Bayesian network and subsequently, the probability of failure was predicted. This process was repeated to get a real time probability profile of CDU failure (shown in Table 6.5). This was used as the probability of an accident initiating event in the event tree (Figure 6.4)

**Table 6.5. Real time probability of failure.**

| **Time (minutes)** | **Real time probability failure** |
|---|---|
| 30 | 0.068 |
| 60 | 0.1333 |
| 90 | 0.028 |
| 120 | 0.039 |
| 150 | 0.022 |
| 180 | 0.049 |
| 210 | 0.012 |
| 240 | 0.019 |
| 270 | 0.018 |
| 300 | 0.01 |
| 330 | 0.015 |
| 360 | 0.044 |
| 390 | 0.014 |
| 420 | 0.028 |
| 450 | 0.057 |
| 480 | 0.061 |

The failure probabilities of safety (prevention) barriers by FT analysis of a CDU unit in previous work Adedigba et al. (2016) are given in Table 6. 6. The probability data of the basic events and associated fault trees for all safety (prevention) barriers are not presented in this paper, this is to decrease the number of tables and figures.

**Table 6. 6. Failure probabilities of prevention through fault tree analysis ( Adedigba et al. 2016).**

| Prevention barriers | Failure probability |
|---|---|
| RPB | 0.0842 |
| DPB | 0.0025 |
| IPB | 0.026 |
| EPB | 0.0286 |
| EMFPB | 0.0229 |
| HFPB | 0.00145 |
| OrFPB | 0.0069 |

The real time end state probability of the consequences for the proposed accident model is presented in Table 6. 7.

**Table 6. 7. Real time occurrence probability of the consequences.**

| Time( minutes) | Safe | Near miss | Mishap | Incident | Accident | Catastrophe |
|---|---|---|---|---|---|---|
| 30 | 6.18E-02 | 6.10E-03 | 5.33E-05 | 4.55E-07 | 1.28E-08 | 2.14E-12 |
| 60 | 1.21E-01 | 1.20E-02 | 1.04E-04 | 8.92E-07 | 2.51E-08 | 4.19E-12 |
| 90 | 2.55E-02 | 2.51E-03 | 2.19E-05 | 1.87E-07 | 5.28E-09 | 8.80E-13 |
| 120 | 3.55E-02 | 3.50E-03 | 3.05E-05 | 2.61E-07 | 7.35E-09 | 1.23E-12 |
| 150 | 2.00E-02 | 1.97E-03 | 1.72E-05 | 1.47E-07 | 4.15E-09 | 6.91E-13 |
| 180 | 4.46E-02 | 4.40E-03 | 3.84E-05 | 3.28E-07 | 9.24E-09 | 1.54E-12 |
| 210 | 1.09E-02 | 1.08E-03 | 9.40E-06 | 8.03E-08 | 2.26E-09 | 3.77E-13 |
| 240 | 1.73E-02 | 1.70E-03 | 1.49E-05 | 1.27E-07 | 3.58E-09 | 5.97E-13 |
| 270 | 1.64E-02 | 1.62E-03 | 1.41E-05 | 1.21E-07 | 3.39E-09 | 5.66E-13 |
| 300 | 9.09E-03 | 8.97E-04 | 7.83E-06 | 6.69E-08 | 1.89E-09 | 3.14E-13 |
| 330 | 1.36E-02 | 1.35E-03 | 1.17E-05 | 1.00E-07 | 2.83E-09 | 4.71E-13 |
| 360 | 4.00E-02 | 3.95E-03 | 3.45E-05 | 2.95E-07 | 8.30E-09 | 1.38E-12 |
| 390 | 1.27E-02 | 1.26E-03 | 1.10E-05 | 9.37E-08 | 2.64E-09 | 4.40E-13 |
| 420 | 2.55E-02 | 2.51E-03 | 2.19E-05 | 1.87E-07 | 5.28E-09 | 8.80E-13 |
| 450 | 5.18E-02 | 5.11E-03 | 4.46E-05 | 3.82E-07 | 1.07E-08 | 1.79E-12 |
| 480 | 5.55E-02 | 5.47E-03 | 4.78E-05 | 4.08E-07 | 1.15E-08 | 1.92E-12 |

It is worth observing that when the conditional dependencies and Bayesian network structure generated from a database are used to predict real time probability of failure, the significantly affects the real time probability of consequences, which highlights the significance of considering conditional dependency that is present among process variables in the historical database. The real time occurrence probability for accident and catastrophe is shown in Figure 6.7.

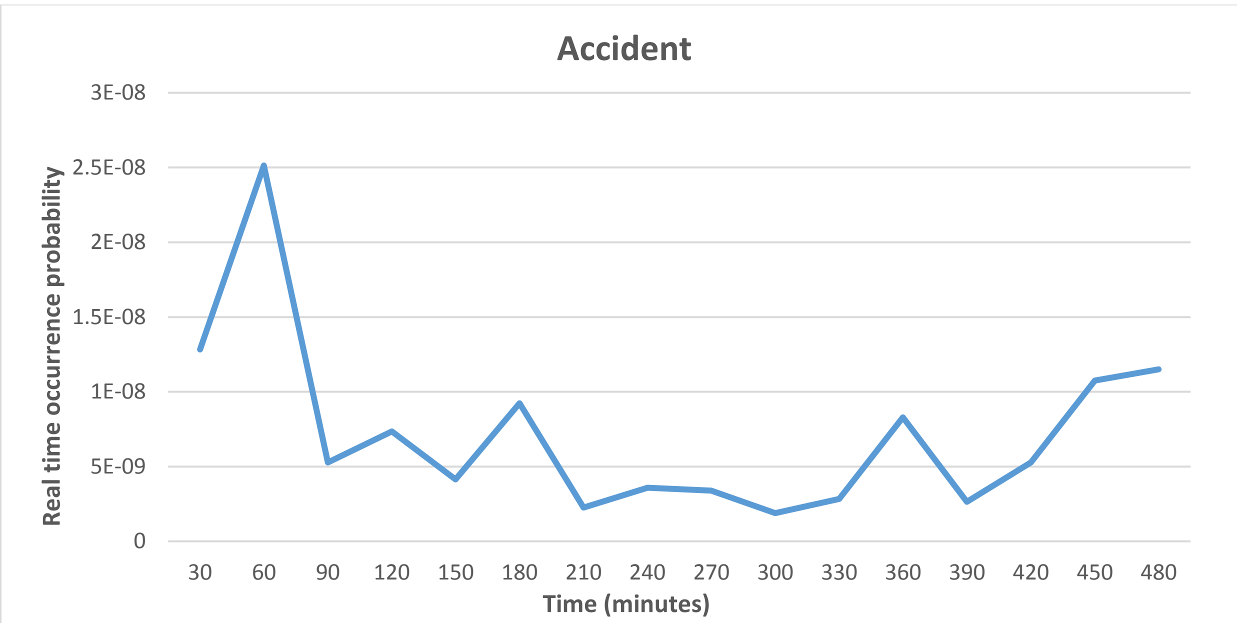



**Figure 6. 7a. Real time occurrence probability of Accident.**

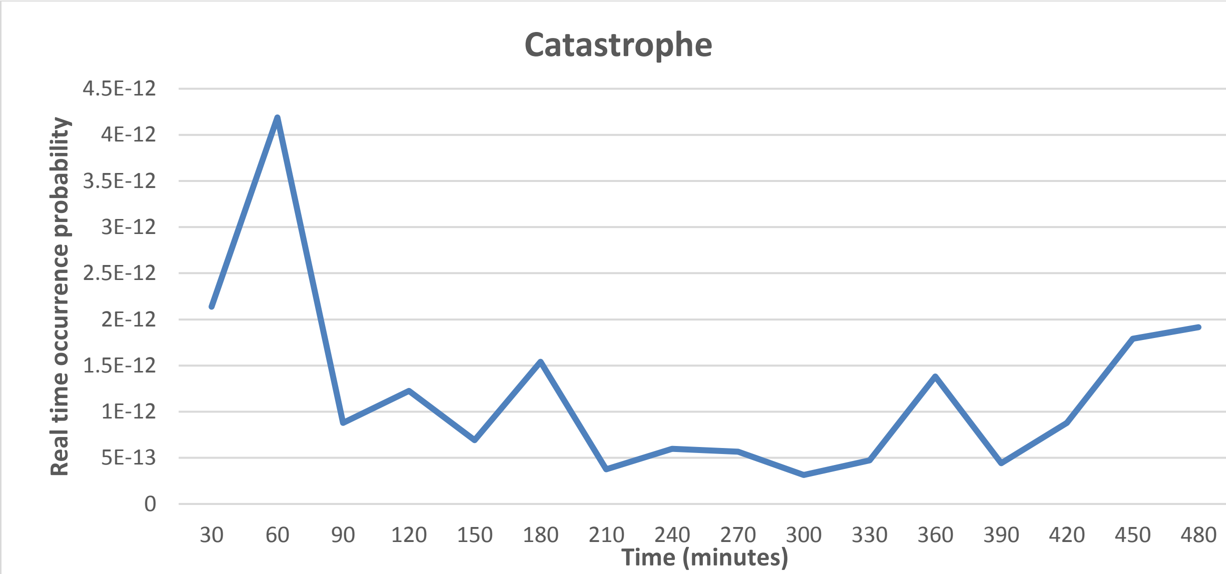



**Figure 6. 7b. Real time occurrence probability of Catastrophe.**

The results indicate that the proposed approach is capable to capture the dynamic effect of process variable deviations on the predicted occurrence probability of a process accident. This provide valuable information for operators to decide when and what process variables to control in order to lower the predicted accident probability to an acceptable level. For example, in this case study, the root causes of abnormal deviation of the feed temperature need to be identified and controlled to prevent the system failure leading to a process accident.

Application of this proposed accident model provides real time early warning and subsequent safety systems can be activated, before the progression of the fault increases the potential devastating impact on the safety of the systems: when occurrence probability of abnormal events in the process exceeds the acceptable threshold limit. This current study has illustrated the importance of BN in modelling the conditional dependencies and probabilistic networks structures that exist among process variables in the database for process accident modeling. The use of predicted time dependent occurrence probabilities of consequences in this methodology will help efficiently to take timely remedial action to avert accidents and guide to develop an effective safety process management plan. Consequently, safe process operation is ensured.

## 6. 7. Conclusions

The application of proposed the PCA-BN based process failure predictive model offers a technique for a predicting real time failure probability profile of a process system. The proposed model has the following strengths:

- The integration of PCA and BN enables the dynamic assessment of the failure probability of a process system completely based on historical and present process operational data.
- The use of PCA provides the capability to identify the key process variables that describe the most variance in process systems and utilize the process data in an efficient way.
- The model is capable of predicting and assessing the real time risk of a process unit by monitoring the deviation of its main variables.

This study has established the usefulness of BN and PCA in modelling the conditional dependencies and probabilistic network structure of historical process variables using the database. There are many potential prospects for using databases to assist constructing probabilistic networks. The probabilistic networks structure generated from databases will provide an exact depiction of probabilistic interaction and dependencies that is present among process variables. The usefulness of the proposed accident model is demonstrated on a simulated system. PCA in conjunction with the TAN algorithm offer an efficient approach to predicting real time probability of failure. Consequently, the real time dynamic risk profile computed based on the proposed process accident model renders valuable guidance in dynamic decision making for process safety management. Adequate process monitoring data are required to efficiently implement this proposed methodology; the model can quantitatively predict a real time dynamic risk profile that will help as a guide for dynamic decision making, before fault devastatingly increasing the potential impact on the safety of the systems. This present study has analyzed accidents and their associated consequences in real time. This provides valuable evidence to support decision making during process safety management. To further improve the model, the following future work has been planned:

(1) Validate the proposed model through real process data; and

(2) Establish the link between process variations and potential loss.

# Appendix

Orthogonality is a feature of the raw variables and it clearly indicates that the raw data (variables) are perpendicular. The zero correlation (uncorrelated) is a feature of the centered variables, which evidently indicates that they (centered variables) are perpendicular. The concept of orthogonal and uncorrelated variables can be vividly explained as follows.

Suppose Y and Z are vector observations of the process variables Y and Z. Algebraically, the following equations are used to describe orthogonal and uncorrelated relationship between Y and Z.

(1) Orthogonal relationship exists between Y and Z if and only if $Y'Z = 0$

(2) Y and Z process variables are said to be uncorrelated if and only if

$$(Y - \bar{Y}1)'(Z - \bar{Z}1) = 0$$

The means of Y and Z are denoted as $\bar{Y}$ and $\bar{Z}$ respectively, while 1 represents vector of ones (Rodgers & Nicewander 1984).

**Orthogonal validation of the case study**

The data matrix for the case study is given below. Matrix Y denotes the scores for Feed temperature and Matrix Z denotes the scores for pressure. It can be observed that $Y'Z = 0$ and $(Y - \bar{Y}1)'(Z - \bar{Z}1) = 0$ . Therefore, the two principal components (Feed temperature and pressure) are orthogonal and uncorrelated.

$$Y = \begin{bmatrix} 5.258532577 \\ -4.940070201 \\ -5.560353458 \\ 26.831214230 \\ -4.241180076 \\ -9.736927382 \\ -0,639439626 \\ -3.950696339 \\ -0.763496277 \\ -0.680791843 \\ 35.05855357 \\ -0.639439626 \\ -5.732939237 \\ -3.534094821 \\ -4.199827859 \\ -1.466483968 \\ -2.707050480 \\ 26.955270880 \\ -1.052961797 \\ -4.774661333 \\ 0.601126886 \\ -2.545739176 \\ -3.534094821 \\ -3.534094821 \\ -13.75117178 \\ -16.71918322 \end{bmatrix} \qquad Z = \begin{bmatrix} -12.54354100 \\ 1.516253280 \\ 2.765049280 \\ 6.520100420 \\ -3.902332200 \\ 11.17360900 \\ -7.142065700 \\ -0.534890400 \\ -6.892306500 \\ -7.058812600 \\ 12.27598730 \\ -7.142065700 \\ -0.958291900 \\ -1.314351000 \\ -3.985585200 \\ -5.477004300 \\ -2.979412300 \\ 6.2703412200 \\ -6.309535000 \\ 1.183241010 \\ -9.639657700 \\ -7.315707900 \\ -1.314351000 \\ -1.314351000 \\ 19.19608640 \\ 24.92359390 \end{bmatrix}$$

## Acknowledgement

The authors thankfully acknowledge the financial support provided by, Natural Science and Engineering Council of Canada and Canada Research Chair (CRC) Program.

## References

Adedigba, S.A., Khan, F. & Yang, M., 2016. Chemical Engineering Research and Design Dynamic safety analysis of process systems using nonlinear and non-sequential accident model. *Chemical Engineering Research and Design*, 111, pp.169–183.

Al-Mayyahi, M. a., Hoadley, A.F. a. & Rangaiah, G.P., 2014. Energy optimization of crude oil distillation using different designs of pre-flash drums. *Applied Thermal Engineering*, 73(1), pp.1204–1210.

Baksh, A. et al., 2015. Network based approach for predictive accident modelling. *Safety Science*, 80, pp.274–287.

BenaditP, J. & FrancisF, S., 2015. ScienceDirect Improving the Performance of a Proxy Cache Using Tree Augmented Naive Bayes Classifier. *Procedia - Procedia Computer Science*, 46, pp.184–193.

Bertolini, M. et al., 2009. Development of Risk-Based Inspection and Maintenance procedures for an oil refinery. *Journal of Loss Prevention in the Process Industries*, 22, pp.244–253.

Bobbio, a. et al., 2001. Improving the analysis of dependable systems by mapping Fault Trees into Bayesian Networks. *Reliability Engineering and System Safety*, 71, pp.249–260.

Cerquides, J. & Lopez De Antaras, R., 2003. Tractable Bayesian Learning of Tree Augmented Naive Bayes Models. In *Proceeding of the Twentieth International Conference on Machine Learning*. Washington, D.C.

Cooper, G.F. & Herskovits, E., 1991. A Bayesian method for constructing Bayesian belief networks from databases. , pp.86–94.

CSB, 2015. *Final Investigation Report: Cheveron Richmond Refinery Pipe rupture and Fire*, Washington, D.C.

Dash, D. & Druzdzel, M., 2002. Robust independence testing for constraint-based learning of causal structure. *Proceedings of the Nineteenth conference on …*, pp.167–174.

Dash, D. & Druzdzel, M.J., 1999. A Hybrid Anytime Algorithm for the Construction of Causal Models From Sparse Data. *Artificial Intelligence*, pp.142–149.

Fei, Z. & Liu, K., 2016. Online process monitoring for complex systems with dynamic weighted principal component analysis. *Chinese Journal of Chemical Engineering*, 24(6), pp.775–786.

Friedman, N., Geiger, D. & Goldszmidt, M., 1997. Bayesian Network Classifiers. *Machine* *learning*, 29, pp.131–163.

He, T. et al., 2006. Process fault detection and diagnosis based on principal component analysis. *Fifth International Conference on Machine Learning and Cybernetics*, 1(August), pp.3551–3556.

Jensen, F.. & Nielsen, T.., 2007. *Bayesian Networks and Decision Graphs* 2 edition., Springer, New York.

Jiang, L. et al., 2012. Improving Tree augmented Naive Bayes for class probability estimation. *Knowledge-Based Systems*, 26, pp.239–245.

Jim Thomson, 2013. June 1999. 2. An accident at the coking plant of the Anacortes refinery, Puget Sound, Washington State, USA, on 25. *2013*, pp.1–22.

Joe Qin, S., 2003. Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 17(8-9), pp.480–502.

Khakzad, N., Khan, F. & Amyotte, P., 2013. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection*, 91(1-2), pp.46–53.

Khakzad, N., Khan, F. & Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety*, 96(8), pp.925–932.

Khan, F., Rathnayaka, S. & Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection*, 98, pp.116–147..

Khan, F.I. & Abbasi, S.., 1998. Techniques and methodologies for risk analysis in chemical process industries. *Journal of Loss Prevention in the Process Industries*, 11(4), pp.261–277.

Kresta, J. V., Macgregor, J.F. & Marlin, T.E., 1991. Multivariate statistical monitoring of process operating performance. *The Canadian Journal of Chemical Engineering*, 69(1), pp.35–47.

Li, N., Feng, X. & Jimenez, R., 2017. Predicting rock burst hazard with incomplete data using Bayesian networks. *Tunnelling and Underground Space Technology*, 61, pp.61–70.

Li, W. & Qin, S.J., 2001. Consistent dynamic PCA based on errors-in-variables subspace identification. *Journal of Process Control*, 11(6), pp.661–678.

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

Mohammad, M., 2006. *Risk Analysis in Engineering: Techniques, Tools and Trend*, CRC.

Neapolitan, R.E., 2004. *Learning Bayesian networks*, Pearson Prentice Hall.

Nývlt, O. & Rausand, M., 2012. Dependencies in event trees analyzed by Petri nets. *Reliability Engineering and System Safety*, 104, pp.45–57.

Pan, Y. et al., 2016. Fault detection with improved principal component pursuit method. *Chemometrics and Intelligent Laboratory Systems*, 157, pp.111–119.

Pearl, J., 1998. *Probabilistic Reasoning in Intelligent Systems*, Morgan Kaufmannn, San Francisco, California.

Rajakarunakaran, S., Kumar, A.M. & Prabhu, V.A., 2015. Journal of Loss Prevention in the Process Industries Applications of fuzzy faulty tree analysis and expert elicitation for evaluation of risks in LPG refuelling station. *Journal of Loss Prevention in the Process Industries*, 33, pp.109–123.

Rathnayaka, S., Khan, F. & Amyotte, P., 2010. SHIPP methodology : Predictive accident modeling approach . Part II . Validation with case study. *Process Safety and Environmental Protection*, 89(2), pp.75–88.

Rathnayaka, S., Khan, F. & Amyotte, P., 2011. SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection*, 89(3), pp.151–164.

Rodgers, J. & Nicewander, W., 1984. Linearly independent, orthogonal, and uncorrelated variables. *The American*.

Shaluf, I.M., Ahmadun, F.R. & Said, A.M., 2003. Fire incident at a refinery in West Malaysia: The causes and lessons learned. *Journal of Loss Prevention in the Process Industries,*

Shlens, J., 2014. A Tutorial on Principal Component Analysis. *ArXiv*, pp.1–13.

Spirtes, P., Glymour, C. & Scheines, R., 1993. Causation, Prediction, and Search. In pp. 272–273.

Venkatasubramanian, V. et al., 2003. A review of process fault detection and diagnosis: Part III: Process history based methods. *Computers & chemical …*, 27(3), pp.327–346

Waheed, M. a. et al., 2014. Thermoeconomic and environmental assessment of a crude oil distillation unit of a Nigerian refinery. *Applied Thermal Engineering*, 66(1-2), pp.191–205.

Waheed, M. a. & Oni, a. O., 2015. Performance improvement of a crude oil distillation unit. *Applied Thermal Engineering*, 75, pp.315–324.

Wang, H., Khan, F. & Ahmed, S., 2015. Design of Scenario-Based Early Warning System for Process Operations. *Industrial & Engineering Chemistry Research*, 54(33), pp.8255–8265.

Wang, Y., Ma, X. & Joyce, M.J., 2016. Reducing sensor complexity for monitoring wind turbine performance using principal component analysis. *Renewable Energy*, 97, pp.444–456.

