

Imperial College London
Royal School of Mines
Department of Earth Science and Engineering

***A systems engineering approach to financial risk quantification in
primary raw materials production***

by

Indranil Pan

A thesis submitted for the degree of Doctor of Philosophy
to Imperial College London

October 2015

Abstract

Financial investments in large scale engineering sectors are subject to a variety of risks which affect the productivity and the profitability of the industry in the long run. Therefore, it is essential to quantify these risks and build mathematical models for analysis, so that better and more informed decisions can be made.

This PhD research work focuses on the financial risk modelling of raw materials production engineering facilities and applies the developed concepts and tools to mining and mineral processing. These sectors are specifically chosen as there are a lot of operational, environmental and other site specific hazards involved and at present there is no streamlined methodology for financial risk assessment. The approach developed in this research uses heuristic methods and expert inputs and is given a quantitative foundation using a robust modelling technique with an engineering basis, as opposed to the qualitative and subjective methods currently employed.

Current methods of estimating risk or forecasting future contingencies rely only on past historical data of failure and insurance claims. These often miss out on the finer details of the specific process and grossly over or under estimate the risks. Many other modelling techniques based on conservation laws employ mass balance and material flow through the whole process chain to obtain a more accurate model. However, the process systems data for mass and material flows through each component of the system is difficult to obtain and in many cases is a trade secret. Thus, it is difficult for an independent assessor or an external insurer to have a clear impression of the risk involved in the operation. The modelling techniques developed in this research aim to identify an optimal balance between these two extremes and

handle missing or inaccessible information in the process flow chain. As such, and beyond its academic value, the work is of significant practical value and especially useful for insurance underwriters, financial risk managers, business consultants and the raw materials production systems operators.

The thesis describes the development of a novel risk modelling tool based on a systems engineering approach. The risk model uses finite state machines with a set of logic for transitioning between the different states of the system. The flexibility of the systems risk modelling tool allows for features like incorporation of maintenance schedules, modelling of storage elements, redundancy schemes etc. which are implemented and validated with short case studies. These bespoke model refinements are useful for the process systems studied in this thesis.

The work presented next then describes how elements from queuing theory have been integrated with the proposed risk modelling methodology. This is especially useful for process systems which can be naturally modelled as a combination of servers and queues (e.g. transportation networks, automated manufacturing operations, computer networks etc.). A validation study is carried out to show the effectiveness of the proposed methodology.

The systems based risk modelling methodology, is augmented with fuzzy logic for incorporating expert knowledge and handling the uncertainties in the model parameters. This is of practical use as it allows for risk quantification in cases where only qualitative inputs are available from an external risk assessor.

Next, a Bayesian network based formalism is coupled with the systems level risk model for incorporating the effects of catastrophic risks. This allows for situations where information for the various parameters are partially available. Validation studies are conducted on both the fuzzy and Bayesian augmented risk frameworks to show the effective working of both methodologies.

Finally, the developed risk modelling tools are used to characterise the risk profiles of a mineral processing operation and various mining operations. The research would help investors and insurers in the mining and mineral processing sector to make appropriate investment decisions, by quantitative assessment of the risk factors, and hence minimise financial losses. It would also be useful for the management of the

process operation to identify the bottlenecks in the system and propose appropriate risk mitigation strategies for the same. To this end, a multi-objective optimisation is also done to obtain various Pareto optimal risk vs. reliability tradeoffs for the mineral processing operation. Since the modelling paradigm used is a generic one, it is expected that it can easily be adapted to other scenarios and would be of benefit for modelling financial risks in other industrial domains.

DEDICATED TO MY PARENTS

Affirmation

The work submitted in this thesis is my own, and has not been submitted previously for any other degree.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgements

I would like to thank my supervisors Professor Anna Korre and Professor Sevket Durucan for giving me the opportunity to conduct this research and for their constant support and supervision over the last three years. The discussions with them have been really useful for grasping the intricacies involved in the engineering process operations and coming up with methodologies to tackle the corresponding challenges.

This PhD was fully sponsored by Sciemus Ltd., London. I would like to thank Andre Finn and Neil Fleming from Sciemus Ltd., London, for their inputs in this research project. The innumerable discussions with Neil Fleming and Ashley Boyd-Lee of Sciemus Ltd. have been very insightful to understand the current needs and requirements of the industry.

Thanks to all my colleagues and friends who have made this stay at Imperial a very enjoyable one and have supported me through these years in so many different ways. I would like to thank Dr. Ji-Quan Shi, Dr. Zhenggang Nie, Dr. Rajesh Govindan, Dr. Amer Syed, Dr. Masoud Babaei, Dr. Giacomo Bacci, Dr. James Smith, Dr. Dong Chen, Dr. Julio Castro Molinare, Guangyao Si, Jorge de La Torre Guzman, Tareq Al-Ansari, Saba Manzoor and Nondas Adrianopoulos. Thank you all!

Finally, I would like to thank my parents for their inspiration and encouragement to pursue and complete my doctoral studies.

Table of Contents

ABSTRACT.....	II
AFFIRMATION.....	VI
COPYRIGHT DECLARATION	VII
ACKNOWLEDGEMENTS	VIII
TABLE OF CONTENTS	IX
LIST OF FIGURES.....	XIII
LIST OF TABLES.....	XVIII
LIST OF ABBREVIATIONS.....	XIX
CHAPTER 1 INTRODUCTION AND OBJECTIVES	1
1.1 INTRODUCTION	1
1.2 RESEARCH OBJECTIVES	2
1.3 THESIS STRUCTURE	3
CHAPTER 2 OVERVIEW OF FINANCIAL RISK FORECASTING AND APPLICATION	
AREAS	7
2.1 INTRODUCTION	7
2.2 DEFINITION OF RISK	7
2.3 MEASURES OF RISK	8
2.3.1 Conventional risk measures	8

2.3.2	<i>Downside risk measures</i>	9
2.4	MANAGING AND TRANSFERRING RISKS.....	11
2.5	MATHEMATICAL METHODS AND TECHNIQUES IN RISK MODELLING AND FORECASTING..	12
2.5.1	<i>Classical time series based risk forecasting techniques</i>	12
2.5.2	<i>Soft computing or computational intelligence based approaches</i>	13
2.5.3	<i>Reliability based risk modelling</i>	15
2.5.4	<i>Extreme Value Theory</i>	16
2.5.5	<i>Other methods</i>	16
2.6	FINANCIAL RISK MODELLING IN THE MINERALS INDUSTRY	17
2.6.1	<i>The mining lifecycle and the associated operations</i>	18
2.7	RISK CLASSIFICATION AND MODELLING IN MINING AND MINERAL PROCESSING SECTOR.	24
2.7.1	<i>Catastrophic natural or environmental risks and their modelling</i>	26
2.7.2	<i>Operational failures or intrinsic risks</i>	28
2.7.3	<i>Risks related to human errors or other management decisions</i>	29
2.8	CONCLUSION	30
 CHAPTER 3 OVERVIEW OF DESIGN PARADIGMS USED IN THE RISK MODEL		
	DEVELOPMENT	32
3.1	INTRODUCTION.....	32
3.2	CONCEPT OF SYSTEMS THINKING.....	32
3.3	FINITE STATE MACHINES	34
3.4	BASICS OF HYBRID DYNAMICAL SYSTEMS.....	35
3.5	DISCRETE EVENT SIMULATION TECHNIQUES.....	37
3.5.1	<i>Basics of queuing theory</i>	38
3.6	BASICS OF FUZZY LOGIC.....	41
3.7	OVERVIEW OF BAYESIAN NETWORKS.....	44
3.8	QUESTIONNAIRE BASED EXPERT APPROACH	46
3.9	BRIEF OUTLINE OF MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM FOR OPTIMISATION	46
3.9.1	<i>Concept of stochastic objective function and its optimisation</i>	49
3.10	CONCLUSION	50
 CHAPTER 4 A SYSTEMS APPROACH FOR RISK MODELLING		
4.1	INTRODUCTION.....	52
4.2	OVERVIEW OF THE PROPOSED SYSTEMS FRAMEWORK.....	52
4.3	MODELLING OF A HYPOTHETICAL RISK ITEM	57
4.3.1	<i>Validation of the model for a risk item with a single failure mode</i>	59
4.4	MODEL IMPROVEMENTS	62

4.4.1	<i>Modelling of a storage element.....</i>	62
4.4.2	<i>Modelling of redundancy schemes.....</i>	65
4.4.3	<i>Incorporating maintenance schedules.....</i>	67
4.4.4	<i>Modelling hybrid systems.....</i>	69
4.5	RELATIVE MERITS AND DE-MERITS OF THE MODELLING PARADIGM.....	72
4.6	CONCLUSIONS	74
CHAPTER 5	QUEUEING THEORY INTEGRATION FOR RISK MODELLING.....	75
5.1	INTRODUCTION	75
5.2	INTEGRATION OF SYSTEMS RISK MODEL WITH QUEUEING THEORY CONCEPTS.....	75
5.3	VALIDATION FOR A SINGLE RISK ITEM.....	77
5.3.1	<i>Case A: Service time is same as the entity generation time and risk item is absent.....</i>	77
5.3.2	<i>Case B: Service time is more than inter-generation time and risk item is absent.....</i>	79
5.3.3	<i>Case C: Risk item present.....</i>	80
5.4	CONCLUSIONS	82
CHAPTER 6	FUZZY LOGIC BASED METHOD TO INCORPORATE EXPERT OPINION IN RISK MODELLING.....	84
6.1	INTRODUCTION.....	84
6.2	NEED FOR UNCERTAINTY QUANTIFICATION OF MODEL PARAMETERS.....	84
6.2.1	<i>Fuzzy logic as a pre-processing tool of the risk model.....</i>	85
6.3	FUZZY INFERENCE SYSTEM FOR INCORPORATING EXPERT DOMAIN KNOWLEDGE	86
6.4	APPLICATION EXAMPLE.....	87
6.5	TESTING WITH A HYPOTHETICAL RISK ITEM	93
6.6	CONCLUSIONS	96
CHAPTER 7	BAYESIAN NETWORKS FOR RISK INFERENCE.....	98
7.1	INTRODUCTION	98
7.2	INTEGRATION OF BAYESIAN NETWORKS WITHIN THE SYSTEMS BASED RISK MODELLING PARADIGM.....	99
7.3	TESTING OF THE METHODOLOGY.....	101
7.4	DISCUSSIONS	107
7.5	CONCLUSIONS	108
CHAPTER 8	SYSTEMS RISK MODELLING CASE STUDY: COPPER PROCESSING AND METAL PRODUCTION OPERATION	109

8.1	INTRODUCTION	109
8.2	BRIEF OUTLINE OF THE COPPER PROCESSING AND METAL PRODUCTION UNIT	110
8.2.1	<i>Copper ore production and other mineral processing operations</i>	110
8.3	MODELLING OF FINANCIAL RISKS OWING TO OPERATIONAL HAZARDS IN COPPER PROCESSING AND METAL PRODUCTION.....	114
8.3.1	<i>Failure modes of the different components</i>	114
8.3.2	<i>Systems framework based risk model development</i>	116
8.3.3	<i>System characterisation and simulation results</i>	117
8.4	SYSTEM OPTIMISATION	120
8.4.1	<i>System description for optimisation and performance objectives</i>	121
8.4.2	<i>Optimisation results and discussions</i>	124
8.5	FUZZY LOGIC BASED MODELLING.....	128
8.6	CONCLUSION	131
CHAPTER 9 SYSTEMS RISK MODELLING CASE STUDY: MINING OPERATIONS		132
9.1	INTRODUCTION	132
9.2	APPLICATION TO OPEN PIT MINING OPERATION	132
9.2.1	<i>Brief overview and system model</i>	132
9.2.2	<i>Simulation results</i>	135
9.3	PROCESS MODELS OF DIFFERENT UNDERGROUND MINING OPERATIONS	143
9.4	RISK MODEL APPLICATION FOR AN UNDERGROUND COAL MINING OPERATION	145
9.4.1	<i>Brief overview of the risk model</i>	145
9.4.2	<i>Simulation results</i>	150
9.5	RISK MODEL APPLICATION FOR AN UNDERGROUND HARD ROCK MINING OPERATION..	152
9.5.1	<i>Brief overview and system model</i>	152
9.5.2	<i>Simulation results</i>	152
9.6	CONCLUSIONS	154
CHAPTER 10 CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH		155
10.1	ACHIEVEMENTS AND CONCLUSIONS FROM THE RESEARCH.....	155
10.2	LIMITATIONS AND UNANSWERED QUESTIONS.....	158
10.3	RECOMMENDATIONS FOR FUTURE RESEARCH.....	160
APPENDIX		163
REFERENCES		178

List of Figures

Figure 1.1: Schematic showing the outline of the thesis.....	4
Figure 2.1: Shortfall risk measures (adapted from (Olivieri and Pitacco, 2010)).....	11
Figure 2.2: Possible risk management options and their probable impact (Olivieri and Pitacco, 2010).....	11
Figure 2.3: A hierarchical view of the processes, machinery and materials involved in the surface preparation phase (modified after LICYMIN, 2003)	20
Figure 2.4: A hierarchical view of the processes involved in the topsoil and overburden removal phase (modified after LICYMIN, 2003)	21
Figure 2.5: Schematic showing a hierarchical classification of the different mining methods (modified after LICYMIN, 2003)	21
Figure 2.6: Overview of the processes, machinery and materials involved in the ore excavation and the auxiliary operations phase (modified after LICYMIN, 2003)	22
Figure 2.7: A hierarchical view of the operations involved in the waste disposal phase for hard rock mining (surface and underground) (modified after LICYMIN, 2003).....	23
Figure 2.8: Overview of the processes and equipment involved in the reclamation phase (modified after LICYMIN, 2003).....	24
Figure 3.1: Concept of a systems thinking framework	33
Figure 3.2: State chart of a finite state machine.....	34
Figure 3.3: Schematic of a hybrid automaton	36

Figure 3.4: Schematic of continuous evolution of state variables with discrete switching instants.....	37
Figure 3.5: Basic queue-server configurations	38
Figure 3.6: Queue-server configurations with scheduling mechanisms	39
Figure 3.7: Queues with admission control	40
Figure 3.8: Time evolution of a priority queue	41
Figure 3.9: Series-parallel interconnections of servers and queues to model complex process systems	41
Figure 3.10: A fuzzy inference system	42
Figure 3.11: A Bayesian network structure for predicting the damage owing to an earthquake. Each of the nodes can take a set of finite values and have a conditional probability table associated with them. These are not shown in the schematic.	45
Figure 3.12: Concept of a stochastic objective function.....	49
Figure 4.1: Hierarchical scheme of the risk model	53
Figure 4.2: Components of each risk item	54
Figure 4.3: Modelling risks in the present framework that can affect multiple risk items.....	56
Figure 4.4: State chart schematic of a failure mode.....	57
Figure 4.5: State chart based schematic of Health-Loading logic	59
Figure 4.6: State chart schematic for flow controller logic.....	59
Figure 4.7: Simulation of a sample risk item with one failure mode. The simulation time-step is taken as one day. The granularity of the time-step can be set according to the application requirements.....	61
Figure 4.8: Schematic of the tank element.....	63
Figure 4.9: Validation of the tank model	64
Figure 4.10: Modelling a redundancy scheme	65
Figure 4.11: State chart schematic of supervisory logic for the redundancy scheme	66
Figure 4.12: Validation of the redundancy scheme modelling	66
Figure 4.13: Different maintenance strategies	68
Figure 4.14: Constant interval maintenance with a period of 50 days.....	69
Figure 4.15: Constant interval maintenance with a period of 20 days.....	69
Figure 4.16: (a) State chart schematic of a hybrid system and (b) implementation for simulation.....	70
Figure 4.17: Risk item with health modelled by differential equation and switching logic with long maintenance interval	71

Figure 4.18: Risk item with health modelled by differential equation and switching logic with short maintenance interval	72
Figure 5.1: Systems approach integrated with queuing theory for risk modeling	76
Figure 5.2: Temporal evolution of the queue statistics	78
Figure 5.3: Output of the entity counter after the gate	78
Figure 5.4: Data carried by each entity which comes out of the gate	78
Figure 5.5: Temporal evolution of the queue statistics	79
Figure 5.6: Output of the entity counter after the gate	80
Figure 5.7: Data carried by each entity which comes out of the gate	80
Figure 5.8: Output of the entity counter after the gate with health of the risk item superimposed	81
Figure 5.9: Data carried by each entity which comes out of the gate	81
Figure 5.10: Temporal evolution of the queue statistics	82
Figure 6.1: A systems level risk model augmented with fuzzy inference	86
Figure 6.2: Membership functions for Company rating	89
Figure 6.3: Fuzzy rule surface for TBF and OD	89
Figure 6.4: Box plots of maintenance and breakdown costs for 1,000 Monte Carlo simulations of different cases.....	95
Figure 6.5: Box plots of material output for 1,000 Monte Carlo simulations of different cases	96
Figure 7.1: A systems level risk model augmented with Bayesian networks.....	101
Figure 7.2: Case B: A BN with no evidential knowledge.....	103
Figure 7.3: Case C: A BN with evidence about high impact risks	103
Figure 7.4: Case D: A BN with evidence about low impact risks	103
Figure 7.5: Costs due to breakdown and environmental risks for the four cases ...	106
Figure 7.6: Stem plot of the costs of catastrophic risks for Case B	106
Figure 7.7: Stem plot of the costs of catastrophic risks for Case C	107
Figure 7.8: Stem plot of the costs of catastrophic risks for Case D	107
Figure 8.1: Schematic for a copper processing unit.....	111
Figure 8.2: Pyro-metallurgical smelting of copper (adapted from (Goonan, 2005))	113
Figure 8.3: Refined copper produced over a period of 1 year. The data is scaled up to represent higher system losses.....	118
Figure 8.4: Total cumulative costs owing to maintenance and breakdown and its individual components over time.....	118
Figure 8.5: One hundred Monte Carlo runs of the evolution of total MB cost over time	119

Figure 8.6: Histogram of the a) total refined Cu output and b) the total MB cost at the end of 1 year for 1,000 Monte Carlo simulations.	120
Figure 8.7: Solutions in the final generation of MOGA showing the Pareto front.	125
Figure 8.8: Cumulative losses in copper metal production for the three representative solutions on the Pareto front.....	126
Figure 8.9: Cumulative costs for the three representative solutions on the Pareto front.....	127
Figure 8.10: Production of copper metal over time for the three representative Pareto solutions	127
Figure 8.11: Amount of materials over time in the stockpile before the smelter ...	128
Figure 8.12: Box-and-whisker plots for 1,000 Monte Carlo runs of the three different cases with total refined Cu output and total MB cost respectively.....	130
Figure 8.13: Multi-objective optimisation results for the three cases.....	130
Figure 9.1: Systems representation of the copper ore production process system operation	133
Figure 9.2: Queuing theory based representation of the process operation coupled with the systems level risk models	134
Figure 9.3: Entity arrivals at different points in the process flow of area A.....	135
Figure 9.4: Entity arrivals and state of muckpile of area A	136
Figure 9.5: Temporal evolution of the final output of copper and total costs.....	137
Figure 9.6: 50 Monte Carlo simulations showing the evolution of the costs over time owing to the various failure modes considered	137
Figure 9.7: 50 Monte Carlo simulations showing the evolution of the copper production over time	138
Figure 9.8: Case P: Bayesian network with evidence about high impact slope collapse risks.....	138
Figure 9.9: Case Q: Bayesian network with evidence about low impact slope collapse risks.....	139
Figure 9.10: Case R: Bayesian network with evidence of high impact flood risks	139
Figure 9.11: Case S: Bayesian network with evidence of low impact flood risks..	140
Figure 9.12: Box plots of annual ore output for all the different cases in 10,000 Monte Carlo evaluations.....	141
Figure 9.13: Box plots of losses for all the different cases in 10,000 Monte Carlo evaluations.....	141
Figure 9.14: Histogram showing the ore output of the extreme events (outliers) for each case	142
Figure 9.15: Histogram showing the financial losses of the extreme events (outliers) in each case	142