Wise, B.M. & Gallagher, N.B., 1996. The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6(6), pp.329–348.

Wolf, F.G., 2009. Operationalizing and Testing Normal Accident Theory in Petrochemical Plants and Refineries. *Production and Operations Management*, 10(3), pp.292–305.

Yang, Y. & Barton, P.I., 2015. Refinery Optimization Integrated with a Nonlinear Crude

Distillation Unit Model. *IFAC-PapersOnLine*, 48(8), pp.205–210.

Yu, H. et al., 2014. Self-Organizing Map Based Fault Diagnosis Technique for Non-Gaussian Processes. *Industrial & Engineering Chemistry Research*, 53(21), pp.8831–8843.

Yu, H., Khan, F. & Garaniya, V., 2015a. Modified Independent Component Analysis and Bayesian Network-Based Two-Stage Fault Diagnosis of Process Operations. *Industrial & Engineering Chemistry Research*, 54(10), pp.2724–2742..

Yu, H., Khan, F. & Garaniya, V., 2015b. Nonlinear Gaussian Belief Network based fault diagnosis for industrial processes. *Journal of Process Control*, 35, pp.178–200.

Yu, H., Khan, F. & Garaniya, V., 2015c. Risk-based fault detection using Self-Organizing Map. *Reliability Engineering & System Safety*, 139, pp.82–96.

Zadakbar, O., Imtiaz, S. & Khan, F., 2012. Dynamic Risk Assessment and Fault Detection Using Principal Component Analysis. *Industrial & Engineering Chemistry Research*, 52, pp.809–816.

Zhang, Y., Li, S. & Teng, Y., 2012. Dynamic processes monitoring using recursive kernel principal component analysis. *Chemical Engineering Science*, 72, pp.78–86.

# Chapter 7

## 7.0 Dynamic Failure Analysis of Process Systems Using Neural Network

### Preface

*A version of this chapter has been published in the* ***Journal of Process Safety and Environmental Protection 2017; 111: 529-543.*** *I am the primary author. Co-author Faisal Khan provided fundamental understanding, assisted in developing the conceptual model and subsequently translated this to the numerical model. Co-author Ming Yang provided much needed support in implementing the concept and testing the model. I carried out most of the data collection and analysis. I prepared the first draft of the manuscript and subsequently revised the manuscript, based on the feedback from co-authors and also a peer review process. The two co-authors assisted in developing the concept and testing the model, reviewed and corrected the model and results. They also contributed to the review and revision of the manuscript.*

### Abstract

Complex and non-linear relationships exist among process variables in a process operation. Owing to these complex and non-linear relationships potential accident modelling using an analytical technique is proving to be not very effective. The artificial neural network (ANN) is a powerful computational tool that assists in modelling complex and nonlinear relationships. This relationship has good potential to be generalized and used for subsequent failure analysis.

This paper integrates ANNs with probabilistic analysis to model a process accident. A Multi-layer perceptron (MLP) is used to define the relationship among process variables. The defined relationship is used to model a process accident considering logical and casual dependence of the variables. The predicted accident probability is subsequently used to estimate the likelihoods of

failure to the process unit. A backward propagation technique is used to dynamically update the variable states and the failure probabilities accordingly.

Integrating ANN with a probabilistic approach provides an efficient and effective way to estimate process accident probability as a function of time and thus the risk can be easily predicted upon quantifying the damage. The updating mechanism of the approach makes the model adaptive and captures evolving process conditions. The proposed integrated approach is applied to the Tennessee process system as a case study.

**Keywords:** Artificial neural network (ANN) analysis; Sequential accident model; Accident prediction; Reliability Analysis; System Safety.

## 7.1. Introduction

In contemporary decades, the complexity and advancement in modern process systems is rapidly increasing. The complexity in the systems being built is directly associated with the number of process variables the system comprises (Yu et al. 2015; Adedigba et al. 2017). This sophistication presents substantial risk of failure. Due to this development, it is of paramount importance to study these systems thoroughly to create a failure forecasting mechanism and to provide early warnings to keep the operations of these systems safe (Zhong et al. 2016). Recently, risk assessment techniques and application in chemical process industries have metamorphosed into dynamic risk analysis (Villa et al. 2016). Dynamic risk techniques present a framework that explicitly captures the impact of time and chemical process dynamics for all scenarios (Labeau et al. 2000). Khan et al. (2016) defined dynamic risk assessment "as a method that updates estimated risk of a deteriorating process according to the performance of the control system, safety barriers, inspection and maintenance activities, the human factors, and procedure" Different techniques have been applied in dynamic risk assessment of chemical process industries. Detail of these technique can be found in (Al-shanini et al. 2014; Khan & Abbasi 1998; Khan et al. 2016; Khan et al. 2015; Meel & Seider 2006; Villa et al. 2016). However, most of the approaches applied in dynamic risks assessment of chemical operations are analytical models which are less effective due to complex and non-linear relationship that exist among process variables ( Adedigba et al. 2016).

Process accident models are frameworks that express the relationship between causes and effects of accidents. Process accident models provide well detailed explanations about how and why accidents happen and they are adopted as tools for process risk assessment (Qureshi 2008; Adedigba et al. 2016). Different types of process accident models have been developed, details of these are given by (Al-shanini et al. 2014; Attwood et al. 2006; Qureshi 2007; Rathnayaka et al. 2011). The strengths and weakness of several dynamic risk assessment techniques developed over the years are presented in ( Khan, et al. 2016). Recently, Adedigba et al. (2016) presented a dynamic safety analysis methodology that "modelled dependency relationships among accident contributory factors within prevention or safety barriers". Afterward permit non sequential failure of the barriers. However, this methodology still suffers a major limitation: it does not account for dependencies among process variables. Process accidents usually occur due to failure of events induced by failure of physical components and abnormalities of process variables (Tan et al. 2013; Adedigba et al. 2016).

Quite recently, artificial neural networks (ANNs) have found wide application as a new computational technique in various fields of studies due to the remarkable characteristics possessed by ANNs. These features are: noise tolerance, high parallelism, nonlinearity and learning and generalization ability (Basheer & Hajmeer 2000; Azizi et al. 2016; Chitsazan et al. 2015).

Artificial neural networks are powerful computational tools for modelling complex and nonlinear relationships that lack mathematical models (Azizi et al. 2016; Ashtiani & Shahsavari 2016). ANNs have the strong ability to map the probability distribution even when the training data is small (Świetlicka et al. 2017).

The ANN model has been used for real time process monitoring of a polymerization plant (Gonzaga et al. 2009). The ANN technique has been applied to learn the correlation between molecular and electrochemical properties (Chen et al. 2016). The ANN has been used to predict the void fraction for a gas-liquid flow (Azizi et al. 2016).The ANN model has been used to monitor deformation behaviour of a metal alloy (Ashtiani & Shahsavari 2016) and the ANN model has been found to be the best to predict the degradation of total petroleum hydrocarbon (TPH) (Sanusi et al. 2016).

Artificial neural network (ANN) regression models have been widely applied in various process control applications for detecting and controlling nonlinear dynamic systems that might not be easily detected and control by conventional controllers. The performance of the controller with respect to nonlinear system dynamics has been modelled with an artificial neural network (Lee et al. 1992). The recurrent neural network is applied in the development of multi-step ahead prediction model for the nonlinear process plant. The model developed proved to be extremely accurate (Zhang & Morris 1995). Neural networks have been applied for the adaptive control of nonlinear process systems. The nonlinear model developed using neural networks is used to calculate the parameters of the adaptive controllers (Yu & Annaswamy 1997). ANNs have been used to give a more precise process model of a vitrification Process based on nonlinear system characteristic. Subsequently, the residual of the model are systematically monitored to identify signs of imminent vessel failure of the vessel used in the vittrificaition process (Lennox & Montague 1997). Table 1 presents the differences between the current work and the following ANNs articles.

**Table 7.1. The differences between the current work and the following ANNs articles.**

| Articles | ANN Input | ANN Output | Process control | Fault diagnosis | Controller design | Probability prediction | Fault tree analysis | Event tree analysis | Risk | Area of Application |
|---|---|---|---|---|---|---|---|---|---|---|
| Lennox & Montague (1997) | Power supplied to the induction coil and the level of waste in the melter vessel | Temperature of the melter | The model developed can be used to control vitrification process | It can be used for fault detection and diagnosis | The PLC controller is design by the nonlinear process model. | It does not predict probability of process deviation based non linear relationship among key process variables | No fault tree analysis | No event tree analysis | No risk prediction | Process control and Automation |
| Zhang & Morris ( 1995) | Process data y(k) is used | To predict y(k+1) | Multiple step ahead prediction & control based on nonlinear system dynamic | It can be used for fault detection and diagnosis | The multiple step ahead prediction are used to design process controller | It does not predict probability of process deviation based non linear relationship among key process variables | No fault tree analysis | No event tree analysis | No risk prediction | Process control and Automation |
| Yu & Annaswamy ( 1997) | System data | Corrections in the controller parameters | Adaptive control of nonlinear system dynamic | It can be used for fault detection and diagnosis | The controller structure is determined by the nonlinear model. | It does not predict probability of process deviation based non linear relationship among key process variables | No fault tree analysis | No event tree analysis | No risk prediction | Process control and Automation |
| Lee et al.( 1992) | System data | Reference input to the controller | Control of nonlinear system dynamic | It can be used for fault detection and diagnosis | The model developed provides effective structure for neural controllers | It does not predict probability of process deviation based non linear relationship among key process variables | No fault tree analysis | No event tree analysis | No risk prediction | Process control and Automation |
| Adedigba et al. (2016) | ANN is not applied | ANN is not applied | Automatic safety control can be activated , when the predicted probability exceed normal range | It can be used for fault detection and diagnosis | It has no valuable information to design process controller | it does not predict probability of process deviation based non linear relationship among key process variable | It uses fault tree to model linear and nonlinear interactions within safety barriers and predict failure probability of safety barriers based on this relationships | It permit both sequential and non sequential failure of safety barriers | It offer opportunity for quantifying risk when the damages are quantified | Process risk Assessments |
| Current work | Cumulative probabilities of key process variables | Probability of system deviation based on nonlinear relationship among process variables | Automatic safety control can be activated , when the predicted probability exceed normal range | It can be used for fault detection and diagnosis | It has no valuable information to design process controller | It predict probability of process accidents based on nonlinear interaction of process variables | It uses fault tree to model linear interactions within safety barriers and predict failure probability of safety barriers based on the linear relationships | It permit sequential failure of safety barriers | It offer opportunity for quantifying risk when the damages are quantified | Process risk Assessments |

Complex and non-linear interaction relationships exist among process variables, so that a fault in a particular component can be hidden by this complex variation and rapidly propagate to multiple upsets. If this multiple upset is not addressed, it could result in devastating consequences that will significantly affect the safety of the entire system (Yu et al., 2014). A vast volume of data is being generated in complex modern industrial systems without appropriate physical models that can be analysed to interpret the data generated for failure analysis and decision making (Xu & Hou, 2009). Due to this development, the authors proposed the application of ANNs for dynamic failure assessment of complex and non-linear relationships that exist among process variables in chemical process operations.

The main contribution of this work is to present a data driven dynamic failure assessment methodology which overcomes the weakness in dynamic safety analysis presented by Adedigba et al. (2016). The probability of process deviation will be predicted from the nonlinear relationship that exists among the real time process monitoring data. This work integrates an ANN- data driven model with a process accident model to form a hybrid model that can be used for fault diagnoses, failure assessments, risk assessment and decision making.

This paper is organized as follows. Section 7.2 presents a brief description of ANNs and its application in different disciplines. The proposed methodology is presented in Section 7.3. In Section 7.4, the Tennessee Eastman Chemical process is used as a case study to demonstrate the application of the proposed methodology. Finally, Section 7. 6 provides the conclusion.

## 7.2. Artificial Neural Network (ANNs)

Artificial neural networks are computational techniques that have a robust capability to detect nonlinear relationships among input and output data without the need for a detailed understanding of the physical systems. ANNs have been widely used as a technique for modelling and forecasting in various disciplines (Chitsazan et al. 2015). The ANN structure consists of tightly interconnected artificial neurons that have a strong capability to execute wide ranging parallel calculations for data processing and adequate knowledge representation. They have been widely used as computational tools in various disciplines because they are more effective than other computational techniques. ANNs learn from example and subsequently capture the functional relationships among the data when the original relationships are unknown or are extremely difficult to describe (Ashtiani & Shahsavari 2016; Zhang et al. 1998).

ANNs give superior prediction accuracy with respect to statistical regression techniques. They have been applied in modelling intricate real world problems in different categories: Clustering, pattern classification, functional approximation, forecasting, optimization, association and control (Basheer & Hajmeer 2000). Different ANN models have been proposed; the most widely used is the multi-layer perceptron (MLP) (Świetlicka et al. 2017; Zhang et al. 1998). An example is shown below to illustrate how ANN can be used to predict the real-time state of a process variable. Consider a simplified flow system of a tank as shown in Figure 7.1. The flow rate of the liquid into the tank is A $m^3/s$ and the flow rate of the liquid out of the tank is B $m^3/s$. The primary interest is to determine the liquid level (C) in the tank at any instant. Mathematically, the liquid level (C) can be expressed as:

$$C\frac{\partial_V}{\partial_t} = A - B \qquad 7.\ 1$$

Where, V is the total volume of the tank.

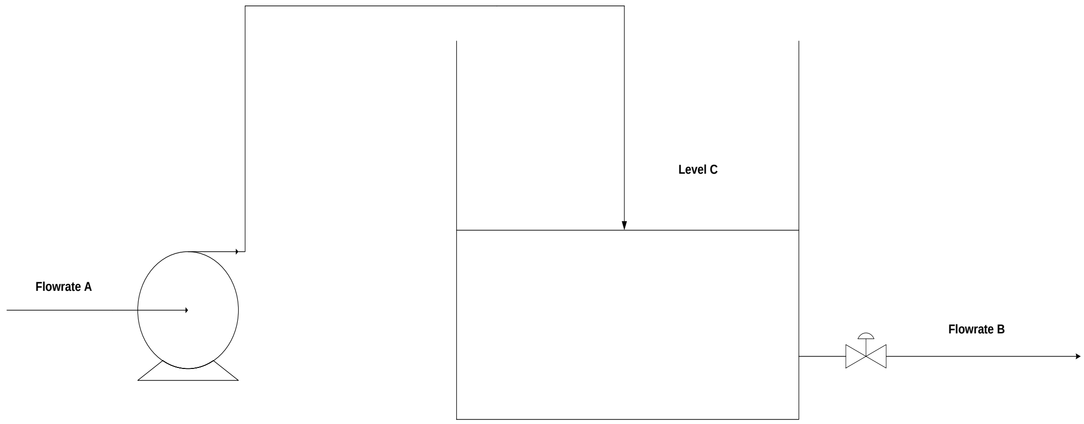



**Figure 7. 1. Simplified flow system of a tank.**

ANN can be used to predict the liquid level (C) at any instant. The following steps summarize the procedure involve in using ANN to predict the liquid level (C) (Gardner & Dorling 1998). The detail calculation for each step is given in the appendix.

1. Initialisation of the networks weights: The values of weights and biases used in this example is given by Table 7. 2. These values of initial weights and biases are assumed.

**Table 7.2. The values of weights, input and biases used in the example demonstrated.**

| Input Values | | Target value (T) | Assumed weights | | | | | | Biases | |
|---|---|---|---|---|---|---|---|---|---|---|
| A($i_1$) | B($i_2$) | | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ | $w_6$ | $b_1$ | $b_2$ |
| 10.0000 | 8.0000 | 2.000 | 0.7000 | 0.6000 | 0.5000 | 0.1000 | 2.2500 | 0.4000 | 0.7500 | 0.6500 |

2. Developing ANN architecture: The ANN architecture of the flow system of a tank is given by figure 7. 2

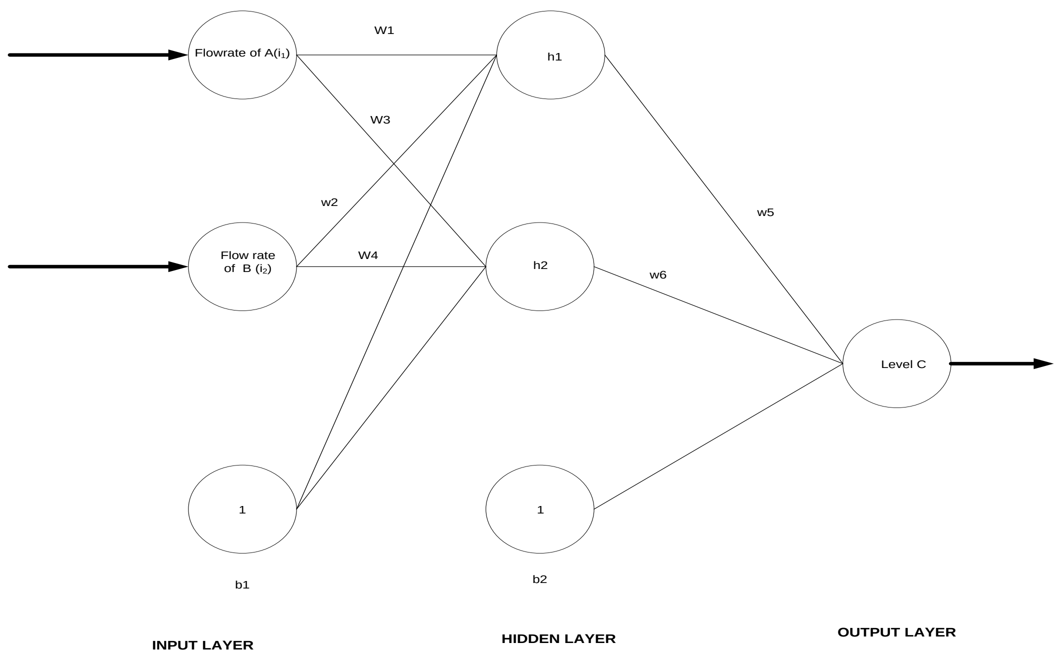



**Figure 7. 2 – Schematic diagram of ANN architecture of flow system of a tank.**

3. Present a part of training data as an input vector to the network and specify the target values: The input and target data for this example is presented by Table 7.2.
4. Calculating the actual output by propagating the input vector through networks (Forward pass): Comprehensive procedure of this step is given in the appendix
5. Computing the error term by finding the difference between the target values and the output values: Detail of this step given in the appendix
6. Propagate the error term back to the network( Backward pass): Comprehensive procedure of this step is given in the appendix
7. Change (update) the weight to decrease overall error term: The updated weight is given by Table 7. 3

**Table 7.3. The updated weights, of the example demonstrated.**

| **Updated weights** | | | | | |
|---|---|---|---|---|---|
| $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ | $w_6$ |
| 0.7000 | 0.6000 | 0.5001 | 0.1001 | 2.2677 | 0.4177 |

8. The entire steps in this procedure is repeated cyclically with all the training input until the overall error term is reasonably small.

### 7.2.1 Multi-layer Perceptron (MLP) using Backpropagation algorithm

The multilayer perceptron is a feedforward neural network. It is comprised of three distinct principal layers: an input layer, one hidden layer or series of hidden layer, and lastly an output layers (Chen et al. 2017; Chitsazan et al. 2015; Azizi et al. 2016). MLP are usually regarded as universal approximators, it has capability to approximate all forms of arbitrary functions to highest degree of accuracy. Function approximation (modelling) in ANN are normally applied to complex real world problems in situation where no definite theoretical model exist (Azizi et al. 2016).

Interconnectivity within a layer do not exist and every neuron in each layer are totally joined to the next layer neurons (Riedmiller 1994).

The input variables are received by the input layer to the network and transfers it through weighted connection to the first neuron in the hidden layer. The first neuron in the hidden layer computes their respective activations and serially passes them to the next neurons respectively. The hidden layer uses activation function for transforming the network input variables to output variables (dependent variables). Activation function are usually bounded and continuous nonlinear functions. Typical examples of activation function are logarithm-sigmoid (logsig) and hyperbolic-tangent-sigmoid (Tansig) (Chitsazan et al. 2015; Riedmiller 1994). The hidden layer do not relate with the external environment and it empowers MLP with the capability to handle nonlinear classification problems which a simple perceptron cannot handle (Basheer & Hajmeer 2000). The output function in the output layer is normally a linear function that sums the input signals of the output layer (Chitsazan et al. 2015; Chen et al. 2016).

The prediction in feedforward ANN is represented mathematically by equations 7.2 and 7.3 (Chitsazan et al. 2015).

$$\mathrm{Q_{jk}} = \mathrm{f_1}\left(\mathrm{b_j} + \sum_{\mathrm{i}} \mathrm{W_{ij}}\ \mathrm{I_{ik}}\right) \tag{7.2}$$

$$\mathrm{Q_k} = \mathrm{b} + \sum_{\mathrm{j}} \mathrm{W_j}\ \mathrm{Q_{jk}} \tag{7.3}$$

Where $f_1$ is the hidden layer activation function, $b_j$ bias for the hidden layer, $b$ is the bias for the output layer, $I_{ik}$ is the $ith$ input for the kth input vector, $Q_{jk}$ is the hidden layer output of the $jth$ node, $W_{ij}$ and $W_j$ are interconnection weight between the layers and $Q_k$ is the $kth$ predicted probability of failure (Chitsazan et al. 2015). Multi-layer perceptron with backpropagation is used to predict the liquid level (C) in the example considered in section 7. 2.

### 7.2.2. Training algorithm

Normally MLP training is supervised learning, where the target values for the input variables are known. (Chen et al. 2017; Ashtiani & Shahsavari 2016). Neural network training is a systematic unrestrained nonlinear minimization process where the arc weight of the networks are repeatedly modified through iteration process to reduce the overall squared error between the target values and the actual output values for every output nodes over every input variable vectors (Chitsazan et al. 2015; Chen et al. 2016; Zhang et al. 1998). The knowledge that is learnt by the neural network is preserved in both the arcs and nodes as arc weight and node biases (Chen et al. 2016; Zhang et al. 1998).

The widely used training algorithm for MLP is the backpropagation algorithm (BP).The backpropagation algorithm essentially uses gradient steepest method to find global minimum error surface. This algorithm computes the local gradient of the error surface and subsequently update the weight along the direction of the steepest local gradient. This permit the weight to converge to the global error of the minimum surface (Chen et al. 2016; Gardner & Dorling 1998).

Each iteration in BP is made of two sweeps: forward and backward sweeps. Forward sweep produce the output and backward sweep propagation calculated the error term to adjust the weight. Both forward and Backward sweep are done repeatedly until the ANN output is the same with the Target value within permissible pre-determined tolerance level (Basheer & Hajmeer 2000). Several performance functions such as ,sum square error (SSE), mean squared error (MSE) , and mean absolute deviation (MAD) are used to determine the weight that decrease the overall error measure (Zhang et al. 1998). To minimise the overall squared error between the target values and output values from ANN model, the MSE performance function is widely used. The equation for MSE performance function is given by equation 7.4.

$$\mathrm{MSE} = \frac{1}{n}\sum_{m=1}^{n}(Y_T - Y_{pred})^2 \qquad (7.4)$$

Back propagation algorithm is very versatile and efficient and it can be generally be used for pattern recognition, data modelling, data and image compression, control, classification and forecasting. The training the example considered in section 7.2 is a supervised training, because the target value for the input variables is known. Detailed computation for both forward and backward passes are given in the appendix.

## 7.3. Proposed Methodology

The flow chart for the proposed methodology is given in Figure 7. 3

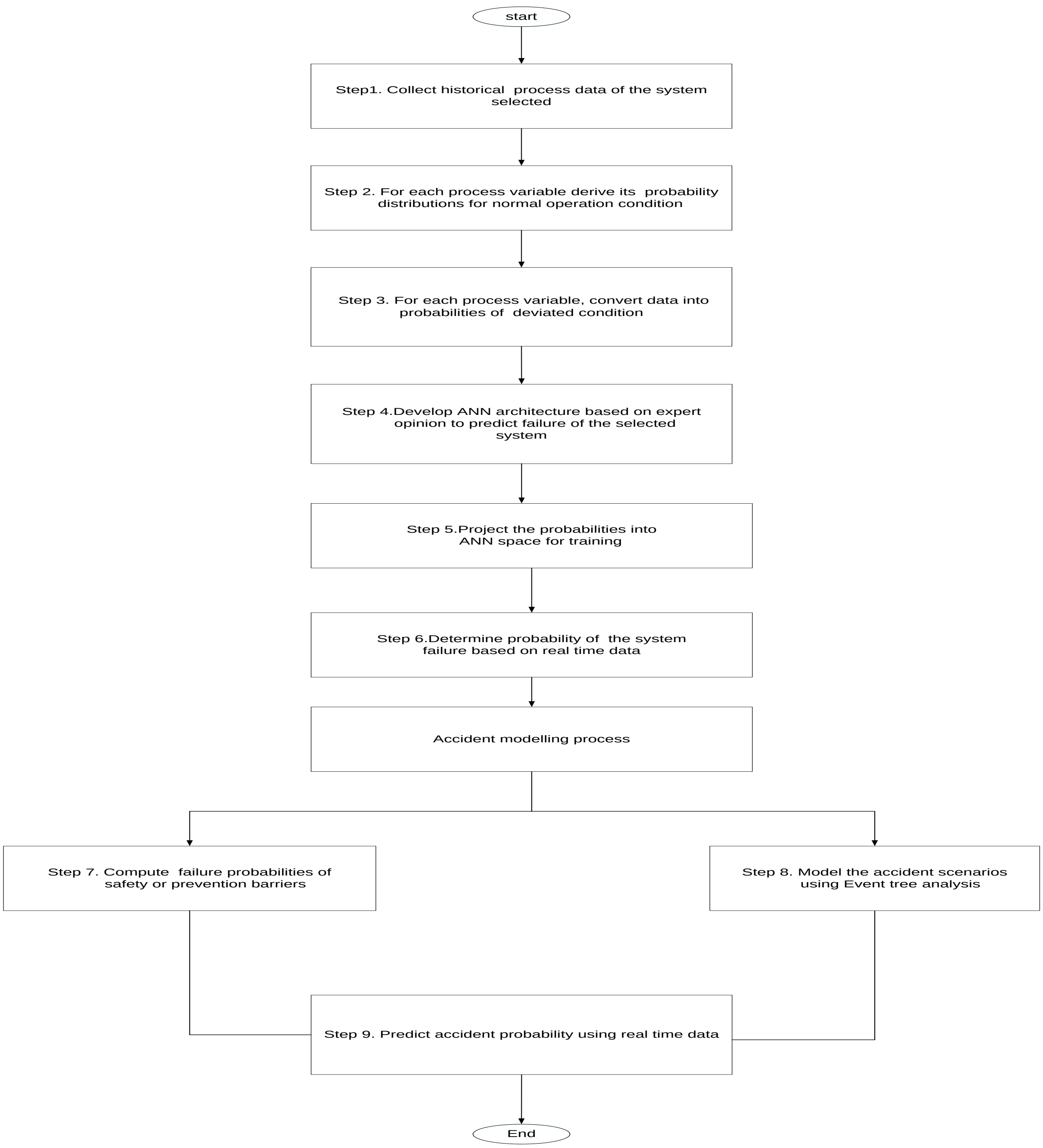



**Figure 7. 3. Proposed accident modelling methodology flowchart.**

### 7.3.1. Historical data collection and determination of probability distribution

The historical data for the plant are acquired or simulated with an appropriate software. An appropriate probability distribution is derived from the process data collected or simulated. The procedure for deriving probability distribution from data is not discussed in this work. The historical process data used in this work perfectly fit into normal distribution. The probability distribution curve of the historical data is generated using the normal distribution. The cumulative probability F (t) of each value of the process variables to it set point (normal operating condition) in the historical data is estimated from the cumulative distribution curve after the set point for each variable has been fixed. This process normalised the input data within the range of 0-1. This method of input normalisation is refer to as external normalisation (Zhang et al. 1998). Equation 7.5 is used for estimating the F (t) from CDF of the all process variables. Cumulative probability estimated is the probability of deviation from normal operation.

$$F(t) = \varphi\left(\frac{t-\mu}{\sigma}\right) \tag{7.5}$$

Where t is the value of the monitor process variable, μ is the mean the set point, $\sigma$ is the standard deviation.

### 7.3.2. ANN model and architecture

The neural network model used for prediction in this study is the three layer feedforward backpropagation neural networks (FNNs). The ANN model consist of three layers: Input, hidden and output layers respectively. The Levenberg- Marquardt optimization algorithm is used for training alongside with cross validation that applied the principle of early stopping to prevent over fitting (Gonzaga et al. 2009). Levenberg- Marquardt optimization algorithm is selected because is most efficient and gives fast convergence and stability in the course of training (Azizi et al. 2016).

The sigmoidal function used as activation function is hyperbolic tangent function, while the performance function selected is the Mean square error (MSE). Table 7.4 present the main features of ANN architecture model developed for the case study

**Table: 7.4. Main features of ANN model developed.**

| ANN model | Parameters |
|---|---|
| **Network architecture** | Feed- forward back propagation |
| **Input data** | Cumulative probability of deviation |
| **Number of hidden neurons** | 10 |
| **Output data** | Probability of process deviation |
| **Training algorithm** | Levenberg- Marquardt |
| **Training function** | Tangsig |
| **performance function** | MSE |

### 7.3.3. Learning and training using Multi-layer perceptron (MLP)

Training is a repetitive iterations process of estimating the arcs weights of artificial neural network. The arc weight are the major elements of artificial neural network. Training in an MLP is a supervised training. A supervised training is a training that has a target value for all input vectors (Ashtiani & Shahsavari 2016; Chen et al. 2017; Zhang et al. 1998).The training input data is the cumulative probabilities values of each process variable to its set point in the historical data that is estimated from the cumulative distribution curve. The target values are a set of acceptable probability of deviation for normal process operation of the plant under consideration.

The total input data is subdivided into training set, test set, validation set. The arc weight is computed using the training set and the test set is used to determine the generalization capability of the network. The validation set is used to prevent overfitting or to determine when the training process should end (Chen et al. 2016; Zhang et al. 1998; Basheer & Hajmeer 2000). 70% of input data is used for training, 15% of input data is used for testing and 15% of input data is used for validation. Both the forward and backward sweeps are carried out repetitively until the ANN output is the same with the target value within initially specify tolerance level. The criteria used to end training is the coefficient of determination ($R^2$). The coefficient of determination ($R^2$) is predicted by the equation 6.

$$\boldsymbol{R^2} = \mathbf{1} - \left\{ \frac{\sum_j^n (Y_j^{Actual} - Y_j^{Predicted})^2}{\sum_j^n (Y_j^{Actual} - \bar{Y}_j^{Actual})^2} \right\} \qquad (\mathbf{7.6})$$

Where $R^2$ is the coefficient of determination, $Y_j^{Actual}$ and $Y_j^{Predicted}$ are the actual and predicted values respectively, $\bar{Y}_j^{Actual}$ is the mean actual value, n is the number of the actual value; the ordinal is represent by subscript j (Chen et al. 2017).

### 7.3.4. Generalisation using real time process monitoring data

ANNs have strong capability to generalize after learning from the sample presented to the network. The real time process monitoring data (current data) is collected from the plant and procedure in 3.1 is repeated on the real time process data. The cumulative probabilities obtained are subsequently used as the new input to the artificial neural network to predict the probabilities of deviation of the plant from normal operation.

### 7.3.5. Hazard identification and analysis

Earlier to the development of process accident model, there is a need for detail hazard evaluation studies. The hazard evaluation studies are performed based on the accessible process information of the plant. Such accessible process information include: flow sheet sketches, data sheet and procedure and piping and instrumentation diagrams (Rathnayaka et al. 2012). Hazard and operability Analysis (HAZOP) is a famously used qualitative hazard identification technique to detect and evaluate equipment failure that could result into accidents and including identification of operability problem (Khan et al. 2015). There are numerous methods that can be used to perform hazard evaluation studies, comprehensive review of these methods including their advantage and disadvantage is given by (Khan & Abbasi 1998). Process accident scenarios are subsequently developed based on hazard evaluation studies performed. Each of process accident scenarios consist of two principal components: initiating event (accident instigating sequence) and consequences (Final effect of accident sequence). The primary reason of scenarios development is to identify principal relevant barriers to be placed along the accident sequences to avert or alleviate the consequences of the accident since accident occurs because of failure of relevant prevention or safety barriers (Rathnayaka et al. 2012). A comprehensive hazard analysis of a process system has been done in previous works. Seven principal prevention barriers has been determined based on hazard analysis. The prevention or safety barrier are define accordingly by (Adedigba et al. 2016) namely: Release prevention barrier (RPB), Dispersion prevention barrier (DPB), Ignition prevention barrier (IPB), Escalation prevention barrier (EPB), Emergency management failure prevention barrier (EMFPB), Human factor prevention barrier (HFPB) and Organization failure prevention barrier (OrFPB). Comprehensive explanation of the barriers and how they are methodically set alongside accident path can be found in (Adedigba et al. 2016).

#### *7.3.5.1. Assessment of failure probability of prevention barriers.*

Fault tree analysis has been widely used as technique for safety analysis and quantitative reliability of process systems. It give a structural (graphical) representation of several combination of basic failure that could result into the occurrence of top event (undesirable top event) (Tan et al. 2013)

Fault tree is a deductive technique for recognising ways in which hazards can combine to cause an accident. The methodology begins with top event (accident instigating event) and systematically works backward toward different scenario that can lead to the accident(Crowl & Louvar 2001). Fault tree analysis could be performed using two basic techniques: qualitative techniques and quantitative techniques. In qualitative technique logical expression is used to derive minimal cut set for the top event while in quantitative technique failure probability of basic event are assigned and they are used to estimate failure probability of top event. Commonly used logic gates in fault trees are AND and OR logics (Tan et al. 2013). Fault tree analysis is used to assess the failure probabilities of the prevention barriers.

The failure probabilities of prevention barriers used in this work is given in Table 7.5. To lessen the number of tables and figures the failure probability of basic event and corresponding fault trees are not shown in this work.

**Table 7.5. Failure probabilities of prevention through fault tree analysis (Adedigba et al. 2016).**

| Prevention barriers | Failure probability |
|---|---|
| RPB | 0.0842 |
| DPB | 0.0025 |
| IPB | 0.0260 |
| EPB | 0.0286 |
| EMFPB | 0.0229 |
| HFPB | 0.0015 |
| OrFPB | 0.0069 |

### *7.3.5. 2. Event tree construction*

Event tree starts with a major accident initiating event and systematically work forwardly towards final consequence. Event tree methodology is inductive in nature. It provides a comprehensive information of how failure of the system occur and subsequently evaluate the occurrence probability of the consequence (Crowl & Louvar 2001). Therefore, and event tree denote a structural logic arrangement of several events that might result from a particular initiating event. The initiating event is methodologically branch into success or failure to propagate event consequence in various branches of the event tree. The path of various branches in the event tree eventually lead to a possible outcome (consequence). Analysis in the event tree can be performed qualitatively and quantitatively. Qualitative analysis merely finds the probable consequence of initiating event, while the quantitative analysis estimate the occurrence probability of consequences (Ferdous et al. 2009).

The occurrence probabilities of consequences $P(C_k)$ in the event tree is estimated by equation 7.

$$\mathbf{P(C_K)} = \prod_{\mathbf{j\,\epsilon\,SB_k}} \mathbf{x_i^{\theta_{i,k}}\,(1 - x_i)^{1-\theta_{i,k}}} \qquad (\mathbf{7.7})$$

Where $SB_k$ signifies prevention barrier associated to level k; and $\theta_{i,k} = 1$ when level k failure moves through the failure branch of prevention barrier i; $\theta_{i,k} = 0$ when level k failure moves through the success branch of prevention barrier i. $x_i$ is the failure probabilities of prevention barriers (Adedigba et al. 2016; Rathnayaka et al. 2010). The event tree for the proposed accident modelling methodology is given in Figure 7. 4.

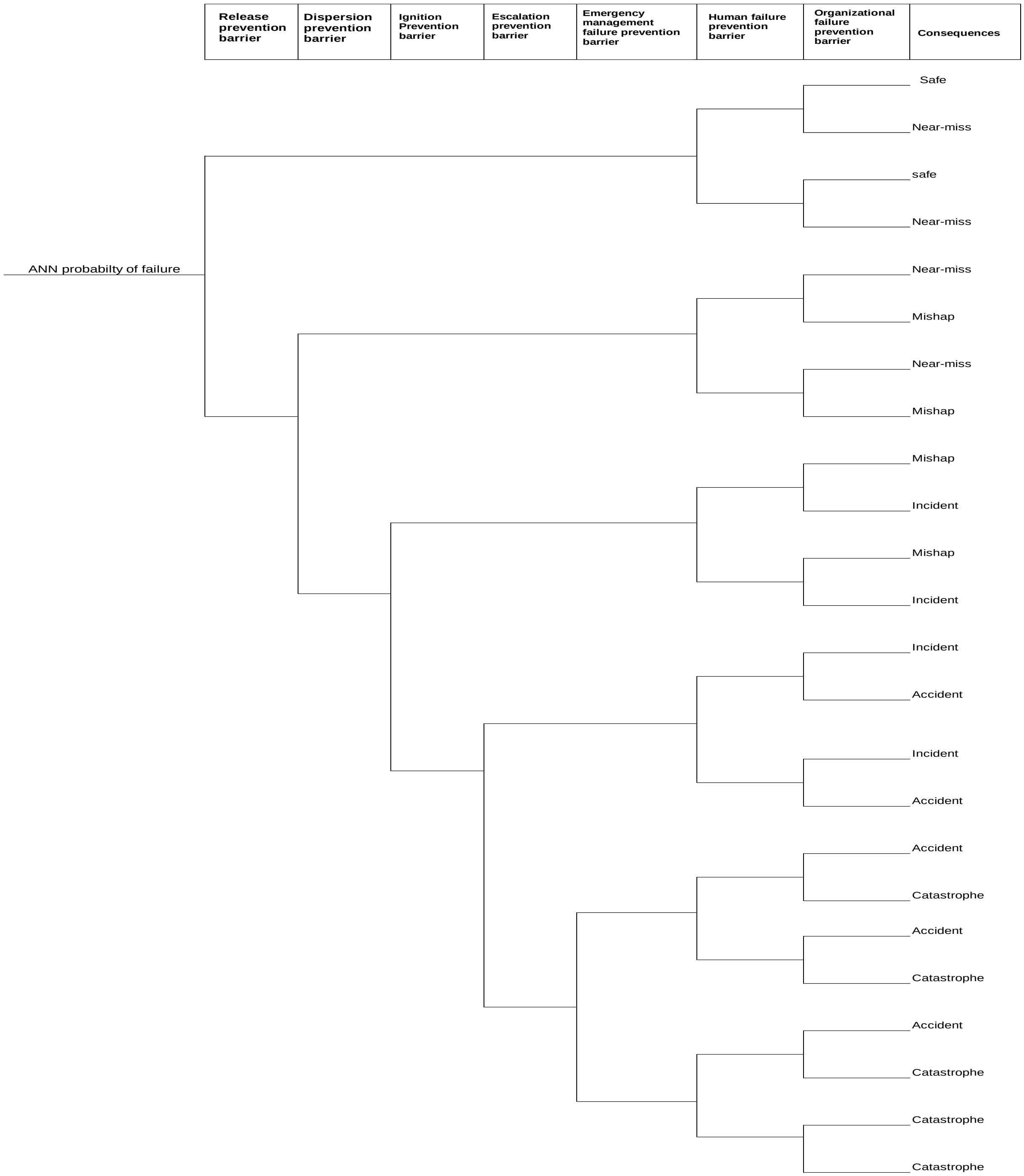



**Figure 7. 4. Event tree for the proposed accident modelling methodology.**

## 7.4. Case study: Tennessee Eastman Chemical process

The proposed methodology is applied to the Tennessee Eastman Chemical Process by monitoring the process variables in the process. There are five principal operating units in the process flow diagram: reactor, product condenser, vapour vapour–liquid separator, a recycle compressor and the product stripper (Downs & Vogel 1993; Yu et al. 2014). Figure 7.5 shows the process flow diagram of the plant.

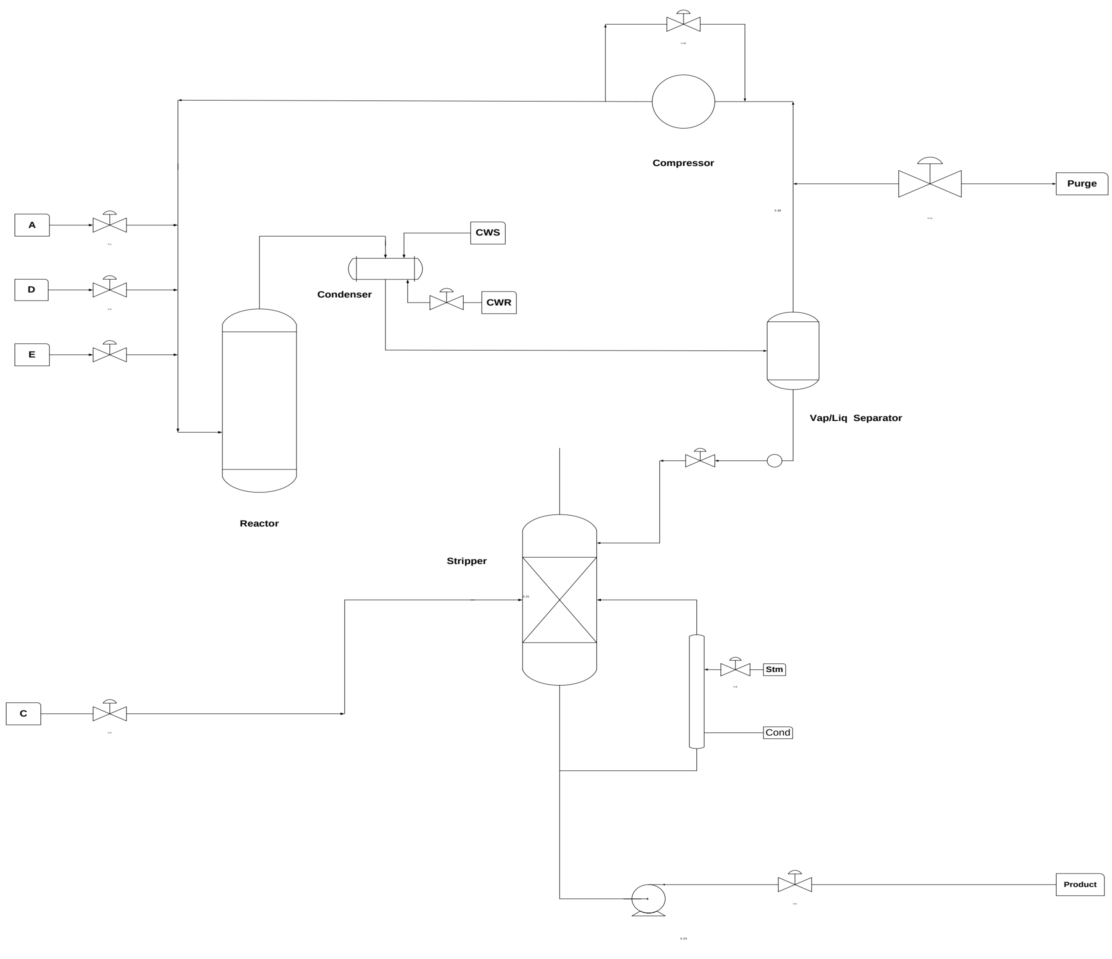



**Figure 7 .5. Flowchart of the Tennessee Eastern chemical process problem. (source (J. J. Downs & Vogel 1993).**

All the reactants (feeds) which are gaseous in nature flow to the catalytic reactor where they react to produce a liquid product. The liquid products leave the reactor in the form of vapour; the condenser condenses the resultant vapour.