Figure 9.16: Process model for an underground longwall coal mining operation..	144
Figure 9.17: Process model for underground hard rock room & pillar and open stoping operation.....	144
Figure 9.18: Risk quantification modelling for underground mining operations leveraging on questionnaire based risk rating approach.....	146
Figure 9.19: Priors for the loss distribution based catastrophic risk model.....	147
Figure 9.20: Histograms of the individual and the composite risk modes as can be obtained from the risk model	149
Figure 9.21: Histograms of annual production of coal from 10,000 Monte Carlo simulations for both cases A and B	151
Figure 9.22: Box plots of total costs (losses) obtained from 10,000 Monte Carlo simulations for both cases A and B	151
Figure 9.23: Histograms of annual production of ore (Zn-Pb) from 10,000 Monte Carlo simulations for both cases C and D.....	153
Figure 9.24: Box plots of total costs (losses) obtained from 10,000 Monte Carlo simulations for both cases C and D	154

List of Tables

Table 6.1: Fuzzy rule base used in the present case.....	88
Table 6.2: Different cases with inputs and outputs from the fuzzy inference system.....	94
Table 6.3: Statistics of 1,000 Monte Carlo simulations of the different cases.....	95
Table 7.1: Probabilities of different impacts expressed as a percentage for the various cases.....	105
Table 8.1: Mean and variance of the important outputs from the model	120
Table 8.2: Representative solutions on the Pareto front.....	125
Table 8.3: Cases with expert inputs about the plant and their outputs from the fuzzy system.....	129
Table 9.1: Probabilities of different impacts of slope collapse as obtained from the Bayesian network	139
Table 9.2: Probabilities of different impacts of flood risk as obtained from the Bayesian network	140
Table 9.3: Different cases of catastrophic and management risk indices simulated for the underground coal mining risk case	150
Table 9.4: Different cases simulated for the underground hard rock mining model	153

List of Abbreviations

AAM: Area Aggregate Method
APARCH: Asymmetric Power GARCH
ARCH: Autoregressive Conditional Heteroskedasticity
BN: Bayesian Network
CV: Coefficient of Variation
DBSIM: Distributed Basin SIMulator
DES: Discrete Event Simulation
EA: Evolutionary Algorithms
EGARCH: Exponential GARCH
EMWA: Exponentially Weighted Moving Average
EVT: Extreme Value Theory
FIS: Fuzzy Inference System
FMECA: Failure Mode Effect and Criticality Analysis
FSM: Finite State Machine
FTM: Failure Time Modelling
GA: Genetic Algorithm
GARCH: Generalised Autoregressive Conditional Heteroskedasticity
GIS: Geographic Information Systems
GJR-GARCH: Glosten-Jagannathan-Runkle GARCH
GP: Genetic Programming

HAZOP: HAZard and OPerability
HAZSCAN: HAZardous SCenario Analysis
HLL: Health-Load Logic
HRA: Human Reliability Analysis
LICYMIN: Life CYcle assessment of MINing projects for waste minimisation and long term control of rehabilitated sites
LT: Lead Time
MB: Maintenance & Breakdown
MC: Monte Carlo
MCMC: Markov Chain Monte Carlo
MGARCH: Multivariate GARCH
MIMO: Multiple Inputs and Multiple Outputs
MINLP: Mixed Integer Non Linear Programming
MOEA/D: Multi Objective Evolutionary Algorithm with Decomposition
MTBF: Mean Time Between Failures
MV: Mean Value
NPGA: Niched-Pareto Genetic Algorithm
NSGA II: Non-dominated Sorting Genetic Algorithm II
NP: Nondeterministic Polynomial time
OD: Outage Duration
PAES: Pareto Archived Evolution Strategy
ppm: parts per million
QRA: Quantitative Risk Assessment
RS: Response Surface
SAG: Semi Autogenous Grinding
semiVar: semi-variance
SHE: Systeme Hydrologique Europeen
SPEA: Strength Pareto Evolutionary Algorithm
SWIFT: Structured What-If Technique
TBF: Time Between Failures
TOPMODEL: a TOPography based hydrological MODEL
TVaR: Tail Value at Risk
VaR: Value at Risk
VMR: Variance-to-Mean Ratio

Chapter 1 Introduction and Objectives

1.1 *Introduction*

Any large scale engineering operation is fraught with diverse risks which can disrupt the smooth operation of the business and result in significant monetary losses. For example, catastrophic events like fire, flood etc. can cause huge losses to an engineering production system. Many of these events might be severe enough to disrupt operation or altogether result in the closure of the business. Other small events like mechanical failures and breakdowns can result in lowered production rates and the business may fail to meet planned production targets. This might not only result in revenue losses, but also lower the credibility and reputation of the business operator for not being able to deliver promised goods to the downstream market in time. To counter the ill effects of these unforeseen risks, a company insures their operation especially for significant and catastrophic events. The idea is to pay a premium every month to the insurance company, and the insurance company then compensates for some or all of the losses in the unfortunate case of business interruption.

However, even though there are significant financial implications, most of the risk assessment techniques for these engineering systems are heuristic in nature and quantitative methods are difficult to come across. One of the largest impediments for

this is that the risk can depend on many external factors which are difficult to quantify in a model. However, relying purely on expert judgement is not good either, since the decisions suffer from personal biases, temperament of the person on a given day etc. Even though humans are capable of intuitive insights, assimilating diverse information to come up with a right decision every time is difficult. So, it is important to encode these intuitive insights into a set of logical algorithmic modules to process the data and obtain consistent numbers each time.

Most financial risk modelling is based on statistical models of the insurance claims data (Embrechts et al. 1997). Say for example, a company which wants to insure a specific business operation from natural catastrophes looks at the historical data of claims related to this and constructs a statistical probability distribution function. The claims data can be at a global level (say across all geographical locations) if there is sparse data, or it can be for a given location, if lot of past claims data is available for that region. Based on this, a probability of the event actually occurring is calculated and then a price is set for the premium and the terms and conditions for the payout is agreed upon. However, this kind of statistical modelling does not take into account the underlying processes which govern the dynamics of the business interruption. If a modelling tool can be made which reflects the uncertainties in the whole process of operation of a complex business chain or an engineering production system, then a more comprehensive idea of the risks involved can be obtained. Such a tool would help the insurers to get an idea of the overall risk profile of the business. On the other hand, the management which runs the business can get an idea of the root causes which give rise to specific risks and can take appropriate counter measures or decide to transfer the risks to a third party. This value of ascertaining the business risks from both sides is a key driver for the PhD research. The focus is to develop a generic tool and tailor it to the specific case of modelling the insurance risks for the mining and the mineral processing sectors.

1.2 Research Objectives

The main objectives of the PhD research were to

- Develop a new risk modelling tool which can be of practical use to insurers, risk managers or other professionals in the field of risk analysis. The tool must be generic in nature and should be adaptable to any large scale

engineering system with slight modifications. The tool must be able to handle partially available qualitative information as obtained from field surveyors or underwriters.

- Incorporate computational intelligence based methodologies allowing to use ‘soft’-information like expert opinion in order to arrive at a more realistic risk profile.
- Validate the tool by modelling the risk for the mining and mineral processing industry.

In order to achieve these objectives, a systems engineering approach was adopted for modelling any generic engineering process risk. Each component or equipment in the engineering process flow was abstracted as a risk item with specific inflows and outflows. A set of logic mechanisms regulate the quantity of outflows and the logical operations are dependent on the reliability of the individual risk items. An array of such individual risk items is connected in series and parallel configurations to represent the overall process flow of the engineering system. The risk profile is then characterised by simulating the composite system multiple times and obtaining a probability distribution over the losses. The underlying idea behind adopting such a strategy is that the methodology is flexible to handle multiple unknown variables in the process chain and also the model can be easily characterised from the reliability information of each component of the process chain. Also the open architecture of the model is useful for extending it with different features which can tackle specific issues. This has been leveraged in the present study, by augmenting the basic risk model with fuzzy logic to incorporate expert opinion in the risk modelling framework. A Bayesian network based approach has also been used to augment the model to predict different levels of catastrophic risks that the operation is exposed to. Finally the risk modelling methodology was applied to characterise the risk profile for the case of various open-pit and underground mining operations and a mineral processing operation (copper processing).

1.3 Thesis structure

The outline of the thesis with the various dependencies between the chapters is presented in Figure 1.1.

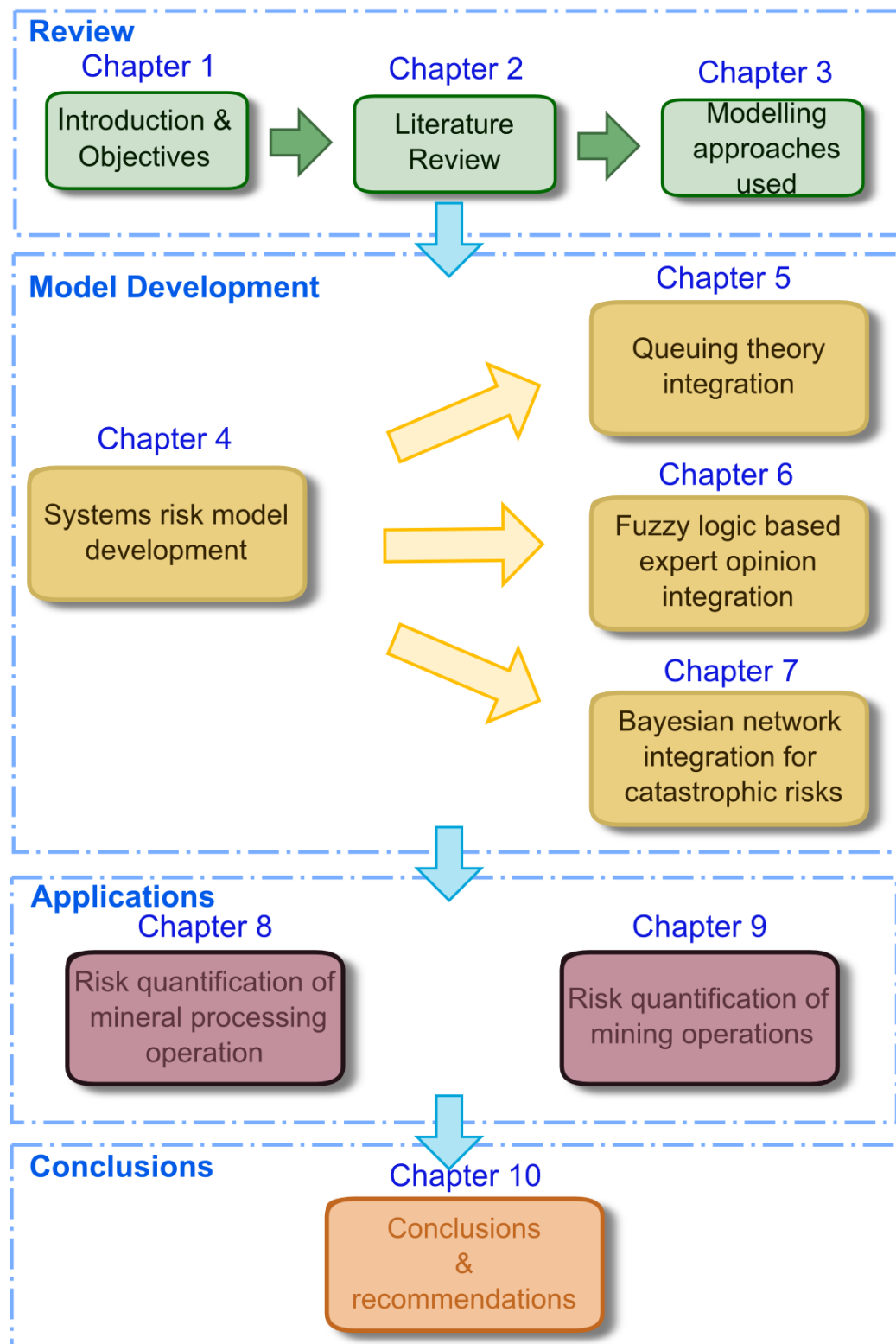


Figure 1.1: Schematic showing the outline of the thesis

Chapter 2 introduces the quantitative notion of risk and outlines the various statistical measures which are commonly used in risk quantification. A detailed literature

review of the various mathematical methods in financial risk modelling and forecasting for different applications is presented in this chapter. The sparse literature on financial risk modelling for engineering and chemical process systems operations, with a specific focus on the mining and the mineral processing industries is also discussed.

Chapter 3 briefly describes the various modelling paradigms used in this work for the development of a novel risk modelling tool along with brief implementations illustrating how these different paradigms address different issues (e.g. model flexibility, parameter uncertainty, expert knowledge etc.) and contribute to different aspects of the model development.

The next four chapters integrate different modelling approaches with this systems based risk framework and validate the models with theoretical case studies.

Chapter 4 proposes a novel risk modelling framework based on systems thinking. The relative merits and limitations of the proposed method and the need to adopt such a methodology for risk modelling of engineering process operations are discussed. Various model enhancements are made and each of them is validated with a theoretical example to prove the functionality of the methodology.

Chapter 5 integrates the concepts from queuing theory with the systems based risk model which was developed in the previous chapter. The need for the queueing theory based paradigm is explained and the model is validated for a theoretical case of a single risk item.

Chapter 6 couples a fuzzy logic based inference mechanism with the systems risk model to incorporate expert domain knowledge for dealing with uncertainty in the model parameters. A validation case study is conducted to show the feasibility of the proposed approach.

Chapter 7 couples the systems based risk modelling approach with Bayesian networks for modelling catastrophic risks from prior information. A model validation is conducted to illustrate the working of the proposed approach.

Chapter 8 applies the systems based risk modelling framework to a copper ore processing and metal production operation. The process system is abstracted into different risk items with inflows and outflows and interconnections between the items. The failure modes of the system are characterised and the system is simulated

multiple times to obtain a risk profile for the overall operation. A multi-objective system optimisation is carried out with different contradictory objectives and various Pareto optimal system configurations are proposed. The fuzzy logic based methodology, leveraging on expert opinion, is also used for dealing with the uncertainty in the model parameters.

Chapter 9 applies the developed modelling methodology for risk quantification of mining operations. The first section deals with an open-pit mining operation which uses the systems framework along with the queuing theory concepts for model development. The Bayesian network based approach is also used in this context to quantify the catastrophic risks. The next section deals with the risk quantification for underground hard and soft rock mining operations. The systems level model is augmented with loss distribution based catastrophic risk models for obtaining the risk profiles of the process operations. An expert questionnaire based approach is used which modulates the parameters of the model to reflect the level of risk inherent in the operation.

Finally, Chapter 10 presents the key contributions of the research and proposes avenues which might be explored in the future.

Chapter 2 Overview of financial risk forecasting and application areas

2.1 *Introduction*

Risk is intrinsically linked with any business operation. Depending on the nature of the operation, the term ‘risk’ can have different connotations. Therefore, it is important to specify the mathematical notion of risk as applicable in the present context. This chapter briefly introduces some of the quantitative definitions and measures of risk.

A wide variety of techniques have been developed to arrive at these quantitative risk measures. Most of them are application specific and each technique is better suited at addressing a specific issue than the others. No single technique is better than the others in all situations. This chapter reviews the different techniques and highlights their relative merits and limitations. A review of the different risk quantification techniques in the mining and the mineral processing sector is also carried out in subsequent sections.

2.2 *Definition of risk*

The concept of risk is often expressed contextually, and various definitions exist in common language, business jargon or other specialised fields. Mathematically speaking, a risk can be characterised by a random number, which can have a finite set of outcomes with some probability. However, the specific outcome of the random

number in each case is unknown. The quantitative measure of risk is the variability of the random number. This variability may be measured by different mathematical or statistical functions.

2.3 Measures of risk

Assume a random variable X , which represents the outcome of some particular event. Specifically, let X denote the risk such that $X < 0 \Leftrightarrow \text{Loss}$ and $X > 0 \Leftrightarrow \text{Profit}$. Depending on the causal effects of the underlying risks, the variable X can have a probability distribution. Assuming that X can be modelled by a probability density function (pdf), $f(x)$, we have

$$P[a \leq X \leq b] := \int_a^b f(x) dx \quad (2.1)$$

The expected value (μ) of X can be expressed as

$$\mu = E[X] := \int_{-\infty}^{+\infty} x f(x) dx \quad (2.2)$$

2.3.1 Conventional risk measures

The variance ($Var[X]$) and the standard deviation ($\sigma[X]$) of X are traditionally used as measures to quantify the risk associated with X .

$$Var[X] := E[(X - \mu)^2] := \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx \quad (2.3)$$

$$\sigma[X] := \sqrt{Var[X]} \quad (2.4)$$

Both of these measures of risk are symmetric, in the sense that they reflect both the positive and the negative aberrations from the expected value. There are other relative measures of risk like the variance-to-mean ratio ($VMR[X]$), coefficient of variation ($CV[X]$) etc. which are defined next.

$$VMR[X] := \frac{Var[X]}{E[X]} \quad (2.5)$$

$$CV[X] := \frac{\sigma[X]}{E[X]} \quad (2.6)$$

In the context of risk management and insurance, the coefficient of variation is known as the risk index. It is dimensionless, as both the numerator and the denominator are expressed in the same units.

2.3.2 Downside risk measures

The conventional risk measures capture the variability of X for both positive and negative deviations. However, in many situations, only the negative deviations, or the loss part of the risk need to be captured. These are characterised by the downside risk measures.