The products from the condenser flow to the vapour-liquid separator where the products are separated into condensed and non - condensed products. (Yu et al. 2014). The non-condensed components at this point are recycled to the reactor with a centrifugal compressor. Additionally, both inert products and by-products are purged as vapour from the system in the vapour –liquid separator. Finally, the condensed components pass through the strippers where they are stripped with stream 4 to eliminate the residual reactants. Further refinement of the products takes place in the downstream section (Downs & Vogel 1993; Yu et al. 2014).

The simulation program applies a decentralized control approach to build a closed loop simulation of the Tennessee Eastman chemical process. 20 fault conditions have been deliberately programmed into the Tennessee Eastman chemical process simulation for the purpose of generating data for process monitoring. A detailed description of how the process dysfunction is created and manipulated for the purpose of generating process monitoring data for the Tennessee Eastman chemical process is provided by (J. J. Downs & Vogel 1993).

Process data about the plant is obtained from the website of the Chemical Engineering Department, at the Massachusetts Institute of Technology: http://web.mit.edu/braatzgroup/links.html .Due to copy right issues, the data is not provided here. Interested readers can access the data on the university website. There are 22 process variables in the Tennessee Eastman Chemical process. All 22 process variables have been classified into two distinct classes: 5 key state variables and 17 manipulate variables by (Khan et al. 2016). The five key state variables data are used for the prediction in this work. Table 7. 6 describes key state process variables with their unit. The proposed methodology is applied to the case study.

**Table 7.6. Key state variables of the Tennessee Eastern chemical process (Khan et al. 2016).**

| Variable no. | Process variable | Set point | unit |
|---|---|---|---|
| X7 | Reactor pressure | 2705 | KPa |
| X8 | Reactor level | 75 | % |
| X9 | Reactor Temperature | 120.4 | $^0$C |
| X12 | separator level | 50 | % |
| X15 | Stripper base level | 50 | % |

The computational procedure of ANN using the backpropagation algorithm has been demonstrated with a simplified example. Comprehensive computation is given in the appendix. The computation technique used in this case study is given in the appendix. In the case study, the training input data is the cumulative probability value for each process variable. The target values are sets of acceptable probability of deviation for normal operation given by experts. The software used for the ANN training and prediction is the Mat lab R2014b program.

## 7.5. Result and discussion

The complex non-linear relationship that exists among process variables in a chemical process operation justifies the application of the ANN model for predicting the probability of deviations. The ANN-based failure prediction technique is proposed to eliminate or reduce subjectivity, using a Bayesian network structure to represent dependency among process variables. ANN empirically gives precise and concise probabilistic dependencies that exist among process variables. The nonlinear correlation structure among process variables is used to predict the probability of deviation and subsequently the likelihoods of failure. Integration of the ANN with a quantitative failure assessment technique enhances accuracy and quantitative power and at the same time reduces uncertainty drastically. The developed model facilitates the dynamic failure quantification of a process operation. The model provides a robust optimization technique for proactively assessing and managing failure in the chemical process operation, when the relationships among process variables are uncertain or ambiguous. The proposed ANN-based failure prediction technique is fast, computationally more efficient and it can be generalized compared with other analytical accident modelling techniques that have been developed. With this developed technique, process dynamics can be satisfactorily monitored at any instant. Due to the dynamic nature of the model, the process operation dynamic is effectively captured and any process upset that can cause a serious safety issue is quickly recognised before affecting the system with consequences. The model exhibit unique advantage over other analytical accident modelling techniques, in that it make use of a vast volume of process monitoring data generated for failure prediction.

The model has strong a capability of diagnosing the related process variables that are responsible for the deviation. The model prediction is able to adequately account for non-linearity and process operational uncertainty. The primary objective of developing an ANN model for a process operation is to have a tool at hand that permits fast and reliable prediction of the probability of process deviation from unlearned process monitoring data, after the model has being trained with historical data.

The input data to the ANN network are cumulative probabilities of 5 key process variables obtained by following the procedure in section 7. 3. The target values are a set of acceptable probabilities of deviation of the process under Consideration. Table 7.7 presents the probabilities of deviations predicted from the ANN model developed.

**Table 7.7. Predicted probabilities of deviation from ANN model developed.**

| ANN Output No | Predicted probability of failure |
|---|---|
| 1 | 2.35E-03 |
| 2 | 2.10E-3 |
| 3 | 1.64E-03 |
| 4 | 1.15E-03 |
| 5 | 6.61E-04 |
| 6 | 2.18E-4 |
| 7 | 1.37E-04 |
| 8 | 3.68-04 |
| 9 | 4.47E-04 |
| 10 | 3.43E-04 |
| 11 | 9.51E-06 |
| 12 | 6.05E-04 |
| 13 | 4.25E-03 |
| 14 | 3.50E-03 |
| 15 | 1.28E-03 |
| 16 | 1.70E-04 |

The predicted probabilities in Table 7.7 are real time probabilities of deviation predicted from the training of unlearned process monitoring data in the ANN model developed, after the model has being trained with process historical data.

The values of Table 7.7 were used as the input to the event tree of Figure7. 4. Table 7.8 provides the occurrence probability of the consequences. The predicted probability of the consequences in Table 7. 8 are real time probabilities predicted from the proposed accident model. Application of the proposed model provides real time probabilities and subsequently real time prompt warning is also provided. When the real time probability predicted surpasses the tolerable threshold limit, safety systems can be promptly activated before a process upset (fault) causes devastating consequences that will significantly affect the safety of the process plant.

One major setback of backpropagation algorithm is overfitting. The ANN model in this case (overfitting) presents a well fitted model for training data set only, however it fails to predict the

**Table 7.8. Occurrence probability of the consequences.**

| ANN model output number | Safe | Near miss | Mishap | Incident | Accident | Catastrophe |
|---|---|---|---|---|---|---|
| 1 | 2.14E-03 | 2.11E-04 | 1.84E-06 | 1.57E-08 | 4.43E-10 | 7.39E-14 |
| 2 | 1.91E-03 | 1.88E-04 | 1.64E-06 | 1.41E-08 | 3.96E-10 | 6.60E-14 |
| 3 | 1.49E-03 | 1.47E-04 | 1.28E-06 | 1.10E-08 | 3.09E-10 | 5.15E-14 |
| 4 | 1.05E-03 | 1.03E-04 | 9.01E-07 | 7.70E-09 | 2.17E-10 | 3.61E-14 |
| 5 | 6.00E-04 | 5.92E-05 | 5.17E-07 | 4.42E-09 | 1.24E-10 | 2.07E-14 |
| 6 | 1.98E-04 | 1.96E-05 | 1.71E-07 | 1.46E-09 | 4.11E-11 | 6.85E-15 |
| 7 | 1.25E-04 | 1.23E-05 | 1.07E-07 | 9.17E-10 | 2.58E-11 | 4.31E-15 |
| 8 | 3.35E-04 | 3.30E-05 | 2.88E-07 | 2.46E-09 | 6.94E-11 | 1.16E-14 |
| 9 | 4.07E-04 | 4.01E-05 | 3.50E-07 | 2.99E-09 | 8.43E-11 | 1.40E-14 |
| 10 | 3.12E-04 | 3.08E-05 | 2.69E-07 | 2.30E-09 | 6.47E-11 | 1.08E-14 |
| 11 | 8.65E-06 | 8.53E-07 | 7.45E-09 | 6.37E-11 | 1.79E-12 | 2.99E-16 |
| 12 | 5.50E-04 | 5.43E-05 | 4.74E-07 | 4.05E-09 | 1.14E-10 | 1.90E-14 |
| 13 | 3.87E-03 | 3.81E-04 | 3.33E-06 | 2.85E-08 | 8.01E-10 | 1.34E-13 |
| 14 | 3.18E-03 | 3.14E-04 | 2.74E-06 | 2.34E-08 | 6.60E-10 | 1.10E-13 |
| 15 | 1.16E-03 | 1.15E-04 | 1.00E-06 | 8.57E-09 | 2.41E-10 | 4.02E-14 |
| 16 | 1.55E-04 | 1.53E-05 | 1.33E-07 | 1.14E-09 | 3.21E-11 | 5.34E-15 |

main feature of the entire data set. This set back of over fitting generally reduces generalization capability of the network (Azizi et al. 2016). Three basic techniques are normally used to prevent over fitting. These techniques are: early stopping, regularization and cross validation (Chen et al. 2017).

In this current work, an early stopping technique is used to prevent overfitting and subsequently the generalization capability of the ANN network is improved. In this approach, the data set is divided into three main categories: training, validation and testing sets. The computation of the gradient and subsequent changing of the weight values and biases are done with the training set. The validation set is used to ensure the precision and generalization capability of the network in the course of the training process. The final performance of the developed network is verified with the testing set. The test set is not used for training and the validation set is used to end the training process of the network (Azizi et al. 2016).

The application of the proposed methodology on the Tennessee process system is declared as an academic analysis which is in need of further investigation. To the best of the author's knowledge, the bow-tie failure analysis of Tennessee process system is not available. Most available publications on Tennessee chemical plant are in the area of fault diagnosis and process control not in the area of failure assessment. Due to this, the results obtained could not be compared. The Tennessee process system is chosen primarily because of availability of data, which is extremely difficult to access for other process systems.

The integration of an ANN based process accident modeling approach is promising and empirical in nature; however, it is important to state categorially the limitations of ANN which can grossly affect the output of ANN. These limitation are: (1) The quantity and quality of data affect the output of the ANNs (2) There are no clear rules for the best ANN architecture design (3) Different training method used (Chen et al. 2017; Basheer & Hajmeer 2000; Zhang et al. 1998). Also, ANN does not consider physical modelling or physical constraints. ANNs are not suited to make predictions outside the range they have been trained for. The data used in ANN prediction need to be very similar to the set of data used in training stage. ANN output is useful in term of statistical estimation on an ensemble of events, not in terms of deterministic estimation of a single even

## 7.6. Conclusions

The artificial neural network is a data driven nonlinear modelling technique with a strong capability to model nonlinear relationships among process variables. In the present study, ANN is used as a tool to define complex non-linear relationships among process variables. These relationships are critical to model the probability of process failure (considered here as a process accident). The probability of a process accident can be subsequently transformed to risk considering the impact of the accident. The proposed methodology has the following strengths and provides unique opportunity to study process behaviour and avoid unwanted condition:

- The ANN based model offers the unique strength of using real time operation data to model process behaviour.
- The probabilistic approach enables consideration of uncertainty in the data and the model for the prediction of accident scenarios.
- The integrated model provide fasts and reliable generalization of the process data.
- The model uses the strength of both data modelling and the physical behaviour of the process.

The proposed process accident model offers a mechanism to study the dynamic failure profile of a process system. Predictive analysis can be performed on the process accident model.

This study has highlighted the effectiveness of the ANN in modelling the conditional relationships that exist among process variables and subsequently predicting the probability of failure empirically. This approach is highly effective and recommended when physical models that represent such dynamic relationships are not available. Integration of an ANN in a process accident model offers a dynamic approach in predicting the probability of fault from database. The dynamic failure profile estimated using the proposed methodology is used as a dynamic decision making parameter for process safety management.

This work could be further improved by: i) considering the performance of different ANNs architectures, ii) analyzing the performance of ANNs compared with other mathematical data driven models, and iii) validating the proposed model with real industrial data, iv) Providing advanced approach to overcome shortcomings of ANNs beyond already known standard approaches and v) modelling dependencies among accident contributory factors within the safety barriers and allowing non- sequential failure of all safety barriers involved to activate end state adverse events randomly.

## Appendix

**Backpropagation algorithm detail procedure**

The primary objective of back propagation is to optimize different weights so that the neural network can learn how to correctly map arbitrary inputs to output. The mathematical procedure is given as follows (Mazur & Marry 2015)**.**

*Forward pass procedure*:

Refer to Figure 2 and Table2.

Starting with hidden layer.

Firstly, estimating the total net input to each hidden layer of neurons:

$$net_{h_1} = w_1 * i_1 + w_2 * i_2 + b_1 * 1 \qquad A-1$$

$$net_{h_1} = \quad 12.5500$$

The activation function chosen is a logistic function. Squashing the total net input with th activation function chosen to estimate output of hidden neuron $h_1$

$$out_{h_1} = \frac{1}{1 + e^{-net_{h_1}}} \qquad A - 2$$

$out_{h_1} = 0.9999$

Similarly, considering hidden neuron $h_2$

$$net_{h_2} = w_3 * i_1 + w_4 * i_2 + b_1 * 1 \qquad A - 3$$

$net_{h_2} = 6.5500$

$$out_{h_2} = \frac{1}{1 + e^{-net_{h_2}}} \qquad A - 4$$

$out_{h_2} = 0.9985$

Considering the output layer, this procedure is repeated for the neuron in the output layer. In this case, the output of hidden layer neurons serves as the input to the output layer neuron.

$$net_C = w_5 * out_{h_1} + w_6 * out_{h_2} + b_2 * 1 \qquad A - 5$$

$net_C = 3.2994$

$$Output_C = \frac{1}{1 + e^{-net_C}} \qquad A - 6$$

$Output_C = 0.9644$

*Computing the total error*

The squared error function is used to estimate the total error. The total error for the output neuron is computed thus,

$$Error_{total} = \frac{1}{2}(target(T) - output_C)^2 \qquad A - 7$$

$Error_{total} = 0.5362$

*Backward pass procedure:*

The primary objective of the phase is to update all the weights so that, they influence the actual output of the ANN to be closer to the target value.

*Output layer*

Considering the output layer, chain rule can be applied to estimate the gradient with respect to $w_5$:

$$\frac{\partial E_{total}}{\partial ww_5} = \frac{\partial E_{total}}{\partial output_C} * \frac{\partial output_C}{\partial net_C} * \frac{\partial net_C}{\partial w_5} \qquad A-8$$

$$\frac{\partial E_{total}}{\partial output_C} = output_C - (target(T) \qquad A-9$$

$$\frac{\partial E_{total}}{\partial output_C} = -1.0355$$

Computing how $output_C$ change with respect to its net input from equation 6:

$$\frac{\partial output_C}{\partial net_C} = output_C(1 - output_C) \qquad A-10$$

$$\frac{\partial output_C}{\partial net_C} = 0.0343$$

Computing how does $total\ net\ input_C$ change with respect to $w_5$ from equation 5:

$$\frac{\partial net_C}{\partial w_5} = outh_1 = 0.9999 \qquad A-11$$

Computing change in total error with respect to $w_5$

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial output_C} * \frac{\partial output_C}{\partial net_C} * \frac{\partial net_C}{\partial w_5} \qquad A \; -12$$

$$\frac{\partial E_{total}}{\partial w_5} = -0.0355$$

Updating $w_5$,

Let the learning rate $(\eta) = 0.5$

$$w_5^+ = w_5 - \eta * \frac{\partial E_{total}}{\partial w_5} \qquad A \; -13$$

$$w_5^+ = 2.2677$$

Computing how does $total\ net\ input_C$ change with respect to $w_6$ from equation 5:

$$\frac{\partial net_C}{\partial w_6} = outh_2 = 0.9985 \qquad A \; - \; 14$$

$$\frac{\partial E_{total}}{\partial w_6} = \frac{\partial E_{total}}{\partial output_C} * \frac{\partial output_C}{\partial net_C} * \frac{\partial net_C}{\partial w_6} \qquad A \; -15$$

$$\frac{\partial E_{total}}{\partial w_6} = -0.0354$$

Updating $w_6$,

$$w_6^+ = w_5 - \eta * \frac{\partial E_{total}}{\partial w_6} \qquad A \; - \; 16$$

$$w_6^+ = 0.4177$$

*Hidden layer*

The process of backward continue in the hidden layer. There is a modification to the approach adopted here. The process put into consideration the fact that, the output of each of the two hidden neuron contribute significant to the output layer neuron and also the output layer neuron error.The main objective here is to update $W_1$, $W_2$, $W_3$ and $W_4$

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial outh_1} * \frac{\partial outh_1}{\partial neth_1} * \frac{\partial neth_1}{\partial w_1} \qquad A-17$$

Computing change in total error with respect with change in $h_1$

$$\frac{\partial E_{total}}{\partial outh_1} = \frac{\partial Error_C}{\partial outh_1} \qquad A-18$$

Re writing $\frac{\partial Error_C}{\partial outh_1}$ mathematically,

$$\frac{\partial Errror_C}{\partial outh_1} = \frac{\partial Error_C}{\partial net_C} * \frac{\partial net_C}{\partial outh_1} \qquad A-19$$

Computing $\frac{\partial Error_C}{\partial net_C}$ from previously calculated values in equations 9 and 10:

$$\frac{\partial Error_C}{\partial net_C} = \frac{\partial Error_C}{\partial output_C} * \frac{\partial output_C}{\partial net_C} \qquad A-20$$

$$\frac{\partial Error_C}{\partial net_C} = -0.0355$$

Computing how does $total\ net\ input_C$ change with respect $\partial outh_1$from equation 5:

$$\frac{\partial net_C}{\partial outh_1} = w_5 \qquad A-21$$

$$\frac{\partial net_C}{\partial outh_1} = 2.2500$$

Evaluating equation 19 by substituting the values of equations 20 and 21:

$$\frac{\partial Errror_C}{\partial outh_1} = -0.0799 \qquad A-22$$

Evaluating derivative of $\frac{\partial outh_1}{\partial neth_1}$ from equation 2

$$\frac{\partial out\ h_1}{\partial net\ h_1} = out\ h_1(1 - out\ h_1) \qquad A-23$$

$$\frac{\partial out\ h_1}{\partial net\ h_1} = 3.544E - 06$$

Evaluating derivative of $\frac{\partial net\ h_1}{\partial W_1}$ from equation 1:

$$\frac{\partial net\ h_1}{\partial W_1} = i_1 = 10 \qquad A-24$$

Substituting all values calculated in equation 22, 23 and 24 into equation 17:

$$\frac{\partial E_{total}}{\partial w_1} = -2.83\mathrm{E} - 06$$

Updating $w_1$

$$w_1^+ = w_1 - \eta * \frac{\partial E_{total}}{\partial w_1} \qquad A-\ 25$$

$$w_1^+ = 0.7000$$

Similarly for $w_2$, evaluating derivative of $\frac{\partial net\ h_1}{\partial W_2}$ from equation 1:

$$\frac{\partial net\ h_1}{\partial W_2} = i_2 = 8 \qquad A\ -26$$

$$\frac{\partial E_{total}}{\partial w_2} = \frac{\partial E_{total}}{\partial outh_1} * \frac{\partial outh_1}{\partial neth_1} * \frac{\partial neth_1}{\partial w_2} \qquad A\ -27$$

Substituting all values calculated in equation 22, 23 and 26 into equation 27:

$$\frac{\partial E_{total}}{\partial w_2} = -2.26\mathrm{E} - 06$$

Updating $w_2$

$$w_2^+ = w_2 - \eta * \frac{\partial E_{total}}{\partial w_2} \qquad A-28$$

$$w_2^+ = 0.6000$$

Repeating similar process for the second hidden neuron $h_2$

$$\frac{\partial E_{total}}{\partial w_3} = \frac{\partial E_{total}}{\partial outh_2} * \frac{\partial outh_2}{\partial neth_2} * \frac{\partial neth_2}{\partial w_3} \qquad A-29$$

Computing change in total error with respect with change in $h_2$

$$\frac{\partial E_{total}}{\partial outh_2} = \frac{\partial Error_C}{\partial outh_2} \qquad A\ -30$$

Re writing $\frac{\partial Error_C}{\partial outh_2}$ mathematically,

$$\frac{\partial Errror_C}{\partial outh_2} = \frac{\partial Error_C}{\partial net_C} * \frac{\partial net_C}{\partial outh_2} \qquad A-31$$

Computing how does $total\ net\ input_C$ change with respect $\partial outh_2$from equation 5:

$$\frac{\partial net_C}{\partial outh_2} = w_6 \qquad A-32$$

$$\frac{\partial net_C}{\partial outh_2} = 0.4$$

Evaluating equation 31 by substituting the values of equations 20 and 32:

$$\frac{\partial Errror_C}{\partial outh_2} = -0.0142$$

Evaluating derivative of $\frac{\partial outh_2}{\partial neth_2}$ from equation 4:

$$\frac{\partial out\ h_2}{\partial net\ h_2} = out\ h_2(1 - out\ h_2) \qquad A-33$$

$$\frac{\partial out\ h_2}{\partial net\ h_2} = 1.43E - 03$$

Evaluating derivative of $\frac{\partial net\ h_2}{\partial W_3}$ from equation 3

$$\frac{\partial net\ h_2}{\partial W_3} = i_1 = 10 \qquad A-34$$

Substituting all values calculated in equation 31, 33 and 34 into equation 29:

$$\frac{\partial E_{total}}{\partial w_3} = -2.03\text{E} - 04$$

Updating $w_3$

$$w_3^+ = w_3 - \eta * \frac{\partial E_{total}}{\partial w_3} \qquad A-35$$

$$w_3^+ = 0.5001$$

Similarly for$w_4$, evaluating derivative of $\frac{\partial net\ h_2}{\partial W_4}$ from equation 3:

$$\frac{\partial net\ h_2}{\partial W_4} = i_2 = 8 \qquad A-36$$

$$\frac{\partial E_{total}}{\partial w_4} = \frac{\partial E_{total}}{\partial outh_2} * \frac{\partial outh_2}{\partial neth_2} * \frac{\partial neth_2}{\partial w_4} \qquad A-37$$

Substituting all values calculated in equation 31, 33 and 36 into equation 37:

$$\frac{\partial E_{total}}{\partial w_4} = -0.0002$$

Updating $w_4$

$$w_4^+ = w_4 - \eta * \frac{\partial E_{total}}{\partial w_4} \qquad A-38$$

$$w_4^+ = 0.1000$$

*Each of the weights has now been updated. These updated weights are feedback to the ANN architecture. This process is repeated continuously until the error is exceedingly small (within the tolerance limit). The learning rate could be varied to accelerate training.

## Acknowledgement

The authors gratefully recognise the financial support provided by Canada Research Chair (Tier I) and Natural Science and Engineering Council of Canada.

## References

Adedigba, S.A., Khan, F. & Yang, M., 2017. Dynamic Failure Analysis of Process Systems Using Principal Component Analysis and Bayesian Network. *Industrial & Engineering Chemistry Research*, 56(8), pp.2094–2106.

Adedigba, S.A., Khan, F. & Yang, M., 2016. Dynamic safety analysis of process systems using nonlinear and non-sequential accident model. *Chemical Engineering Research and Design*, 111, pp.169–183.

Adedigba, S.A., Khan, F. & Yang, M., 2016. Process accident model considering dependency among contributory factors. *Process Safety and Environmental Protection*, 102, pp.633–647.

Al-shanini, A., Ahamad, A. & Khan, F., 2014. Accident modelling and analysis in process industries. *Journal of Loss Prevention in the Process Industries*, 32, pp.319–334.

Ashtiani, H.R.R. & Shahsavari, P., 2016. A comparative study on the phenomenological and artificial neural network models to predict hot deformation behavior of AlCuMgPb alloy. *Journal of Alloys and Compounds*, 687, pp.263–273.

Attwood, D., Khan, F. & Veitch, B., 2006. Occupational accident models-Where have we been and where are we going? *Journal of Loss Prevention in the Process Industries*, 19, pp.664–682.

Azizi, S., Ahmadloo, E. & Awad, M.M., 2016. Prediction of void fraction for gas–liquid flow in horizontal, upward and downward inclined pipes using artificial neural network. *International Journal of Multiphase Flow*, 87, pp.35–44.

Basheer, I.A. & Hajmeer, M., 2000. Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1), pp.3–31.

Chen, F.F., Breedon, M., White, P.,Chu, C.,Mallick, D.,Thomas,S., Sapper, E.,& Cole, I., 2016. Correlation between molecular features and electrochemical properties using an artificial neural network. *Materials & Design*, 112, pp.410–418.

Chen, Z., Ma, W., Wei, K., Wu, J., Li, S., Xie, K., & Lv, G., 2017. Artificial neural network modeling for evaluating the power consumption of silicon production in submerged arc furnaces. *Applied Thermal Engineering*, 112, pp.226–236.

Chitsazan, N., Nadiri, A. & Tsai, F., 2015. Prediction and structural uncertainty analyses of artificial neural networks using hierarchical Bayesian model averaging. *Journal of Hydrology*.

Crowl, D.. A. & Louvar, J.F., 2011. *Chemical process safety:fundamentals with appplication* Third edit., Pearson Education.

Downs, J. & Vogel, E., 1993. A plant-wide industrial process control problem. *Computers & chemical engineering*, 7(3), pp.245–255.

Ferdous, R.,Khan, F., Sadiq, R., Amyotte, P., & Veitch, B., 2009. Handling data uncertainties in event tree analysis. *Process Safety and Environmental Protection*, 87(5), pp.283–292.

Gardner, M.W. & Dorling, S.R., 1998. Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14-15), pp.2627–2636.

Gonzaga, J.C.B., Meleiro, L.A.C., Kiang, C., &  Filho, R.M.,2009. ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerizationg process. *Computers & Chemical Engineering*, 33(1), pp.43–49.

Khan, F., Hashemi, S.J., et al., 2016. Dynamic risk management: a contemporary approach to process safety management. *Current Opinion in Chemical Engineering*, 14, pp.9–17.

Khan, F., Rathnayaka, S. & Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection*, 98, pp.116–147.

Khan, F., Wang, H. & Yang, M., 2016. Application of loss functions in process economic risk assessment. *Chemical Engineering Research and Design*, 111, pp.371–386.

Khan, F.I. & Abbasi, S.., 1998. Techniques and methodologies for risk analysis in chemical process industries. *Journal of Loss Prevention in the Process Industries*, 11(4), pp.261–277.

Labeau, P.E., Smidts, C. & Swaminathan, S., 2000. Dynamic reliability: towards an integrated platform for probabilistic risk assessment. *Reliability Engineering & System Safety*, 68(3), pp.219–254.

Lee, M., Lee, S.. & Park, C.., 1992. Neural controller of nonlinear dynamic systems using higher order neural networks. *Electronic letters*, 28(3), p.276.

Lennox, B. & Montague, G.A., 1997. Dynamic process modelling and fault prediction using artificial neural networks. In E. Fiesler & R. Beale, eds. *Handbook of Neural computation*.

Mazur, M. & Marry, L., 2015. A step by step Backpropagation example.

Meel, A. & Seider, W.D., 2006. Plant-specific dynamic failure assessment using Bayesian theory. *Chemical Engineering Science*, 61(21), pp.7036–7056.

Qureshi, Z.H., 2007. A Review of Accident Modelling Approaches for Complex Socio-Technical Systems. In *12th Austrialian Conference on safety Related progarammable Systems*. pp. 47–60.

Rathnayaka, S., Khan, F. & Amyotte, P., 2012. Accident modeling approach for safety assessment in an LNG processing facility. *Journal of Loss Prevention in the Process Industries*, 25(2), pp.414–423.

Rathnayaka, S., Khan, F. & Amyotte, P., 2010. SHIPP methodology : Predictive accident modeling approach . Part II . Validation with case study. *Process Safety and Environmental Protection*, 89(2), pp.75–88.

Rathnayaka, S., Khan, F. & Amyotte, P., 2011. SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection*, 89(3), pp.151–164.

Riedmiller, M., 1994. Advanced supervised learning in multi-layer perceptrons—from backpropagation to adaptive learning algorithms. *Computer Standards & Interfaces*, 16(3), pp.265–278.

Sanusi, S.N.A., Halmi, M.I.E.,Abdullah, S.R.S.,Hassan, H.A., Hamzah, F.D., & Idris, M., 2016. Comparative process optimization of pilot-scale total petroleum hydrocarbon (TPH) degradation by Paspalum scrobiculatum L. Hack using response surface methodology (RSM) and artificial neural networks (ANNs). *Ecological Engineering*, 97, pp.524–534.

Świetlicka, I., Sujak, A., Muszynski, S., & Swietlicki, M., 2017. The application of artificial neural networks to the problem of reservoir classification and land use determination on the basis of water sediment composition. *Ecological Indicators*, 72, pp.759–765.

Tan, Q.,Chen, G., Zhang, L., Fu, J., & Li, Z., 2013. Dynamic accident modeling for high-sulfur natural gas gathering station. *Process Safety and Environmental Protection*, 92(6), pp.565–576.

Villa, V., Paltrinieri, N., Khan, F., & Cozzani, V., 2016. Towards dynamic risk analysis: A review of the risk assessment approach and its limitations in the chemical process industry. *Safety Science*, 89, pp.77–93.

Xu, J.-X. & Hou, Z.-S., 2009. Notes on Data-driven System Approaches. *Acta Automatica Sinica*, 35(6), pp.668–675.

Yu, H.,Khan, F., Garaniya, V., & Ahmad, A., 2014. Self-Organizing Map Based Fault Diagnosis Technique for Non-Gaussian Processes. *Industrial & Engineering Chemistry Research*, 53(21), pp.8831–8843.

Yu, H., Khan, F. & Garaniya, V., 2015. Risk-based fault detection using Self-Organizing Map. *Reliability Engineering & System Safety*, 139, pp.82–96.

Yu, S.-H. & Annaswamy, A.M., 1997. Adaptive control of nonlinear dynamic systems using θ-adaptive neural networks. *Automatica*, 33(11), pp.1975–1995.

Zhang, G., Eddy Patuwo, B. & Y. Hu, M., 1998. Forecasting with artificial neural networks: *International Journal of Forecasting*, 14(1), pp.35–62.

Zhang, J. & Morris, A.., 1995. Dynamic process modelling using locally recurrent neural networks. In *American control conference*. Seattle,Washington, p. 2767.

Zhong, Y. et al., 2016. The Method of Oilfield Development Risk Forecasting and Early Warning Using Revised Bayesian Network. , 2016.

# Chapter 8

## 8.0 An Integrated Approach for Dynamic Economic Risk Assessment of Process System

### Preface

*A version of this chapter has been published in* ***Journal of Process Safety and Environmental Protection 2018; 116:312-323.*** *I am the primary author. Co-author Faisal Khan provided fundamental understanding, assisted in developing the conceptual model and subsequently translated this to the numerical model. Co-author Ming Yang provided much needed support in implementing the concept and testing the model. I carried out most of the data collection and analysis. I prepared the first draft of the manuscript and subsequently revised the manuscript, based on the feedback from co-authors and also a peer review process. The two co-authors assisted in developing the concept and testing the model, reviewed and corrected the model and results. They also contributed to the review and revision of the manuscript.*

### Abstract

This paper proposes a dynamic economic risk analysis methodology for process systems. The Bayesian Tree Augmented Naïve Bayes (TAN) algorithm is applied to model the precise and concise probabilistic dependencies that exist among key operational process variables to detect faults and predict the time dependent probability of system deviation. The modified inverted normal loss function is used to define system economic losses as a function of process deviation. The time dependent probability of system deviation owing to an abnormal event is constantly updated based on the present state of the relevant process variables. The integration of real time probability of system deviation with potential losses provides the risk profile of the system at any instant. This risk profile can be used as the basis for operational decision making and also to

activate the emergency safety system. The proposed methodology is tested and verified using the Richmond refinery accident.

**Keywords:** Dynamic failure prediction, Loss functions, economic consequences, Process safety, Structure learning of Bayesian network from data and Risk analysis.

## 8.1. Introduction

Recently, industrial technological systems have become extremely complex and more susceptible to process accidents. The risk posed by these systems is alarming and more worrisome. Operating modern industrial technological systems requires high levels of safety and reliability. Realising a high level of safety and reliability calls for effective management of these industrial technological systems' performance alongside management of process safety (Al-shanini et al. 2014; Adedigba et al. 2016). This development justifies the need for efficient and effective process safety and risk management techniques that will drastically reduce both the chance (probability) and the severity (consequences) of process accidents. Major research has been conducted on how to better monitor chemical process operations, evaluating risk and subsequently the development of safety systems for chemical process operations (Khan et al. 2015). A vital part of a safety system is the identification of possible hazards related to a process and the evaluation of both the likelihood of their occurrence and the associated consequences. This method is referred to as quantitative risk assessment (QRA) (Kalantarnia et al. 2010).

QRA is a widely used technique that provides dependable quantitative information on risk initiated by conventional accidents in a chemical process operation. It vividly describes the best available analytic predictive data to compute the risk of a chemical process. The QRA technique consists of various steps: hazard identification, frequency estimation, and consequence analysis and risk quantification. The foremost step of QRA techniques is hazard identification, which is of paramount importance because it highlights potential faults of the system, identifies undesired top events and finally explains possible scenarios related to the top events (Villa et al. 2016). Quite a few methods have been developed for hazard identification. Details of these methods and their description is given by ( Khan & Abbasi 1998).

Risk analysis is defined as the “ the development of a quantitative estimate of risk based on an engineering evaluation and mathematical techniques for combining estimates of incident consequences and frequencies” (Crowl & Louvar 2011). Two principal parts of risk assessment are incident identification and consequence analysis. A vivid explanation of exactly how and why an accident occurs is established by incident identification. In most cases, it involves estimating occurrence probabilities. Consequences analysis explains and quantifies the expected damages. Over past decades, numerous techniques have been proposed and developed for quantitative risk assessment in chemical process industries ( Khan & Abbasi 1998). However, the majority of the methodologies developed over the past decades focuses mainly on estimation or prediction of occurrence probabilities (CCPS 1999; Crowl & Louvar 2011; Khan & Abbasi 1998; Khan et al. 2015). While significant advancement has been made in assessing abnormal process conditions, fewer efforts have been made to analyze the economic impact of these abnormal conditions. This study attempts to fill the gap.

The primary objective of the study is to develop a dynamic economic risk assessment framework which integrates probability with consequences assessment. The proposed approach links process deviations (from target/normal operation conditions) to accident probability and potential losses.

This paper is organised as follows. Section 8.2 discusses the effectiveness of assessing process operational risk based on process deviations and briefly reviews loss functions. Section 8. 3

presents the proposed methodology. Section 8.4 provides the validation of the proposed methodology with a case study of the Richmond refinery accident with the results and discussion. Lastly, the conclusion is given in Section 8.5.

## 8.2. Process deviation, risk assessment and loss functions

Modern industrial technological systems are exceptionally complex and consist of several dynamic process variables for their operations. Process safety management includes thorough identification of all abnormal process conditions and subsequent application of corrective actions prior to the time the abnormal conditions lead to devastating consequences (Khan, et al. 2016). Risk analysis is an appropriate technique to evaluate the performance of process safety management. Process safety measures manage risk associated with a system through risk evaluation, design improvement and risk based decision making. (Khakzad et al. 2013).

Process deviations combined with failure of protection layers and control systems induce failures that increase the likelihood of an accident and subsequently increase the operational risk associated with process systems (Hashemi, et al. 2014). The effectiveness of assessing process operational risk based on process deviations needs serious consideration. Conventional risk assessment techniques lack the capability to account for risk variations due to process deviations because the conventional methods are static in nature (Wang et al. 2016; Khan, et al. 2016).

In reality, there is a relationship between process deviation with both accident probability and consequence assessment. Deviations are needed to cause violations or disruptions of safety objectives. Process deviation causes failure of safety barriers which affects both accident probability and the impact of the consequences (Khorsandi & Aven 2017). Therefore, the impact of process deviations on both the accident occurrence probability and associated economic consequence is of paramount importance in risk assessment of process systems. This approach provides variation in risk due to process deviations.

### 8.2.1. A Brief Review of loss functions

Loss functions (LFs) are commonly used to compute losses related to deviation of a product from optimal value. Quite recently, loss functions have gained wide acceptance among researchers and quality assurance practitioners due to Taguchi's philosophy and his quality improvement strategies (Zadakbar et al. 2015; Spiring 1993). Several types of loss functions have been discussed in the literature. Below is a brief review of some types of loss functions.

#### *8.2.1.1. Quadratic loss function*

The quadratic loss function was proposed by Taguchi (Taguchi 1986). It has been extensively used to compute losses connected with deviation of products from their optimal value. Mathematically, the quadratic loss function is represented as:

$$L(y) = B(y - T)^2 \tag{8.1}$$

L(y) represents the actual loss at y, T represents the desired target value and B is the proportionality constant (Khan, et al. 2016; Hashemi, et al 2014). However, the quadratic loss function exhibits some weaknesses and has been severely criticized by researchers and quality assurance practitioners. The weaknesses include: (1) the inability to give quantifiable maximum loss (unbounded nature) (2) the symmetric nature of the quadratic loss function, making it impossible to predict the magnitude of losses connected with extreme departures from the target (Zadakbar et al. 2015). Because of these weaknesses, the quadratic loss function has been improved or revised to give quantifiable maximum loss by truncating the quadratic loss function at the point at which it intersects the maximum loss. The mathematical representation of this modified quadratic loss function is:

$$L(y) = \begin{cases} B\big((y-T)^{2,} & |y-T| \leq \sqrt{K/B,} \\ K, & |y-T| > \sqrt{K/B,} \end{cases} \tag{8.2}$$

where L(y) represents the actual loss at y, T represents the desired target value, B is the proportionality constant, K denotes maximum loss and $\Delta = \sqrt{K/B}$ represents the total distance from T (Target) to the point where K (maximum loss) first occurs (Khan, Wang, et al. 2016; Spiring 1993).

#### *8.2.1.2. Quartic loss function*

Due to the weaknesses exhibited by the quadratic loss function: its symmetric nature and lack of a specific functional profile (Khan, et al. 2016), quartic loss function was proposed by Fathi and Poonthanomsook (2007) to denote continuous loss functions. The quartic loss function shape is either symmetric or asymmetric within the specification limit, by applying Taylor series expansion and its parameters. Mathematically, quartic loss function can be represented as:

$$L(y) = K_2(y-T)^2 + K_3(y-T)^3 + K_4(y-T)^4 \tag{8.3}$$

The second, third and fourth order loss coefficients are represented as: $k_{2,}$ $k_{3,}$ and $k_4$ respectively. Interestingly, both the quartic loss and quadratic functions lack the capability to adequately account for the nonlinear nature of chemical process operations (Khan, et al. 2016).

#### *8.2.1.3. Inverted normal loss function*

Normal distribution is normally used to describe random variables operations (Khan, et al. 2016). The inverted normal loss function (INLF) is proposed by Spiring (1993) to overcome the obvious weakness of quadratic loss functions. A transformation or simple modification of the normal density function give an alternative that permits a single function to define the loss. The inverted or reflected normal loss function can be mathematically represented as:

$$L(y) = K\left\{1 - exp\left(-\frac{(y-T)^2}{2\gamma^2}\right)\right\} \tag{8.4}$$

where, y denotes the quality, K represents the maximum loss, $\gamma$ represents a shape parameter and T represent the target value. The shape parameter is normally used to modify or alter the general form of INLF. The shape parameter is mathematically defined as:

$$\gamma = \frac{\Delta}{4} \tag{8.5}$$

where $\Delta$ represents the total distance from T (Target) to the point where K (maximum loss) first takes place. The curve produced or generated with the inverted normal loss function is smooth; it gives a minimum of zero at the target and subsequently gives a quantifiable maximum. It satisfactorily meets the basic requirements of the loss function: (1) decreasing at the interval (- ∞, T] and also increasing at [T,∞); (2) it gives a specific minimum; and finally (3) it gives general shapes that can be changed or altered. Additionally, the INLF can be adopted to a situation of asymmetric loss around the target and can also be used to denote the quadratic loss function over a defined region (Spiring 1993). The shape parameter alongside the maximum loss permits the modification of the loss function to accommodate various incremental losses that occur for various processes and therefore makes it more flexible, and giving a more realistic assessment of losses associated with the departures from optima value or the target. The inverted normal loss function offers a better alternative to the modified quadratic function (Zadakbar et al. 2015; Spiring 1993).

#### *8.2.1.4. Modified inverted normal loss function (MINLF)*

Sun et al. (1996) carried out the modification of INLF to a modified inverted normal loss function (MINLF). MINLF offered a more reasonable loss representation and at the same time presented a technique for revising the modified inverted normal loss function graph to replicate the operator's actual loss. The MINLF adopts a nonlinear least squares technique for computing the shape of the MINLF. The shape parameter $\gamma$ is determined by the user; the value of this shape parameter decides the slope of the function around the optimal or desired value (Hashemi, et al 2014; Khan, et al. 2016). The MINLF can be mathematically represented as:

$$L(y) = EML \frac{1}{1 - e^{-\Delta^{\frac{2}{}} 2\gamma^2}} (1 - e^{-(y-T)^{\frac{2}{2\gamma^2}}}) \qquad (8.6)$$

where y represents the quality measurement, L(y) represents the actual loss at y, T represents the optimal value (target) and EML is the estimated maximum loss. $\Delta$ represents the total distance from T (Target) to the point where K (maximum loss) first takes place and $\gamma$ represents the shape parameter (Khan, et al. 2016).