The semi-variance can be used as a downside risk measure. It characterises the negative deviations from the expected value (μ) and in essence reflects the probable outcomes of X which are below the expected value.

$$semiVar[X] := E\left[\left(\min\{X - \mu, 0\}\right)^2\right] := \int_{-\infty}^{\mu} (x - \mu)^2 f(x) dx \quad (2.7)$$

The semi-standard deviation is given by

$$semi\sigma[X] := \sqrt{semiVar[X]} \quad (2.8)$$

This concept of semi-variance can be generalised by replacing μ with any arbitrarily chosen value (τ) and considering any generic power k , instead of the square of the deviations. Then the lower partial moment of degree k can be defined as

$$LPM_{\tau}^k[X] := E\left[\left(\min\{X - \tau, 0\}\right)^k\right] := \int_{-\infty}^{\tau} (x - \tau)^k f(x) dx \quad (2.9)$$

In a certain sense, τ characterises the concept of shortfall risk measures. Where X denotes a financial outcome, τ can be chosen to be negative and $X < \tau$ would then denote the ‘tail loss’, which is greater than the chosen target loss. The shortfall probability is then given by

$$P[X < \tau] := \int_{-\infty}^{\tau} f(x) dx \quad (2.10)$$

The expected shortfall is the expected value of the loss exceeding the target, given that the loss exceeds the target itself.

$$ES_{\tau}[X] := E[X | X < \tau] := \frac{\int_{-\infty}^{\tau} xf(x)dx}{P[X < \tau]} \quad (2.11)$$

The downside risk measures, in a sense, indicate the amount of capital that must be set aside so as to cover up for any potential losses that might be incurred due to a financial transaction having some uncertainty in the result. This might need to be done to prevent the loss from jeopardising the other aspects of business. The Value at Risk (VaR) is thus defined as the (negative) amount VaR_{α} such that

$$P[X \leq VaR_{\alpha}] := \alpha \quad (2.12)$$

where α is a heuristically chosen low probability. The amount VaR_{α} can be comprehended as the maximum loss in case an extreme event (tail event) does not occur. If the capital corresponding to VaR_{α} is set aside, then all non-tail losses are completely funded. However, the knowledge of VaR_{α} itself does not ascribe a loss value in the unfortunate event of a tail loss. The Tail Value at Risk (TVaR) or the Conditional Tail Expectation is given by the (negative) amount $TVaR_{\alpha}$ as

$$TVaR_{\alpha} := E[X | X < VaR_{\alpha}] := \frac{\int_{-\infty}^{VaR_{\alpha}} xf(x)dx}{P[X < VaR_{\alpha}]} \quad (2.13)$$

$TVaR_{\alpha}$ is the expected shortfall to the target VaR_{α} and a capital allocation of $-TVaR_{\alpha}$ covers, at least partially, a loss caused due to a tail event. Figure 2.1 shows a schematic for the various shortfall risk measures as described above.

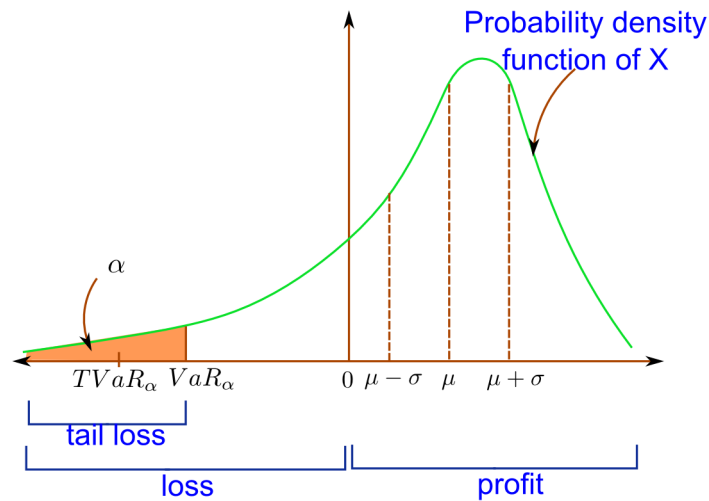


Figure 2.1: Shortfall risk measures (adapted from (Olivieri and Pitacco, 2010))

2.4 Managing and Transferring risks

The broad discipline of risk management aims to identify the risks associated with a particular firm and gives possible remediation measures for risk amelioration. These remediation measures might include loss financing through retention, insurance, hedging or other contract based risk transfers. Internal risk reduction measures like diversification or investment in information might also be applied. Figure 2.2 shows a frequency vs. severity based framework which is commonly employed to understand what risks can be retained and what can be transferred to a third party (e.g. an insurance company).

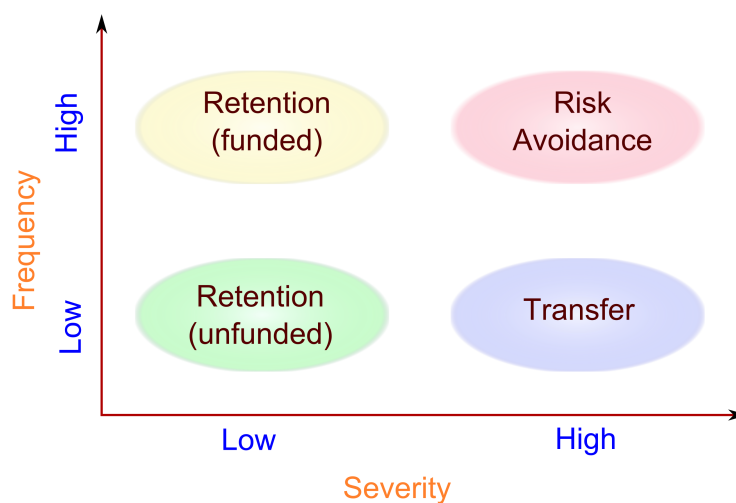


Figure 2.2: Possible risk management options and their probable impact (Olivieri and Pitacco, 2010)

Thus, if certain classes of risks are low in both frequency and severity, i.e. the number of occurrences of losses and the amount of money involved in the loss are both low, then these risks can be retained by the company and no specific contingency arrangement is required to be made. On the other hand, if these low severity risks recur a significant number of times, then the risk can be retained but a specific fund can be pre allocated for meeting these losses. This fund can be internally financed by current cash flows or share-holder's capital etc. or they might be externally financed by borrowing. For low frequency events which have a high monetary impact, it is best to transfer the risk through some contractual agreement with a third party like an insurance company. Risks which are high impact and have a high frequency should be avoided by reducing the level of risky actions. These actions might lead to a reduction in expected profits, but are necessary to prevent huge losses and bankruptcy.

2.5 Mathematical methods and techniques in risk modelling and forecasting

Forecasting in a risk context refers to the prediction of the expected outcome of a random variable. Data driven techniques are commonly used for quantitative forecasting of volatility. The jargon volatility is used in the genre of finance to represent the standard deviation of the returns and is a measure of the risk. The various classical risk modelling techniques are briefly outlined in the following section. This is followed by other approaches towards risk modelling and prediction.

2.5.1 Classical time series based risk forecasting techniques

Classical time series based modelling techniques have been used widely in the field of empirical finance to forecast volatility. Various applications in financial derivatives, quantitative trading and risk modelling have augmented their importance in a wider sphere. Univariate volatility modelling refers to the volatility forecasting of a single time series. Multivariate volatility modelling on the other hand refers to forecasting of the whole covariance matrix of the asset returns (Danielsson, 2010). The simplest volatility models are the moving average (MA) ones which calculate the standard error from a sample of returns and predict the next possible outcome. This performs badly in several cases and can be improved by assigning more weightage to the most recent observation than the previous ones. The exponentially

weighted moving average model (EMWA) (Holt, 2004) does this and is the most expedient among this class of models. The other advanced models that are used in present day quantitative prediction are the GARCH (Generalised Autoregressive Conditional Heteroskedasticity) family of models. These models are a subset of the more generalised conditional volatility models and use optimal exponential weighting of previous returns to realise a volatility forecast. The parameters of the model are estimated using maximum likelihood and the returns at a particular time are a function of the prior returns with heavy weightages assigned to the more recent returns than the previous ones. The ARCH (Autoregressive Conditional Heteroskedasticity) model was introduced by Engle (Engle, 1982) and was extended to its generalised version of GARCH by Bollerslev (Bollerslev, 1986). Numerous variations of the GARCH have been proposed to address specific characteristics of time series. A GARCH model that can model the leverage effect has been proposed by Glosten (Glosten et al., 1993) and is known as the GJR-GARCH or the threshold-GARCH. Another GARCH model which can incorporate the leverage effect has been proposed in Nelson (1991) and is known as the exponential GARCH (EGARCH). In cases where the absolute returns have stronger autocorrelations than squared returns, the power GARCH is used (Taylor, 1986). A more complicated model which can consolidate both the power effects and the leverage effects as proposed in Ding et al. (1993) is known as the asymmetric power GARCH or APARCH model. The multivariate GARCH (MGARCH) models are used in the study of volatility and co-volatility of numerous time series. A survey of the various submodels in this class of MGARCH type models are discussed by Bauwens et al.(2006). These concepts have also been extended to a family of models known as ARMAX (Autoregressive Moving Average with exogenous inputs) which can take into account other independent (or exogeneous) variables. These techniques are suitable for modelling risk if a large quantity of past historical data is available. However, for the problem studied here, it is not possible to get a large amount of data for extreme events and, hence, this method is generally not suitable.

2.5.2 Soft computing or computational intelligence based approaches

Soft computing techniques have been widely used in a varied number of disciplines like physics, biology, engineering and finance. These techniques differ from conventional computing in that they are able to tolerate imprecision, uncertainty and

vagueness in the problem. They are generally used for NP (non-deterministic polynomial time) complete problems for which no known solution exists in polynomial time. Thus these techniques are a viable alternative in many real world problems and give robust solutions at low computational cost. The main components of soft computing are fuzzy logic, artificial neural networks, evolutionary and swarm based computation, machine learning and probabilistic reasoning. These techniques complement each other and address a specific nature of the problem. Thus a synergism between these techniques is expedient in modelling real world scenarios. For example, neuro-fuzzy systems or Genetic Algorithm (GA) based fuzzy systems incorporate one or more of these techniques to harness the power of both domains. A discussion of how the synergism can help in the insurance industry is documented in Shapiro (2002).

The individual components of soft computing have also found applications in the insurance industry. Fuzzy logic, which deals with imprecision or vagueness in the data, has found applications in insurance classification, underwriting, projected liabilities, pricing, asset allocations etc. (Shapiro, 2004). Feed forward neural networks have been used as a pattern classification instrument in contractors risk assessment system (Bakheet, 1995). Multi-layer perceptrons have been used to classify standard and non-standard risks in life insurance (Vaughn et al., 1997). In Brockett et al. (1994), neural networks have been employed to ascertain the risks associated with insurer insolvency and an early warning system has been developed. A Kohonen self-organising map has been developed to detect fraudulent claims for risks associated with the automobile sector (Brockett et al., 1998). Neural networks have also been used for predicting insolvency of property and casualty insurers (Brockett et al., 1997). Neural networks designed through the help of Genetic Algorithms have been used in Park (1993) for predicting bankruptcy of financial companies. Modelling with neural networks can be useful in situations where large datasets are available. However, for the types of risk modelling as pursued in the present work, the lack of large datasets makes these techniques difficult to implement. Moreover, for small datasets neural networks suffer from the problem of overfitting and might not generalise well to new data.