The MINLF gives more flexibility in demonstrating symmetric loss situations by specifying their shape parameters. This parametrization of MINLF aids its comparison with the quadratic loss function (Sun et al. 1996).

#### *8.2.1.5. Inverted Beta loss function (IBLF)*

A class of symmetric and asymmetric loss is developed by Leung & Spiring (2002) using an inverted beta probability distribution function. The IBLF can be mathematically represented as:

$$L(y) = EML(1 - (T(1-T)^{\frac{1-T}{T}})^{1-\alpha}(y(1-y)^{\frac{1-T}{T}})^{\alpha-1}) \quad (8.7)$$

where y denotes the quality measurement, L(y) represents the actual loss at y, T represents the desired value, and EML represents the estimated maximum loss. $\alpha$ and $\beta$ represent the shape parameters that are quantified for the IBLF. The $\beta$ in IBLF is denoted by $\alpha$ and T respectively. The mathematical relationship between $\alpha$ and T is given below (Khan, Wang, et al. 2016).

$$\beta = 1 + \frac{(\alpha-1)(1-T)}{T} \quad (8.8)$$

The family of the Inverted Beta loss function (IBLF) provides flexibility to the shape their loss functions can reach and provides an adequate technique for designing the IBLF to show the actual loss suffered (Leung & Spiring 2002). IBLF provides both the traditional (basic) features of loss functions and asymmetrical loss situations (Hashemi, Ahmed & F. Khan 2014).

### *8.2.1.6. Inverted Gamma loss function (IGLF)*

The inverted gamma loss function (IGLF) provides the base for a classification of loss functions that is used to describe process operations with asymmetric loss. The shape of the IGLF is asymmetric around the target (Spiring & Yeung 1998; Leung & Spiring 2004). The IGLF is mathematically represented as:

$$L(y) = EML\left(1 - \left(\frac{ye^{1-(y/T)}}{T}\right)^{\alpha-1}\right) \quad (8.9)$$

where EML is the estimated maximum loss, y represents the process measurement, T denotes the desired value, and $\alpha$ is the shape parameter. The shape parameter $\alpha$ permits practitioners to modify the Inverted gamma loss function to precisely reflect losses connected with the deviations from the target (Spiring & Yeung 1998).

Hashemi et al. (2014) thoroughly reviewed the applicability of various kinds of loss functions to process operation safety assessment and deduced that; MINLF and IBLF are more flexible (adaptable) to represent losses connected with process deviations. However, Hashemi et al. (2014) emphasied that MINLF is better compared to IBLF because it is simple to formulate and at the same time gives a more robust perfomance during sensitivity analysis; consequently the MINLF function is highly endorsed for predicting loss associated with process accidents.

The choice of which loss function to select to translate the process deviation into real time economic loss depend on the prevailing process conditions and the users' objectives. Characteristics of various loss functions alongside their sensitivity analysis can be adopted in selection of appropriate loss functions (Hashemi et al. 2014).

## 8.3 The proposed methodology

The flowchart for the proposed methodology is given in Figure 8.1.

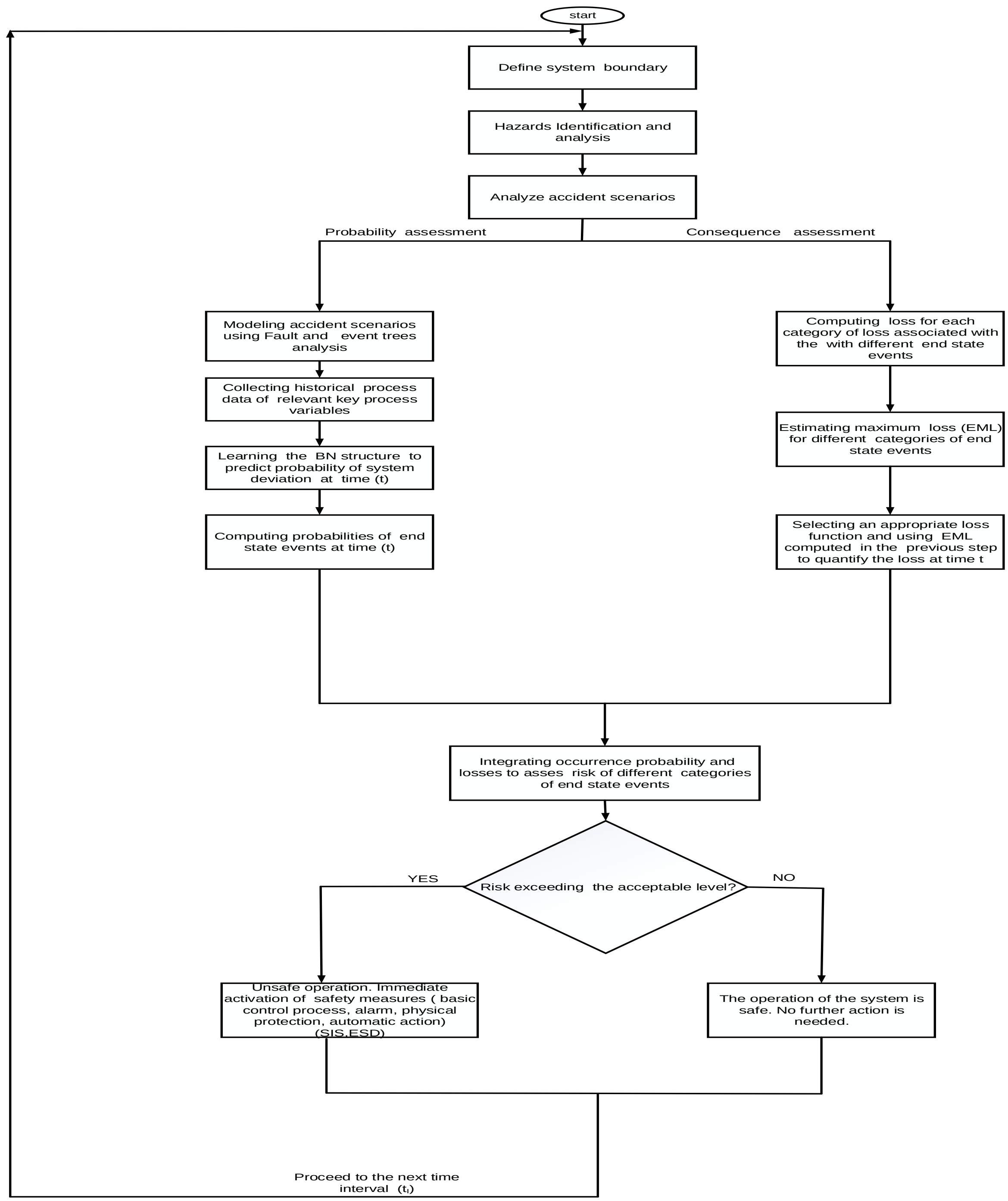



**Figure 8 .1.The flowchart for the proposed methodology.**

### 8.3.1 Hazard identification and analysis

The main aim of this phase is identification of potential hazards and the underlying initiating factors that may be responsible for the hazards. Various hazard identification techniques can be applied at this phase. Different techniques for hazard identification such as the hazard operability study (HAZOP) and failure mode and effect analysis (FMEA) can be applied at this stage. A thorough review of the hazard identification technique is given by Khan et al.( 1998). The accident sequence pattern also needs to be determined; it helps to identify the necessary safety barriers that are needed to be placed along the accident path to inhibit the initiation and progression of the accident process. Accidents usually happen due to failure of relevant safety barriers (Rathnayaka et al. 2011). One common process hazard is process deviations which include high pressure, high temperature and steam hammering (Rathnayaka et al. 2011; Adedigba et al. 2017a)

### 8.3.2. Failure probability assessment of safety barriers.

Fault tree analysis (FTA) is a very reliable technique to predict the probability of a hazard due to sequences of failure events. Fault tree analysis has been widely used for both quantitative reliability and safety analysis (Durga Rao et al. 2009). A fault tree is a logical graphical model representing various combinations of sequential and/or parallel faults events that could cause the occurrence of the pre-determined undesired top event. It vividly represents logical connections among basic events to the topmost undesired event. Normally, Boolean gates are used to represent the relationship between the basic events and the top event. The occurrence probability of the topmost unwanted event of the fault tree depends strongly on the reliability data of the causative events. A fault tree analysis could be carried out in various ways: qualitatively, quantitatively and with a combination of both qualitative and quantitative techniques (Adedigba et al. 2016; Khakzad et al. 2011). Fault tree analysis is carried out quantitatively in the proposed approach and failure probability data of the basic event are sourced from several journals and expert opinions.

The safety barriers identified during hazards identification analysis are methodically analyzed with the fault tree to denote causative relationships. Subsequently, the occurrence probability of the uppermost unwanted event is predicted from the reliability data of the causative events.

### 8.3.3. Analysis of potential scenario using event tree analysis

The event tree offers a systematic technique for thoroughly investigating all potential accident scenarios involving a complex system. Event tree analysis is an inductive technique that begins with a specified initiating event (Nivolianitou et al. 2004). An event tree describes a logical combination of several events that may follow after initiation of an accident sequence. The event sequence in the event tree is influenced by either success or failure of various prevention or safety barriers placed along the accident path (Ferdous et al. 2009). The event sequences in the event tree eventually lead to a set of possible outcomes generally called the consequences. In the present case study, the initiating event is the process deviation.

### 8.3.4. Collecting historical process data of relevant key process variables

In this phase, the relevant key process variables are selected. The historical process data of key process variables are methodologically analysed and set points are fixed for each of the key process variables to determine their state. The states of the historical data of relevant process variables are used to learn the Bayesian network structure. The detailed procedure is given in Adedigba et al.( 2017b).

### 8.3.5. Structure learning of Bayesian network using state operational historical data

A Bayesian network (BN) is a graphical modelling and inferencing technique for problems involving uncertainty. Bayesian networks vividly model the probabilistic relationships among a group of variables (Heckerman 1997). A Bayesian network consists of directed acyclic graphs with a group of conditional probability tables (CTPs) related to them. Usually CTPs are allotted to the numerous nodes to indicate probabilistic relations among the linked nodes.

The joint probability distribution P (U) of a collection of random variables U= $\{A_1\ldots,A_n\}$, is integrated as:

$$\mathbf{P(U) = \prod_{i=1}^{n} P(A_i | \ P_{a(A_i)}))} \qquad \mathbf{(8.10)}$$

where Pa ($A_i$) is the parent of variable $A_i$ and P (U) is the joint probability distribution of variables (Jensen & Nielsen 2007).

The Bayesian probability theorem offers a unifying methodology for data modelling. A Bayesian data modeller's offers a unique opportunity to develop probabilistic models that adequately fit the data presented to them (Mackay 1995). Formerly, the direct acyclic graph (DAG) of the BN was normally hand-constructed by the experts and the conditional probabilities table (CTP) was provided based on expert opinions. However, recently researchers have developed methods that can both learn DAG (structure) and provide the CPT from data. The methods are classified into two categories: constraint based methods and Bayesian based methods (Jensen & Nielsen 2007). The detailed procedure involved in generating both the structure and the CTP of BN using one of the Bayesian score methods, called the tree augmented naïve Bayes algorithm, is presented in Adedigba et al.( 2017b). To avoid duplication of information, the detailed methodology is not presented here. The probabilistic network's structure generated from operational process data will give an exact representation of probabilistic dependencies that exist among process operational data. The primary reason to use operational process data to construct a probabilistic networks structure is to predict the probability of process deviation based on deviation of process variables from their optimal values.

### 8.3.6. Computing the probability of end state event consequence at time (t)

The time dependent probabilities of system deviation predicted from the learned structure of the BN are used as an input to the event to compute the time dependent probability of end state consequences.

The probability of end state consequences in the event tree P ($C_k$) is computed by equation 11.

$$\mathbf{P(C_K)} = \prod_{\mathbf{j \in SB_k}} \mathbf{x_i^{\theta_{i,k}}} \, (\mathbf{1} - \mathbf{x_i})^{\mathbf{1-\theta_{i,k}}} \qquad (\mathbf{8.11})$$

where $SB_k$ signifies the safety barrier linked to level k; and $\theta_{i,k} = 1$ every time a level k failure travels through the failure branch of safety barrier i; $\theta_{i,k} = 0$ every time a level k failure goes through the success branch of safety barrier i. $x_i$ denotes the failure probability of safety barriers.(Adedigba et al. 2016; Rathnayaka et al. 2010).

### 8.3.7. Consequence assessment using loss functions

Loss functions have been adopted to predict losses connected with the deviation of operational variables. Quite recently, the application of inverted probability distributions for quantifying losses is gaining wide acceptance among researcher and practitioners. Hashemi et al. (2014) thoroughly reviewed the applicability of various kinds of loss functions to process operation safety assessment and deduced that MINLF and IBLF are more flexible (adaptable) to represent losses connected with process deviations. However, Hashemi et al. (2014) emphasised that MINLF is better compared to IBLF because it is very simple to formulate and at the same time gives a more robust performance during sensitivity analysis; consequently, the MINLF loss function is used for predicting loss associated with accidents in the proposed methodology. The following procedures are proposed for loss modelling of a process unit.

(1) Identification of key process variables

In loss modelling of a process unit, there is a need to identify key process variable/variables that might be responsible for undesired events among all process variables peculiar to the process unit under consideration. Common examples of process variables are; temperature, pressure and concentration. A process variable or variables with a direct bearing on the undesired events should be chosen as key variable/variables in loss modelling.

(2) Identification of loss categories and computing loss for various categories of losses

Generally, for any incident scenario in a process operation, four basic categories of losses can be established. The four categories of loss are well defined in (Khan & Amyotte 2005). The following equations are used to predict each category of loss.

Production loss (PL)

$$C_{PL} = Likely\ downtime\ (hours)\ X\ Production\ value\ \left(\frac{\$}{hours}\right) \tag{8.12}$$

Asset loss (AL)

$$C_{AL} = Asset\ density\left(\frac{\$}{area}\right) X\ Damage\ area \tag{8.13}$$

Human health loss (HHL)

$$C_{HHL} = Damage\ area\ X\ Population\ density \left(\frac{people}{area}\right) X\ Cost\ of\ \frac{fatality}{injury}\ (\$) \qquad (8.14)$$

Environmental cleanup cost (ECC)

$$C_{ECC} = C_{Soil} + C_{Water} + C_{Air} \qquad (8.15)$$

where each category of cleanup cost is predicted as:

$$C_{Soil} = Mass\ of\ contaminated\ soil\ X\ Cleanup\ cost \left(\frac{\$}{mass}\right) X\ NH \qquad (8.16)$$

$$C_{Water} = Volume\ \ of\ contaminated\ waterX\ Cleanup\ cost \left(\frac{\$}{volume}\right) X\ NH \qquad (8.17)$$

$$C_{Air} = Volume\ \ of\ contaminated\ air\ X\ Dilution\ or\ cleanup\ cost \left(\frac{\$}{volume}\right) X\ NH \qquad (8.18)$$

The symbol NH denotes the NFPA ranking of chemicals (Khan & Amyotte 2005). The fire explosion index methodology is applied to compute the damage radius and damage area. The damage radius is computed using equation 19 (Jafari et al. 2012).

$$R = 0.256\ X\ FEI \qquad (8.19)$$

where FEI is the fire explosion index

(3) Estimation of maximum loss.

Estimated maximum loss (EML) is defined as “ the loss that could be sustained under an abnormal condition with the failure of all protective systems” (Marsh Risk 2011). Therefore, the estimated maximum loss is the summation of all these loss.

$$EML = C_{PL} + C_{AL} + C_{HHL} + C_{ECC} \qquad (8.20)$$

(4) Selecting an appropriate type of loss function and determination of its shape parameter

An appropriate loss function is selected to quantify each category of losses associated with any given incident scenario. Upon selection of the desired type of loss function there is a need to determine its shape parameter. Usually regression techniques that adopt non- linear search approaches are used to find the shape parameters. A good example of these techniques is the least square technique.

### 8.3.8. Risk Estimation

The time dependent risk profile is developed by combining the real time occurrence probability of consequences predicted from process variables deviation with the loss associated with an accident at any time (t). Plant operators can easily access the risk associated with the process operations at any given time. The acceptable risk of the process operations should be bounded: acceptable risk should have lower and upper limits. The operators will be able to alert the decision maker whenever the upper bound of the acceptable risk is exceeded. This will enable the decision maker to take remedial action before the process deviation results in devastating losses.

## 8.4. Case study: Richmond refinery accident

### 8.4.1. Description of the case

The Richmond refinery occupies nearly 2900 acres of the San Pablo Peninsula in the U.S.A. The refinery processed 250,000 barrels of crude oil per day and nearly 1200 people were employees of the refinery (CSB 2015).

On August 6, 2012, the Richmond refinery suffered a disastrous accident due to a pipe rupture in the crude distillation unit. The accident originated from one of many process streams commonly called the “4 - side cut” leaving the Richmond refinery’s C- 1100 Crude unit atmospheric column.

Very flammable, high temperature light oil at the rate of 10,800 barrels per day was released in the course of the incident. The released light oil  kindled into flame approximately two minutes later (CSB 2015). The process flow diagram of the Richmond refinery crude unit and its associated up stream process is shown Figure in 8.2.

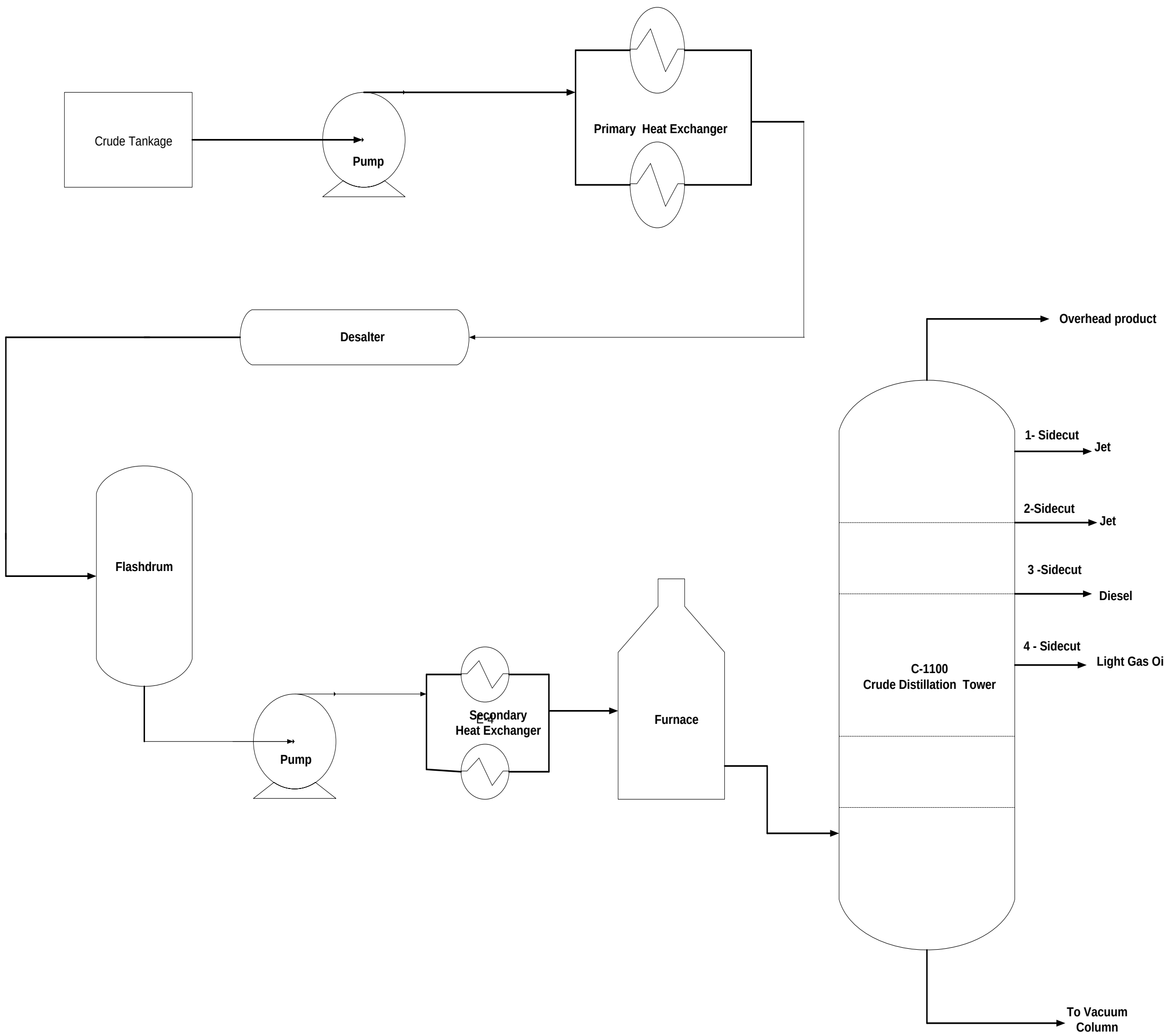



**Figure 8.2. The process flow diagram of Richmond refinery crude unit and it associated up stream equipment( Adedigba et al. 2016; CSB 2015).**

This incident was comprehensively investigated by the Chemical Safety Board (CSB). The CSB clearly emphasized several technical issues and safety system deficiencies that were responsible for the accident. Using the information made available by the CSB with the simulated operational process variables, the Richmond refinery accident is modeled using the proposed methodology.

Detailed hazards analysis of the Richmond refinery crude oil distillation unit (CDU) and associated accident scenarios has been previously undertaken by the authors. The relevant prevention barriers needed to be placed along the accident pathway have also been identified. A detailed description of these prevention barriers is given by (Adedigba et al. 2016).

Based on hazard analysis of the Richmond refinery, six consecutive safety barriers have been identified and placed in the accident path to inhibit or mitigate the effects of the accident. These barrier are: “Design error prevention barrier (DPB), Operational failure prevention barrier (OPB), Equipment failure prevention barrier (EPB), Human failure prevention barrier (HPB), Organisation failure prevention barrier (OrPB) and External factor prevention barrier (ExPB)” (Adedigba et al. 2016). Based on available information, different events that caused the failure of the safety or prevention barriers were adequately investigated using fault tree analysis

### 8.4.2. Results and discussions

The proposed methodology can be effectively demonstrated on a CDU unit. The CDU incident that nearly have total representativeness of the proposed model is the Richmond refinery accident. Due to lack of real life process monitoring data from the industry, the operation of CDU unit was simulated using Aspen HYSYS (8.8). The following key process variables were monitored: Feed pressure, feed temperature, reboiler temperature and condenser temperature. The process monitoring data generated were thoroughly analyzed using the proposed methodology. The organizational and technical causes of the accident were thoroughly analyzed using fault and event tree analysis. The authors assumed that Richmond refinery accident have environmental impact to effectively demonstrate this model on the accident.

The failure probability of prevention barriers using fault tree analysis of the Richmond refinery crude unit accident is shown Table 8.1. Fault tree construction is common; therefore, the failure probability of the basic event and associated fault trees are not displayed for all the prevention barriers. This minimized the numbers of figures and tables in this work.

**Table 8.1. Failure probability of prevention barriers ( Adedigba et al. 2016).**

| Prevention barrier | Failure probability |
|---|---|
| Design error prevention barrier (DPB) | 0.2567 |
| Operational failure prevention barrier (OPB) | 0.2700 |
| Equipment failure prevention barrier (EPB) | 0.2628 |
| Human failure prevention barrier (HPB) | 0.2870 |
| Organisational failure prevention barrier (OrPB) | 0.2959 |
| External factor prevention barrier (ExPB) | 0.0171 |

The states of historical operational data were used to build a probabilistic network structure among the main operational process variables, applying the Bayesian Tree Augmented Naïve Bayes (TAN) algorithm. This algorithm models the precise and concise probabilistic dependencies that exist among key operational process variables. The probabilistic structure generated using the Bayesian Tree Augmented Naïve Bayes (TAN) algorithm is used to detect faults and predict the time dependent probability of system deviation. The detailed methodology for the use of the Bayesian Tree Augmented Naïve Bayes (TAN) algorithm to generate a probabilistic network is clearly presented by (Adedigba et al. 2017b). To avoid replication of information, the detailed methodology for using the Bayesian Tree Augmented Naïve Bayes (TAN) algorithm to generate probabilistic network is not presented here. Following the methodology given by Adedigba et al. (2017b), the Bayeseian network structure of Figure 8.3 is generated. Also, the dynamic failure probability of deviation of the Richmond crude distilation unit is preicted using Figure 8. 3

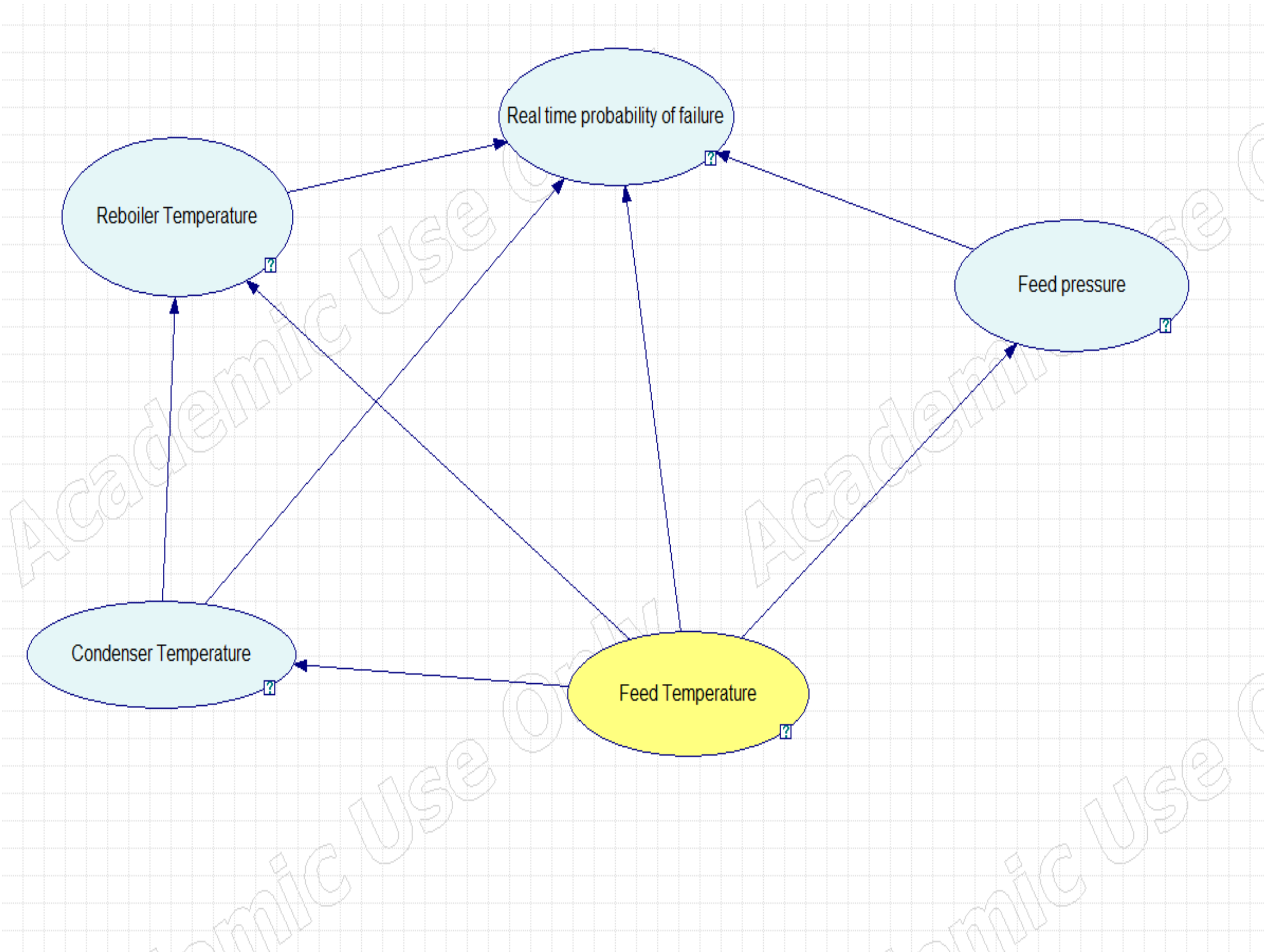



**Figure 8. 3. Probabilistic networks structure generated from operational historical data (Adedigba et al. 2017).**

The predicted real time dependent probability of devaiton of the Richmond CDU unit is given in Table 8. 2. The time dependent probability of deviation predicted by this methodology can be adopted to support automatic process control or activation of a layer of protections as soon as the predicted probability of deviation exceeds the acceptable range.The event tree for the case study is presented in Figure 8. 4. The values of real time probability of deviation given by Table 8.2 are used as an input to the event tree of Figure 8.4. The time dependent probability of deviation predicted by this methodology can be adopted to support automatic process control or activation of a layer of protections as soon as the predicted probability of deviation exceeds the acceptable range.

**Table 8. 2. Real time probability of deviation (Adedigba et al. 2017).**

| Time (minutes) | Real time probability of failure |
|---|---|
| 30 | 0.068 |
| 60 | 0.1333 |
| 90 | 0.028 |
| 120 | 0.039 |
| 150 | 0.022 |
| 180 | 0.049 |
| 210 | 0.012 |
| 240 | 0.019 |
| 270 | 0.018 |
| 300 | 0.01 |
| 330 | 0.015 |
| 360 | 0.044 |
| 390 | 0.014 |
| 420 | 0.028 |
| 450 | 0.057 |
| 480 | 0.061 |

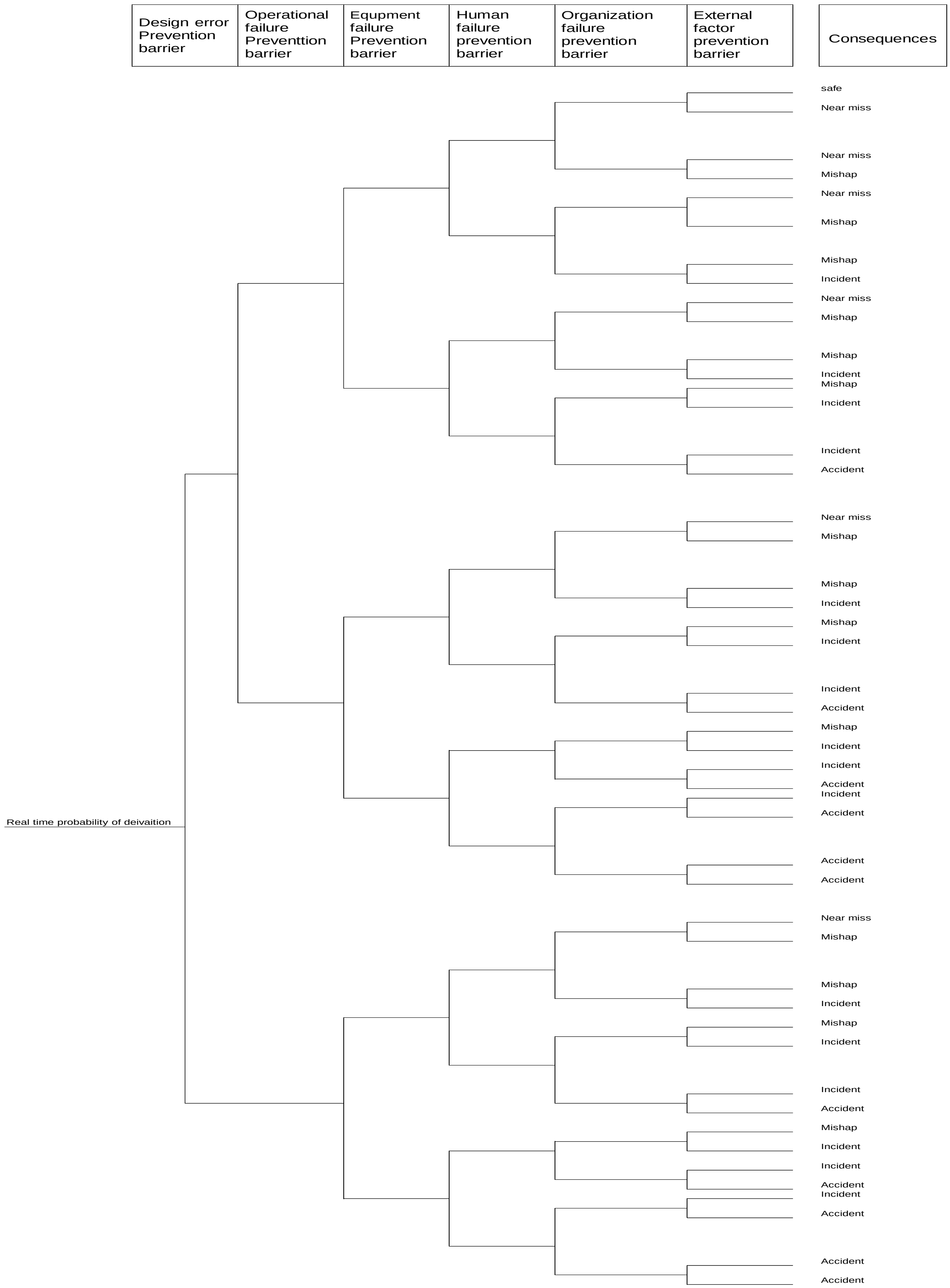



**Figure 8.4. Revised event tree for the proposed methodology ( Adedigba et al. 2016).**

Five principal end state consequences are considered in the event tree of Figure 8. 4. The real time occurrence probabilities of the end state consequences are given in Table 8.3.

**Table 8.3. Real-time occurrence probability of Consequence.**

| Time (Minutes) | Safe | Near miss | Mishap | Incident | Accident |
|---|---|---|---|---|---|
| 30 | 1.34E-02 | 2.74E-02 | 2.00E-02 | 6.38E-03 | 7.99E-04 |
| 60 | 2.63E-02 | 5.35E-02 | 3.92E-02 | 1.25E-02 | 1.53E-03 |
| 90 | 5.53E-03 | 1.13E-02 | 8.24E-03 | 2.63E-03 | 3.29E-04 |
| 120 | 7.70E-03 | 1.57E-02 | 1.15E-02 | 3.66E-03 | 4.58E-04 |
| 150 | 4.34E-03 | 8.86E-03 | 6.48E-03 | 2.06E-03 | 2.59E-04 |
| 180 | 9.67E-03 | 1.97E-02 | 1.44E-02 | 4.60E-03 | 5.76E-04 |
| 210 | 2.37E-03 | 4.83E-03 | 3.53E-03 | 1.13E-03 | 1.41E-04 |
| 240 | 3.75E-03 | 7.65E-03 | 5.59E-03 | 1.78E-03 | 2.23E-04 |
| 270 | 3.55E-03 | 7.25E-03 | 5.30E-03 | 1.69E-03 | 2.12E-04 |
| 300 | 1.97E-03 | 4.03E-03 | 2.94E-03 | 9.39E-04 | 1.18E-04 |
| 330 | 2.96E-03 | 6.04E-03 | 4.42E-03 | 1.41E-03 | 1.76E-04 |
| 360 | 8.68E-03 | 1.77E-02 | 1.30E-02 | 4.13E-03 | 5.17E-04 |
| 390 | 2.76E-03 | 5.64E-03 | 4.12E-03 | 1.31E-03 | 1.65E-04 |
| 420 | 5.53E-03 | 1.13E-02 | 8.24E-03 | 2.63E-03 | 3.29E-04 |
| 450 | 1.13E-02 | 2.29E-02 | 1.68E-02 | 5.35E-03 | 6.70E-04 |
| 480 | 1.20E-02 | 2.46E-02 | 1.80E-02 | 5.72E-03 | 7.17E-04 |

Different categories of loss are calculated independently using equations 8.12-8.20. The exact parameters used in predicting various category of loss are given in Table 8. 4. It is important to state that some of these parameters are estimated from available information provided in CSB report. The other parameters are sourced from expert opinion where the data is not available and various journals (Khan & Amyotte 2005; Zadakbar et al. 2015; Jafari et al. 2012; Yang et al. 2015; Yang & Barton 2015; CSB 2015).

**Table 8.4. Exact parameters used in estimating different categories of loss** (Khan & Amyotte 2005; Zadakbar et al. 2015; Jafari et al. 2012; Yang et al. 2015; Yang & Barton 2015; CSB 2015).

| No | Parameter | Assigned value |
|---|---|---|
| 1 | Likely downtime (hr) | 168 |
| 2 | Production value ($/hr) | 1,041666.7 |
| 3 | Asset density ($) | 500,000 |
| 4 | Damage radius (m) | 25.088 |
| 5 | Population density per metre square | 8 |
| 6 | Cost of fatality ($) | 1,000000 |
| 7 | Clean up cost of soil ($) | 480,000 |
| 8 | Clean up cost of water ($) | 12,000000 |
| 9 | Clean up cost of air ($) | 9,000000 |
| 10 | NFPA ranking of chemical (NH) | 2 |
| 11 | Fire explosion index (FEI) | 98 |

Table 8.5 presents the values for each category of loss. Estimated maximum loss is the sum of all the categories of loss. For the case study, the quality loss is $0.1million, the production loss is $175million, the asset loss is $12.5 million, the environmental loss is $21.5million and human health loss is $201million. Different end state consequences are considered with respect to their associated losses. Zero loss is associated with safe (normal operation). Quality loss, the cost of replacement and repair, is associated with a process near miss, which is assumed based on expert opinion. Production loss is associated with process mishap. Production, asset, and environmental losses are associated with the incident. For an accident, all categories of losses are assumed to be associated.

**Table 8. 5. Value of each category of loss.**

| **Categories of losses** | **Value of losses in (USD)** |
|---|---|
| Quality loss | $0.1Million |
| Production loss | $175 Million |
| Asset loss | $12.5 Million |
| Environmental loss | $21.5Million |
| Human health loss | $201 Million |
| | Estimated maximum loss = $ 401 Million |

Table 8. 6 shows the end state consequences of the event tree and the associated losses. The real time economic risk associated with the Richmond refinery Crude distillation unit are determined by combining the time dependent probabilities of end state consequence and their respective associated losses respectively.

This work assigns a univariate key process characteristic to the system. Sulfidation corrosion is responsible for the pipe rupture of the C-1100 Crude unit of the Richmond refinery. Sulfidation is a reaction that occurs at high temperature. Consequently, feed temperature is considered the critical variable in loss modelling. In a situation where there are several critical process variables to a system operation, this model can still be applied to accurately analyse the system risk. In this situation, principal component analysis (PCA) can be used to determine the process variable that most significantly affect process performance variation (Adedigba et al. 2017b). Multivariate loss function is used in this case to translate the process deviation into real time economic loss. Detailed information about multi variate loss functions is given by (Suhr & Batson 2001; Hsu 2001; Chan & Ibrahim 2004).

**Table 8. 6. End state consequences and associated losses. (Hashemi et al 2014).**

| End state consequences | Definition | Associated loss | Value of losses in (USD) |
|---|---|---|---|
| **Safe** | Normal operation | No loss | $0.000 |
| **Near miss** | Events that do not lead to an actual loss however has the possibility to cause the actual loss | Quality loss | $0.1Million |
| **Mishap** | Events that could result into slight asset and environment losses | Production loss | $175Million |
| **Incident** | Events that can result to significant harm or loss property, people and environment. | Production loss, asset loss, and<br><br>Environmental loss | $209Millon |
| **Accident** | An incident that lead to devastating consequences on people and asset. It normally receives publicity from national media | Production loss, asset loss, environmental loss and<br><br>human health loss | $410 Million |

This work assigns a univariate key process characteristic to the system. Sulfidation corrosion is responsible for the pipe rupture of the C-1100 Crude unit of the Richmond refinery. Sulfidation is a reaction that occurs at high temperature. Consequently, temperature is considered the critical variable in loss modelling.

The loss associated with temperature deviation in the CDU is estimated using the Modified inverted normal loss function and the estimated maximum loss (EML) alongside other parameters and user defined values given by Table 8.7. Figure 8.5 shows the CDU loss using modified inverted loss function.

The economic loss associated with feed temperature deviation in the Richmond crude distillation unit is estimated using the Modified inverted normal loss function and the estimated maximum loss (EML) alongside other parameters and user defined values given by Table 8.7.

**Table 8.7. Process information of CDU used to develop MINLF.**

| **Symbol** | **Description** | **Values** |
|---|---|---|
| T | The target temperature in the CDU | 350 $^0$C |
| $T_S$ | Set point for high temperature | 360 $^0$C |
| USL | Upper specified limit for high temperature in the CDU | 380 $^0$C |
| LSL | Lower specified limit for low temperature in the CDU | 330 $^0$C |
| $T_{max}$ | Maximum tolerable temperature in the CDU. It is presumed that the CDU fails catastrophically at this temperature. | 420 $^0$C |
| $T_{min}$ | Minimum tolerable temperature in the CDU | 310 $^0$C |
| $EML_{USL}$ | Loss connected with upper specified limit temperature in the CDU | 4000000 |
| $EML_{LSL}$ | Loss connected with lower specified limit temperature in the CDU | 2000000 |
| $EML_1$ | Estimated maximum loss as a result of high temperature in the CDU including all losses categories (quality ,asset, production, environmental cleanup and human health) | 410,000000 |
| $EML_2$ | Estimated maximum loss as a result of low temperature in the CDU | 82,00000 |

Figure 8.5 shows the Richmond CDU loss using modified inverted loss function. It show the losses related with the progression of the process deviation (high temperature) in Richmond refinery crude distillation unit. Figure 8.5 clearly shown that zero loss is associated with Richmond refinery crude distillation unit at 350 $^{o}$C. Beyond this temperature, the loss associated with the progression of the deviation of the CDU unit progressively increases.