Bankruptcy predictions using Genetic Algorithms have also been studied in Shin and Lee (2002). In other works, the GA was used to optimise the parameters of a support

vector machine for predicting the risk of bankruptcy in Wu et al. (2007). Analysis of insolvency risk using GA has been studied in Varetto (1998) and a Pareto frontier of the profitability and risk competitiveness using GA has been documented. GAs have also been coupled with fuzzy logic for financial risk management in Robinson and Yager (1996). Genetic Programming (GP) which is a variant of GA, evolves new programs which can efficiently solve a particular task. GP along with rough sets have been used in a hybrid manner to classify the risks associated with bankruptcy in McKee and Lensberg (2002). In general GA variants can be useful for automatically tweaking the parameters of a model for calibrating it with real data.

2.5.3 Reliability based risk modelling

Many reliability based analysis for component failure have been investigated in large scale complex systems. Modelling systems, based on the failure rates of individual components, can give an idea of the overall risk involved in the complete process chain. A survey of probabilistic methods in reliability, risk and uncertainty analysis has been reported in Robinson (1998). Most of the methods related to this category try to characterise the uncertainty in system response due to uncertainty in internal or external system parameters. The mean value (MV) method (Kapur and Lamberson, 1977) is widely adopted and uses a Taylor series expansion about the mean of a random variable to estimate the first and the second moments. The moments of the obtained approximate functions can then be calculated and these can give an idea regarding the probability of failure. The Response Surface (RS) method differs from the MV method in that it tries to approximate the original system response or performance by a simpler and computationally tractable polynomial model. The appropriate constants to the polynomial equation are fitted using linear regression. Both these methods are affected by the sensitivity to the expansion point and might grossly underestimate the failure probability. Also analysis becomes cumbersome if the random variables are not statistically independent.

Differential analysis methods for approximating performance functions and hence estimating risk and reliability has been documented by Zimmerman et al. (1990). Differential analysis can also be coupled with Monte Carlo techniques for reliability analysis. A comparison of this technique with the Latin hypercube sampling can be found in Iman and Helton (1988). There are various other first order reliability methods like Hasofer-Lind (Hasofer et al., 1973), Rackwitz-Fiessler (Rackwitz and

Flessler, 1978) etc. These methods compute a linear approximation to the performance function in an iterative fashion. Advanced first order reliability methods were proposed by Hohenbichler and Rackwitz (1981) to overcome the various deficiencies of the original first order reliability methods. Here any non-Gaussian dependent variables can be mapped to a set of independent variables using the Rosenblatt transformation (Rosenblatt, 1952). These reliability based methods are a viable alternative to risk modelling for the present case. Some of the basic concepts from this field are introduced into the proposed risk modelling paradigm.

2.5.4 Extreme Value Theory

Most of the risk modelling techniques like GARCH and other methods based on past historical data are appropriate for events that occur with probabilities of around 1-5%. However, for more rare catastrophic events whose probability of occurrence is of the order of $\sim 0.1\%$, the Extreme Value Theory (EVT) is used. The EVT focusses on the tail zones of the probability distributions and does not require a-priori specification of the response distribution. EVT has been used in many fields of engineering (Castillo, 1988), climate change (Katz, 1999), electricity markets (Byström, 2005) etc. and more recently in measuring financial risk (Gençay et al., 2003; Gilli and Këllezi, 2006). The applications of EVT as a risk management tool for insurance and finance have been documented in Embrechts et al. (1999). A comparison of the various EVT models for determining VaR have been investigated in Brooks et al. (2005) and the EVT based methods have been shown to give better results than GARCH based methods. The EVT methods are a good alternative for risk modelling and can be further delved into for possible application in the risk analysis model development.

2.5.5 Other methods

There are many other mathematical techniques which have been adapted for risk analysis and modelling depending on specific requirements of the problem domain. Bayesian techniques in risk management for insurance applications have been investigated in Abbaspour (1994 and Hardy (2002). Bayesian techniques make the best use of the available data, can be used on both small and big datasets and can incorporate subjective preferences of the experts in the modelling paradigm. Therefore these methods are suitable for the risk modelling problem pursued in this thesis. Markov chain models have been used widely for credit risks in (Jarrow et al.,

1997; Kijima and Komoribayashi, 1998; Kijima, 1998). These models assume that the systems present state at time ' t ' is only dependent on its previous state at time ' $t-1$ ' and therefore does not incorporate the effect of long range memory in the system. This is more of a mathematical convenience and might not be a true representation of the underlying system. Therefore this is not pursued in the present work. Markov Chain Monte Carlo (MCMC) methods have been used in risk modelling in Ades and Lu (2003), stochastic volatility modelling in Chib et al. (2002) and operational risk management in Dalla Valle (2009). MCMC is a sampling algorithm which helps in the evaluation of high dimensional integrals involved in Bayesian inference and facilitates their real world applications. Therefore their use in conjunction with Bayesian approaches can be used to exploit the full potential of these methods.

2.6 *Financial risk modelling in the minerals industry*

The focus of the present thesis is in the area of financial risk modelling for mineral processing and mining operations. The following sub-sections discuss the salient points in mining engineering, mineral processing operations and their risk classifications. It is important to know the process work flow and have a basic idea of the workings of various sub-units within the mining operations and the processing of minerals. The key concepts are succinctly presented in the following discussion and are of relevance to the development of the risk model for mining and mineral processing. The major mining operations along with the commonly used equipment and materials are presented in a modular structure. This classification is also important from the model development point of view as the systems level modelling tool developed in this research benefits from the nested subsystem representation of the entire mining operation. A short overview of the copper processing system is also presented as a sample mineral processing operation that has been used in the case study illustrated in this thesis. Copper is widely used in many engineering applications like cables and wires for electrical and electronic equipment, heating and air conditioning tubing, electroplating, automotive and machine parts etc. It is also one of the largest metals produced by volume. Therefore, copper production and extraction using surface mining methods is taken up as a case study in this thesis.

2.6.1 The mining lifecycle and the associated operations

The mining industry and its associated activities are engaged in the profitable extraction of mineral resources from the earth's crust. The actual mining operation is preceded by geological exploration for mineral prospecting. Successful discovery of mineral resources is followed up by economic feasibility studies. When the economic profitability of the mineral deposit is ascertained within sufficient confidence intervals, then the mineral deposit (which is generally referred to as ore in the case of metal mining) is mined. When the excavation of the mine is wholly on the surface, it is termed as a surface mining operation. Underground mining refers to those cases where underground operations are necessary to extract the minerals from ore deposits. The mining method refers to the technique, equipment and layout design used for various mining operations. Following the extraction of the ore, this is further concentrated in mineral processing operations. These may or may not be a part of the mineral production facility. The mineral processing produces high concentration minerals that can be further refined and marketed to industries which manufacture consumer goods.

The whole life cycle of the mine production can be divided into five major constituents which also refer to the five stages in the life of the mine. They are prospecting, exploration, development, exploitation and reclamation.

The first stage known as prospecting refers to the search for ores or mineral deposits. Both direct and indirect methods are employed for this as the minerals can be found at the surface or underground. Direct techniques rely on visual examination of the rock surface, geological maps, structural assessment of the area, microscopic and chemical studies of rocks etc. to ascertain probable locations for the mineral deposits and infer their characteristics from the mineralization of the rocks. Indirect techniques employ physical measurement data of the gravitational, seismic and electromagnetic signatures of the earth, employed at some distance from the deposit itself.

The exploration phase seeks to ascertain the size and volume of the mineral deposit. A wide array of measurements is taken to determine the extent and grade of the ore within a small uncertainty bound. Samples are collected from the entire spread of the ore body using techniques like drilling, tunnelling, trenching etc. Various tests like spectrography, X-ray spectroscopy, radiometry and other chemical and metallurgical

techniques are used to infer the grade of the ore. Diamond core drilling is preferred over rotary or percussion drilling for exploration since the cores obtained from these drills can be used to determine the geologic structure of the ore body. The cylindrical ores are split in two along their longitudinal axis, and only one half is used for analysis while the other half is stored away for further study and archival purposes. The determination of the richness and location of the mineral deposits helps in the calculation of the reserve, the exploitation method (production), the mining costs, environmental costs and the long term profitability of the mineral reserve. This process is called reserve estimation and based on its final conclusions, the project either goes forward for further development or it is relinquished. The exploration phase does not entail significant risks, is usually lengthy and preceeds the establishment of a mine production operation.

The development phase relates to the actions that must be performed for the opening of the mineral deposit before the exploitation is implemented. The top soil and the overburden must be removed to expose the near surface ore for mining and to gain access to the deposit. In the case of underground mining, sinking of shafts and access roads allows to approach the deposit. Therefore it is not considered further in the design of the risk modelling system developed in this research.

Furthermore, the development phase also involves site preparation, which entails construction of access roads, infrastructure etc. in preparation for the exploitation phase of the mine. Figure 2.3 shows a summary of the important processes, associated machinery and materials involved in the site preparation operation.