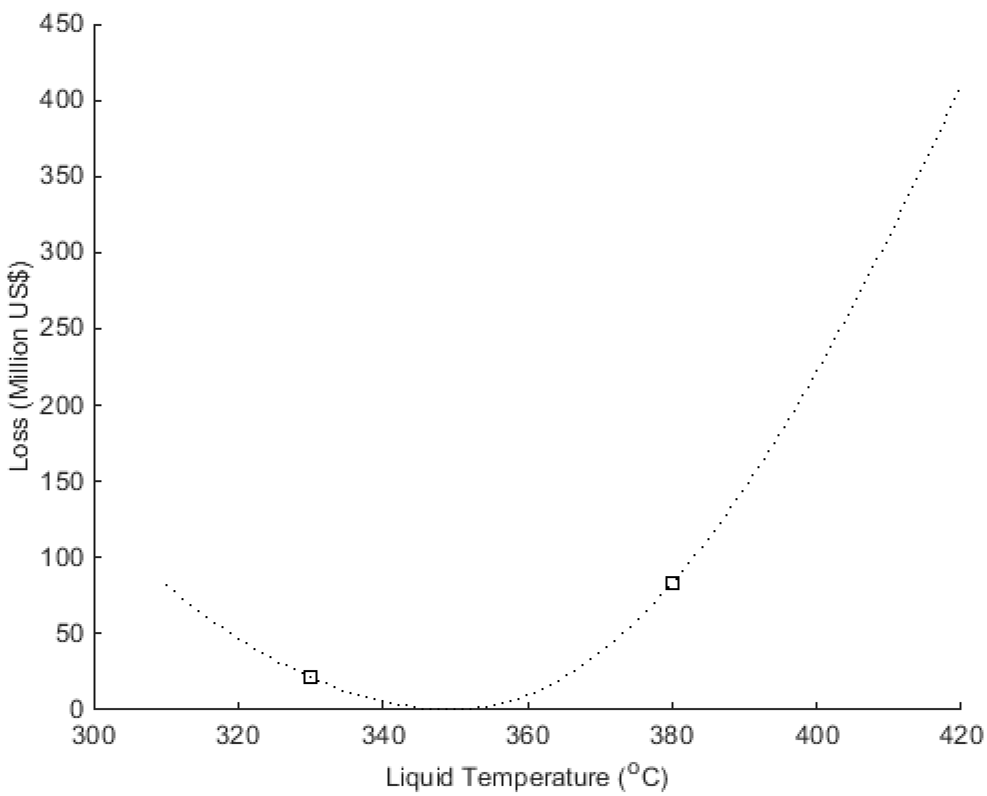



**Figure 8.5. CDU loss using MINLF.**

The real time risk associated with the incident and accident end state is shown in Figure 8.6. Figure 8.6b, shows that the economic risk at any instant due to deviations is beyond $10million. Also, different acceptable risk levels can be assigned to different end state events. The acceptable risk level for different end states can be used for decision making and the activation of layer of protections by the operators.

The real time economic risk (the combination of the consequences with associated losses) associated with the incident and accident end state for Richmond refinery crude distillation unit is shown in Figure 8.6

**(a)**


**(b)**

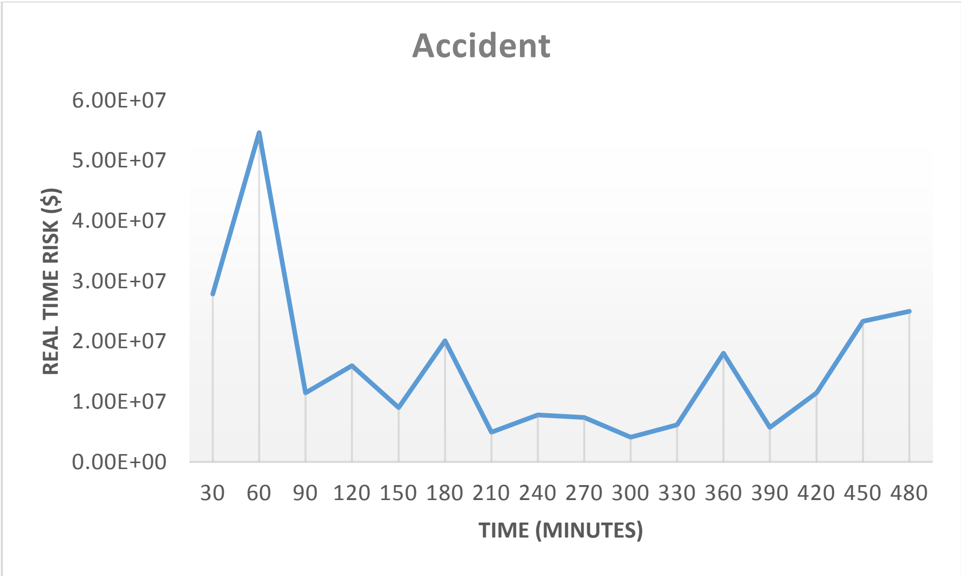



**Figure 8. 6. Real time risk of the Richmond refinery CDU.**

Figure 8.6b, shows that the economic risk at any instant due to deviations in Richmond refinery CDU is less than $10million. However, the risk association with the operation of Richmond refinery CDU at 60 minute is significantly high. This indicate that the probability of the system deviation is significantly high at 60 minutes. The Crude distillation unit need to be properly diagnosed to detect the root cause of the deviation at this time. The real time risk associated with the Richmond refinery CDU fluctuates progressively between the 90-480 minutes. This notified the operators of the abnormality of the CDU operation and also known the risk of the system as function of time as the operation progresses. Also, different acceptable risk levels can be assigned to different end state events. The acceptable risk level for different end states can be used for decision making and the activation of layer of protections by the operators.

Using the Richmond refinery accident to validate the proposed methodology has demonstrated that the proposed methodology has a strong capability to detect dynamic consequences of process variables' deviations from the optimum condition. This methodology provides risk profiles connected to various process end state consequences. This offers timely and valuable risk information for the operators and decision makers to decide when the operation of the plant will be shut down when the real time risk predicted exceeds the acceptable level. The methodology also provides an opportunity for root cause diagnosis of the deviations, since the dependability of the main key operational variables is used to predict the probability of deviation. The operator could take correct action by manipulating these variables with a basic control system to curtail the deviations. One obvious advantage of this developed methodology is that it dynamically captures the real time changes occurring in the process unit. The real time risk profile provided by the proposed methodology serves as a performance indicator for operational decision making. Each category of loss can also be updated whenever new information is available.

## 8.5. Conclusion

The economic risk model proposed in this work vividly relates economic risk (losses incurred) to main operational variables' deviation. It adequately computes economic risk connected with the probability of system deviation predicted using the performance variation of the characteristic system variables. The methodology proposed reflects dynamic variations of economic risk or loss for various kinds of abnormal events. The proposed methodology has the following features:

- The proposed methodology establishes the link between the process deviations with not only the probability estimation but also the potential loss prediction due to such deviations.
- Time dependent probability of system deviation is predicted based on exact representation of probabilistic dependencies among the process variables and makes use of both historical and present system operational data.
- The real time risk profile is predicted by monitoring the deviation of its key process variables.
- The methodology demonstrates the importance of using a loss function model to relate economic losses to process deviation.
- The real time economic consequences of process operation are predicted.

The methodology developed has a strong capability of generating a real time risk profile based on main operational process variables. Valuable performance information could be extracted from the risk profile generated. Also, the risk profile generated by the application of this methodology can serve as a performance indicator guide for daily operational decision making to prevent accidents, before the deviation devastatingly affects the system and the environment, resulting in great economic loss.

The proposed methodology is tested and verified with the Richmond refinery accident case study.

## Acknowledgements

The authors thankfully acknowledge the financial support provided by, the Natural Science and Engineering Council of Canada and Canada Research Chair (CRC) Program.

## References

Adedigba, S.A., Khan, F. & Yang, M., 2017a. Dynamic failure analysis of process systems using neural networks. *Process Safety and Environmental Protection*, 111, pp.529–543.

Adedigba, S.A., Khan, F. & Yang, M., 2017b. Dynamic Failure Analysis of Process Systems Using Principal Component Analysis and Bayesian Network. *Industrial & Engineering Chemistry Research*, 56(8), pp.2094–2106.

Adedigba, S.A., Khan, F. & Yang, M., 2016. Process accident model considering dependency among contributory factors. *Process Safety and Environmental Protection*, 102, pp.633–647.

Adedigba, S.A., Khan, F. & Yang, M., 2016. Dynamic safety analysis of process systems using nonlinear and non-sequential accident model. *Chemical Engineering Research and Design*, 111, pp.169–183.

Al-shanini, A., Ahamad, A. & Khan, F., 2014. Accident modelling and analysis in process industries. *Journal of Loss Prevention in the Process Industries*, 32, pp.319–334.

CCPS, 1999. *Guidelines for Consequence Analysis of Chemical Releases*, New York: AMerical Institute of Chemical Engineers.

Chan, W.M. & Ibrahim, R.N., 2004. Evaluating the quality level of a product with multiple quality characterisitcs. *The International Journal of Advanced Manufacturing Technology*, 24(9-10), pp.738–742.

Crowl, D.. A. & Louvar, J.F., 2011. *Chemical process safety:fundamentals with appplication* Third edit., Pearson Education.

CSB, 2015. *Final Investigation Report: Cheveron Richmond Refinery Pipe rupture and Fire*, Washington, D.C.

Durga Rao, K. et al., 2009. Dynamic fault tree analysis using Monte Carlo simulation in probabilistic safety assessment. *Reliability Engineering & System Safety*, 94(4), pp.872–883.

Fathi, Y. & Poonthanomsook, C., 2007. A Quartic Quality Loss Function and Its Properties. *Journal of Industrial and Systems Engineering*, 1(1), pp.8–22.

Ferdous, R. et al., 2009. Handling data uncertainties in event tree analysis. *Process Safety and Environmental Protection*, 87(5), pp.283–292.

Hashemi, S.J., Ahmed, S. & Khan, F., 2014. Loss functions and their applications in process safety assessment. *Process Safety Progress*, 33(3), pp.285–291.

Hashemi, S.J., Ahmed, S. & Khan, F.I., 2014. Risk-based operational performance analysis using loss functions. *Chemical Engineering Science*, 116, pp.99–108.

Heckerman, D., 1997. Bayesian Networks for Data Mining. *Data Mining and Knowledge Discovery*, 1(1), pp.79–119.

Hsu, C.-M., 2001. Solving multi-response problems through neural networks and principal component analysis. *Journal of the Chinese Institute of Industrial Engineers*, 18(5), pp.47–54.

Jafari, M., Zarei, M. & Movahhedi, M., 2012. The credit of fire and explosion index for risk assessment of iso-max unit in an oil refinery. *International Journal of*.

Jensen, F.. & Nielsen, T.., 2007. *Bayesian Networks and Decision Graphs* 2 edition., Springer, New York.

Kalantarnia, M., Khan, F. & Hawboldt, K., 2010. Modelling of BP Texas City refinery accident using dynamic risk assessment approach. *Process Safety and Environmental Protection*, 88(3), pp.191–199.

Khakzad, N., Khan, F. & Amyotte, P., 2013. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. *Process Safety and Environmental Protection*, 91(1-2), pp.46–53.

Khakzad, N., Khan, F. & Amyotte, P., 2011. Safety analysis in process facilities: Comparison of fault tree and Bayesian network approaches. *Reliability Engineering and System Safety*, 96(8), pp.925–932.

Khan, F., Hashemi, S.J., et al., 2016. Dynamic risk management: a contemporary approach to process safety management. *Current Opinion in Chemical Engineering*, 14, pp.9–17.

Khan, F., Rathnayaka, S. & Ahmed, S., 2015. Methods and models in process safety and risk management: Past, present and future. *Process Safety and Environmental Protection*, 98, pp.116–147.

Khan, F., Wang, H. & Yang, M., 2016. Application of loss functions in process economic risk assessment. *Chemical Engineering Research and Design*, 111, pp.371–386.

Khan, F.I. & Abbasi, S.A., 1998. Techniques and methodologies for risk analysis in chemical process industries. , 11, pp.261–277.

Khan, F.I. & Amyotte, P.R., 2005. I2SI: A comprehensive quantitative tool for inherent safety and cost evaluation. *Journal of Loss Prevention in the Process Industries*, 18(4), pp.310–326.

Khorsandi, J. & Aven, T., 2017. Incorporating assumption deviation risk in quantitative risk assessments: A semi-quantitative approach.

Leung, B.P.K. & Spiring, F. a., 2002. The inverted beta loss function: properties and applications. *IIE Transactions (Institute of Industrial Engineers)*, 34, pp.1101–1109.

Leung, B.P.K. & Spiring, F.A., 2004. Some Properties of the Family of Inverted Probability Loss Functions. *Quality Technology & Quantitative Management*, 1(1), pp.125–147.

Mackay, D.J.C., 1995. Probable networks and plausible predictions — a review of practical Bayesian methods for supervised neural networks. *Network: Computation in Neural Systems*, 6(3), pp.469–505.

Marsh Risk, C., 2011. *The 100 Largest Losses. Technical Report*, London.

Nivolianitou, Z.S., Leopoulos, V.N. & Konstantinidou, M., 2004. Comparison of techniques for accident scenario analysis in hazardous systems. *Journal of Loss Prevention in the Process Industries*, 17, pp.467–475.

Rathnayaka, S., Khan, F. & Amyotte, P., 2010. SHIPP methodology : Predictive accident modeling approach . Part II . Validation with case study. *Process Safety and Environmental Protection*, 89(2), pp.75–88.

Rathnayaka, S., Khan, F. & Amyotte, P., 2011. SHIPP methodology: Predictive accident modeling approach. Part I: Methodology and model description. *Process Safety and Environmental Protection*, 89(3), pp.151–164.

Spiring, F.A., 1993. The reflected normal loss function. *Canadian Journal of Statistics*, 21(3), pp.321–330.

Spiring, F.A. & Yeung, A.S., 1998. A general class of loss functions with industrial applications. *Journal of Quality Technology*, 30(2).

Suhr, R. & Batson, R.G., 2001. Constrained multivariate loss function minimization. *Quality Engineering*, 13(3), pp.475–483.

Sun, F.-B. et al., 1996. On Spiring’s Normal Loss Function On Spiring's norm. *Source: The Canadian Journal of Statistics / La Revue Canadienne de Statistique The Canadian Journal of Statistics La Revue Canadienne de Statistique*, 24(2), pp.241–249.

Taguchi, G., 1986. *Introduction to Quality Engineering: Designing Quality into products and Processes*, New York: Kraus.

Villa, V. et al., 2016. Towards dynamic risk analysis: A review of the risk assessment approach and its limitations in the chemical process industry. *Safety Science*, 89, pp.77–93.

Wang, H. et al., 2016. Dynamic quantitative operational risk assessment of chemical processes. *Chemical Engineering Science*, 142, pp.62–78.

Yang, M., Khan, F. & Amyotte, P., 2015. Operational risk assessment: A case of the Bhopal disaster. *Process Safety and Environmental Protection*, 97, pp.70–79.

Yang, Y. & Barton, P.I., 2015. Refinery Optimization Integrated with a Nonlinear Crude Distillation Unit Model. *IFAC-PapersOnLine*, 48(8), pp.205–210.

Zadakbar, O., Khan, F. & Imtiaz, S., 2015. Development of Economic Consequence Methodology for Process Risk Analysis. *Risk Analysis*, 35(4), pp.713–731.

# Chapter 9

# 9.0 Summary, Conclusions and Recommendations

## 9.1. Summary

This present work has demonstrated the use of a Bayesian network, Principal Component Analysis, artificial neural networks and loss function in the dynamic risk assessment of complex process systems. The majority of risk assessment techniques for process systems adopt traditional logic-based sequential techniques which are static and unable to cover an evolving process dynamic in real time. Dynamic risk analysis techniques for process operations have been developed to integrate advanced technique, addressing the knowledge gap and incorporating dynamic nature of the process.

This thesis presents an innovative predictive probabilistic model to assess hazardous processes operation accident likelihood. The model accounts for inter dependency of accident contributory factors within a safety barrier and also accounts for other contributory factors that have not been accounted for in the fault tree models of prevention barriers using both Noisy –OR and Leaky Noisy –OR gates. This serves as an effective tool to facilitate risk assessment and management of dynamic process operations.

Also, a new predictive nonlinear and non-sequential Bayesian network is introduced, based on the process accident causation model. This model effectively captures the non-sequential sequence of accident progression.

A novel integrated dynamic failure prediction model using principal component analysis (PCA) and the Bayesian TAN algorithm has been developed for process operation. This model is capable of predicting probabilistic relationships among process monitoring data and subsequently predicting and updating the risk profile dynamically, using real time process information. Similarly, an innovative ANN based model capable of predicting a risk profile empirically from process monitoring data has been proposed and tested as part of this work.

Finally, an integrated approach for dynamic economic risk assessment for a process system is developed. The model relates economic losses with process variable deviation. It helps to quantify economic loss under different abnormal conditions, which will be very useful for developing risk minimization strategies.

## 9.2. Conclusions

The major conclusions of this study are:

### 9.2.1. Development of an innovative predictive probabilistic model:

This study presented a new process accident model with emphasis on interdependency of contributory factors that lead to the failure of a particular prevention barrier. Six barriers were defined to prevent process accidents before they escalate into catastrophic events. The effectiveness of the proposed model was partially validated through the application of the model to the Richmond refinery accident. The BN model is capable of modelling the dependencies among these accident contributory factors. The application of Noisy-OR and leaky Noisy-OR gates helps to represent the uncertainties of the probabilities that are used in the CPTs of the BN model. Consequently, the proposed model is able to provide the lower and upper boundaries of the failure probability of a process accident. The accident model provides a mechanism for predicting a process accident based on the interdependency and nonlinear interaction of contributory factors. Process monitoring data is needed to effectively implement this accident model; with process monitoring data, the model can quantitatively estimate the dynamic risk profile that will greatly guide dynamic decision making. The use of predicted accident probabilities based on this model will help to take early corrective actions to prevent process accidents and developed an effective process safety management plan.

### 9.2.2. Development of a new predictive nonlinear and non-sequential Bayesian network based process accident model:

This present study has demonstrated the use of BN in modelling conditional dependencies among accident contributory factors within safety barriers and non-sequential failure of safety barriers which cause adverse events. This model highlights the importance of modelling the interdependency of accident contributory factors, nonlinear event sequences, and the selection of appropriate logic gates to the reduce uncertainty. BN is more appropriate to represent complex dependencies among prevention barriers and to include uncertainty in modelling. The BN has high a capability of abductive reasoning and the ability to handle uncertainty makes it a more appropriate technique for analyzing accidents. This accident model provides methodology for predicting a process accident based on nonlinear interactions within prevention barriers and non-sequential failure of prevention barriers which cause adverse events. Application of this method in models to predict accident occurrence probability will enable early remedial actions to prevent process accidents and consequently provide additional valuable information for process safety management.

### 9.2.3. Development of dynamic failure prediction model using Bayesian TAN algorithm:

This study has developed a risk assessment methodology based on the Bayesian TAN algorithm for safety analysis of the process system. The PCA-BN based process failure predictive model offers a technique for a predicting the real time failure probability profile of a process system. The model provides the capability to identify the key process variables that describe the most variance in process systems. The model is capable of predicting and assessing the real time risk of a process unit by monitoring the deviation of its main variables.

### 9.2.4. Development of dynamic failure prediction model using ANN:

This study has developed a failure prediction model for analyzing a process operation using an artificial neural network. Artificial neural networks are data driven nonlinear modelling techniques with the strong capability to model nonlinear relationships among process variables. In the study, ANN is used as a tool to define complex non-linear relationships among process variables

This study integrates ANNs with probabilistic analysis to model the process accident. Multi-layer perceptron (MLP) is used to define the relationship among process variables. The defined relationship is used to model a process accident considering logical and causal dependence of the variables. The predicted accident probability is subsequently used to estimate the risk of the process unit. A backward propagation technique is used to dynamically update the variable states and the risk accordingly. The updating mechanism of the approach makes the model adaptive and captures evolving process conditions.

### 9.2.5. Development of dynamic economic risk model for a process system:

This study developed a dynamic economic risk analysis methodology for a process system. The Bayesian Tree Augmented Naïve Bayes (TAN) algorithm is applied to model the precise and concise probabilistic dependencies that exist among key operational process variables, to detect faults and predict the time dependent probability of system deviation. The modified inverted normal loss function is used to relate system deviations in any given scenario to economic consequences (losses) from system deviation. The time dependent probability of system deviation due to abnormal event is continuously updated based on the current state of the characteristic variable of the process system.

## 9.3. Recommendations

This study has endeavored to introduce innovative concepts to the dynamic safety analysis of process systems. Nevertheless, many knowledge gaps and scope of the work could be further addressed. These include, though are not limited to:

- Techniques for capturing uncertainty in the safety analysis of a process system should be vigorously explored. Hierarchical Bayesian networks and fuzzy Bayesian network may be explored for dynamic safety analysis of the process system to minimize uncertainties in the input data.
- Model uncertainty should also be closely investigated. Uncertainty plays an important role and may affect results significantly. Advance data-driven models will help to better assess and manage model uncertainty.
- The performance of constraint and Bayesian score methods of learning the structure of a Bayesian network from historical data should be investigated for safety analysis purposes.
- The performance of different ANNs architectures for safety analysis of a process operation should be investigated and compare with other mathematical data driven model.
- Validation of the developed models with real industrial data is of paramount importance. The methods proposed in this thesis may be further tested and validated using real-life case studies.