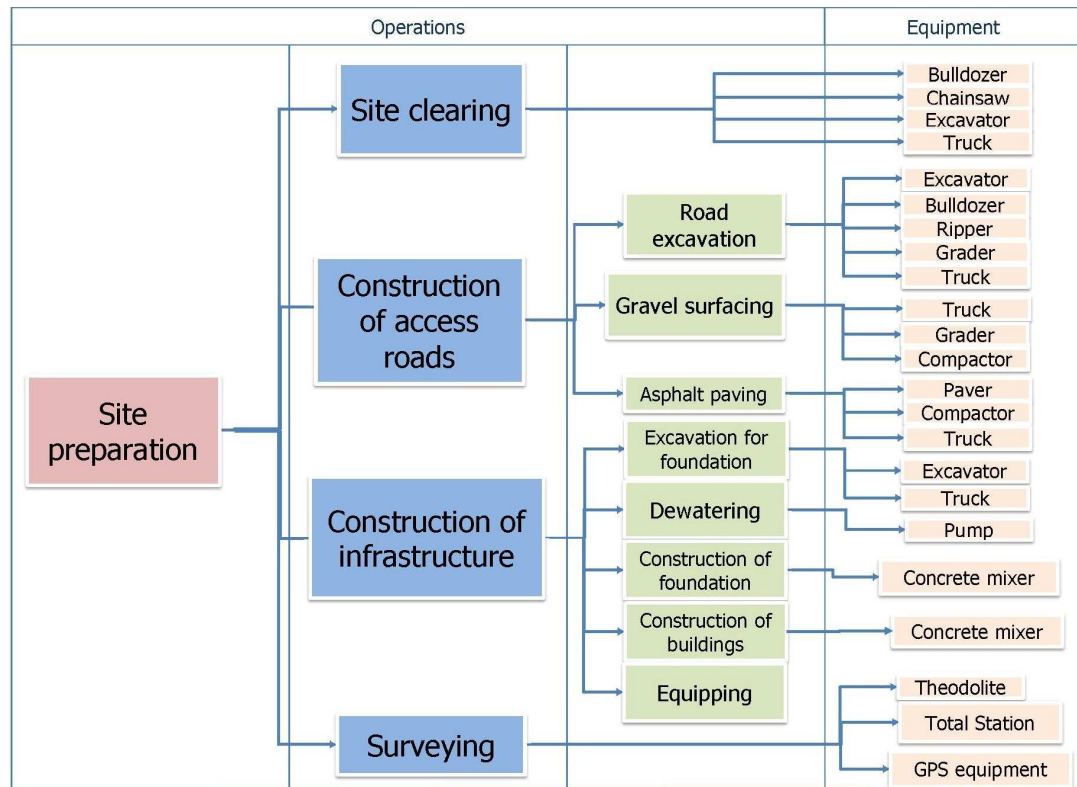


Figure 2.3: A hierarchical view of the processes, machinery and materials involved in the surface preparation phase (modified after LICYMIN, 2003)

Figure 2.4 shows a systems level breakdown of the important operations and the equipment involved in the process of topsoil and overburden removal.

The exploitation phase of the mine focusses primarily on the extraction of the mineral from the ground. Once started, it generally goes on uninterrupted for the whole life of the mine. Depending on the depth of the ore body from the earth's surface, two different methods are used - surface mining and underground mining. Figure 2.5 shows a schematic representation of the various mining methods that may be implemented. The actual ore excavation operation consists of drilling, blasting, loading, crushing and hauling operations. Various auxiliary operations like haul road construction, mobile equipment maintenance, dust suppression etc. must be carried out in parallel with these. The schematic of the various operations involved in ore excavation and auxiliary operations is illustrated in Figure 2.6. Surface mining is the most widely used method of exploitation. It requires significant capital investment, but produces more tonnage of ore with low operating costs.

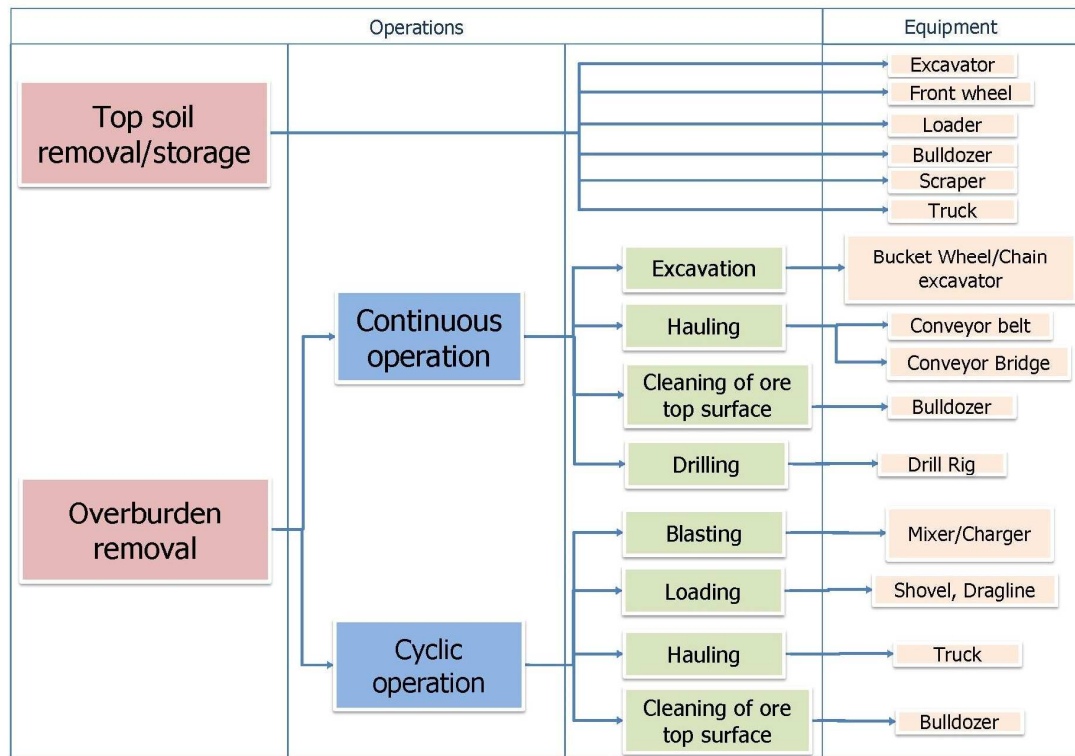


Figure 2.4: A hierarchical view of the processes involved in the topsoil and overburden removal phase (modified after LICYMIN, 2003)

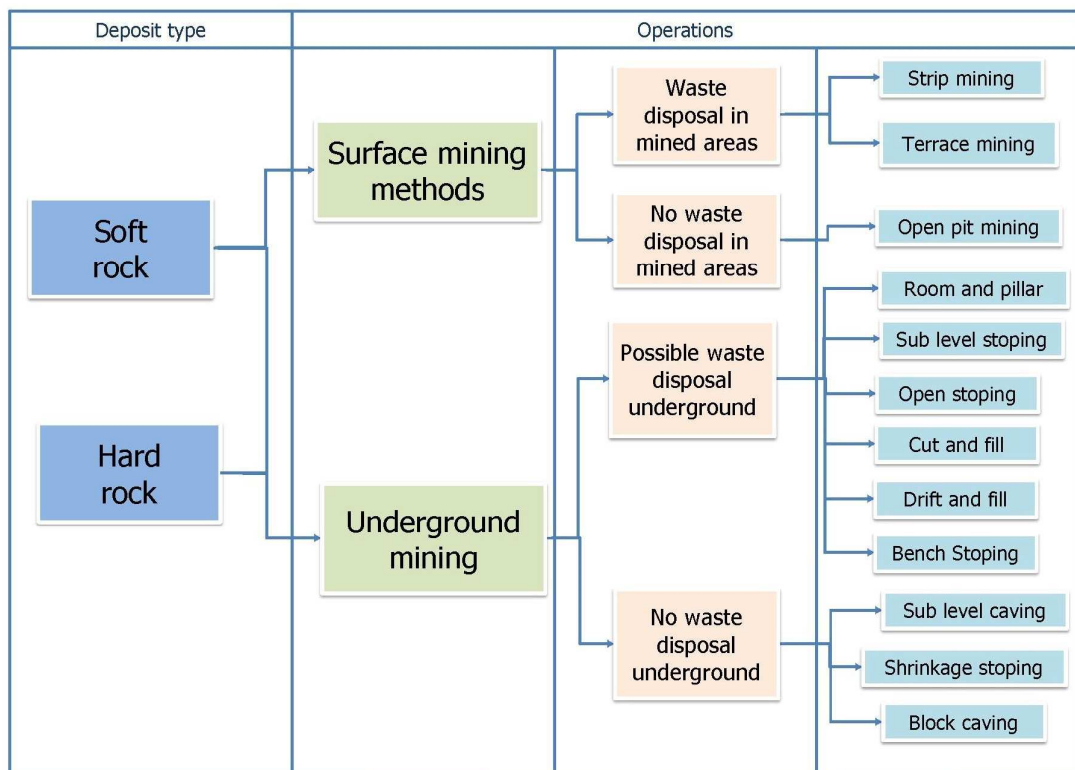


Figure 2.5: Schematic showing a hierarchical classification of the different mining methods (modified after LICYMIN, 2003)

The underground mining methods have a high labour cost and by nature induce more significant safety concerns. They are employed in cases where the mineral is very deep in the earth and it is economically unviable to extract the whole overburden and proceed with surface mining techniques, or where environmental considerations do not allow mineral operations to expose the ore and undertake mining activities on the surface.

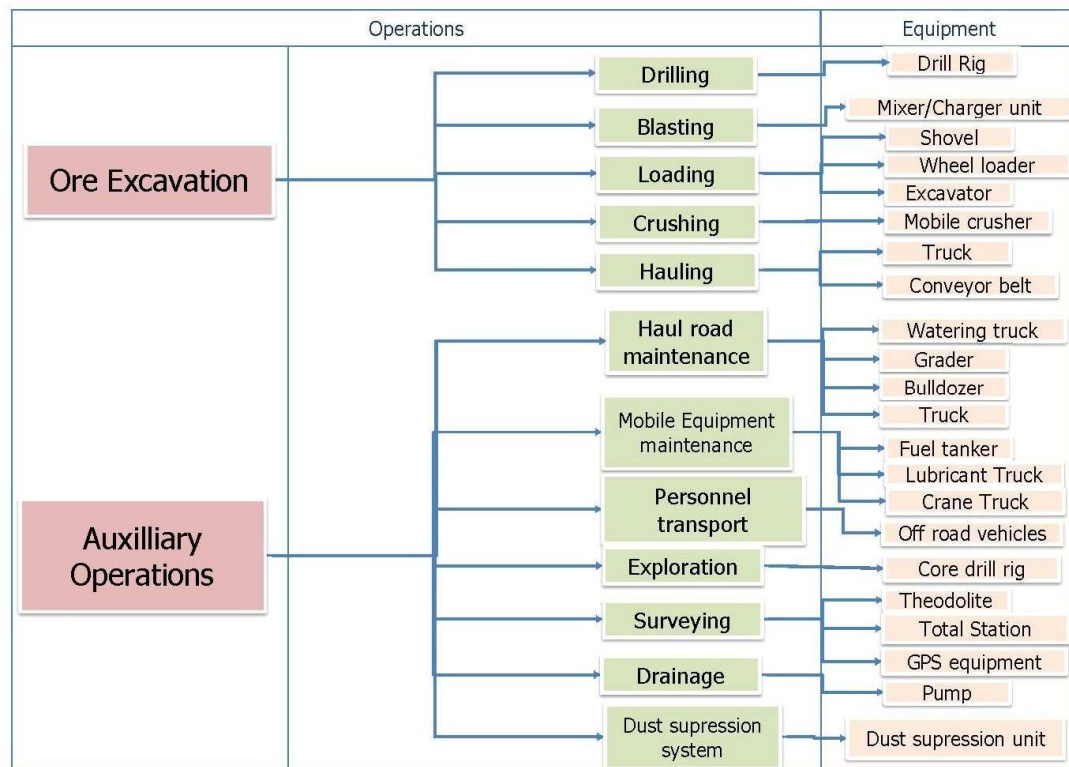


Figure 2.6: Overview of the processes, machinery and materials involved in the ore excavation and the auxiliary operations phase (modified after LICYMIN, 2003)

The waste generated during the mineral production operation needs to be disposed safely to ensure that the impacts from the mineral production on the environment are minimised. Figure 2.7 illustrates the operations involved in waste disposal during the active life of mineral production.

In the final reclamation phase, the mine is closed down and the affected areas are re-vegetated and restored. The principles of sustainability and environmental protection are adhered to. The major processes involved in this phase are outlined in Figure 2.8.

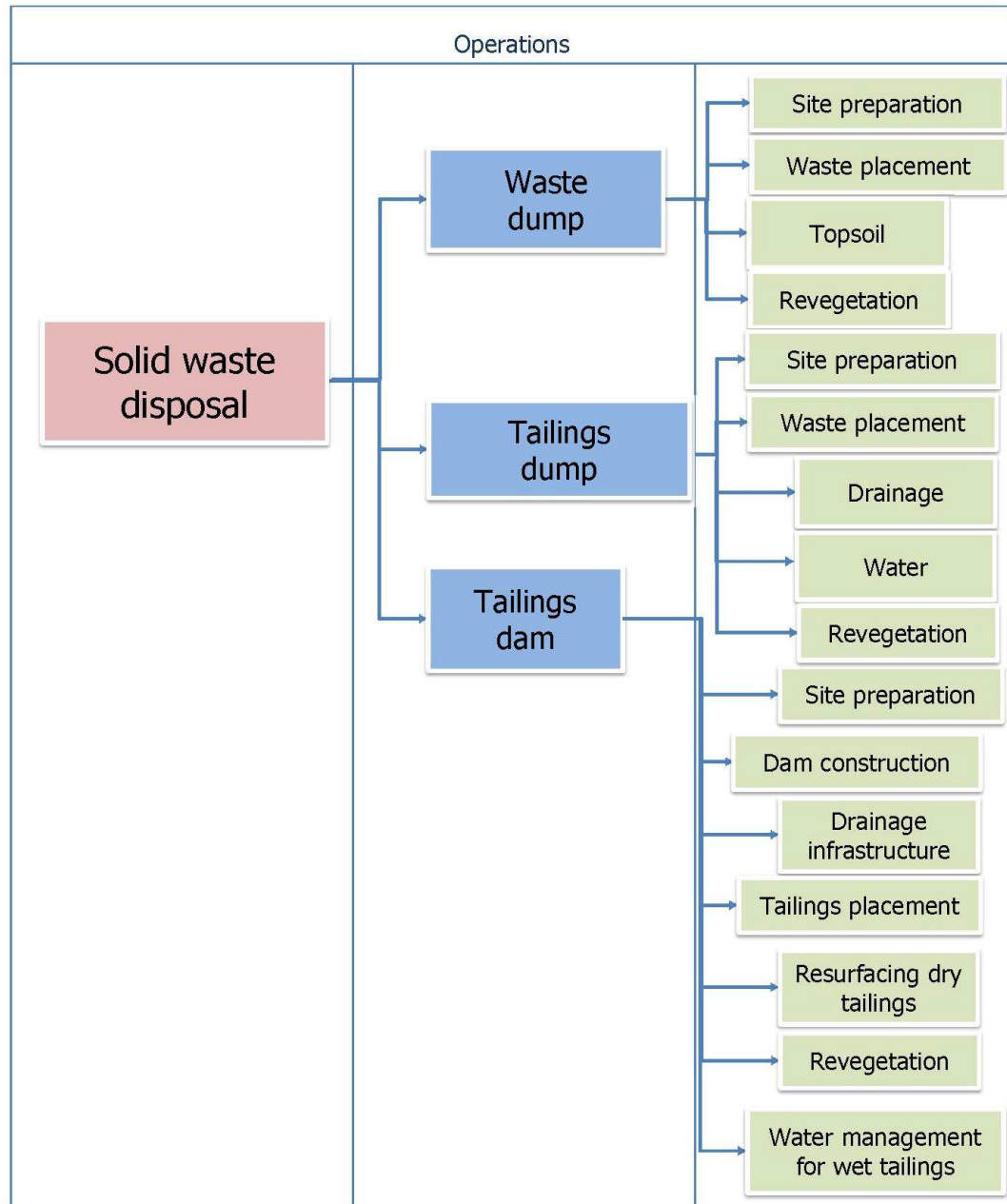


Figure 2.7: A hierarchical view of the operations involved in the waste disposal phase for hard rock mining (surface and underground) (modified after LICYMIN, 2003)

In the present research a mine in the exploitation phase, i.e. one which is producing minerals, is the one of primary interest. The risks associated with the actual operation and production period are therefore studied. The same or similar techniques may be extended to the other phases including the closure of the mining operation.

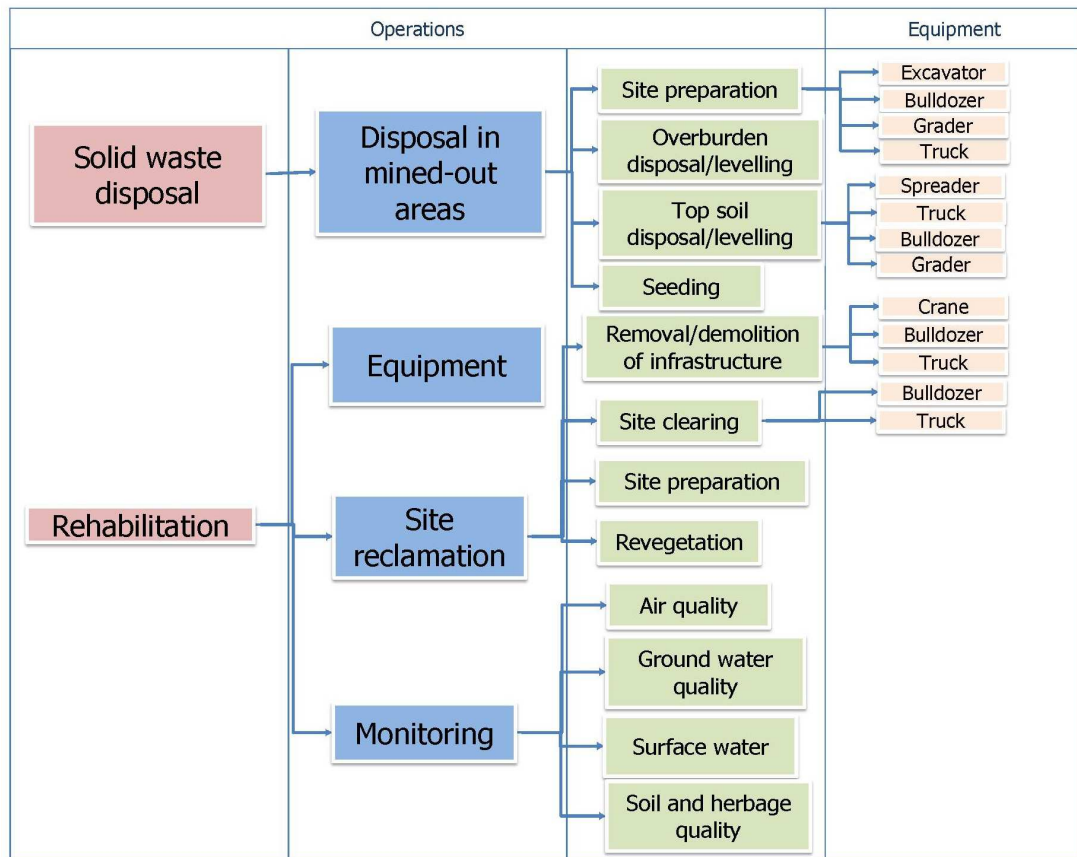


Figure 2.8: Overview of the processes and equipment involved in the reclamation phase (modified after LICYMIN, 2003)

2.7 Risk classification and modelling in mining and mineral processing sector

Some of the general mathematical tools described in the previous section, have been used to a limited extent, for risk modelling of these systems. Though a lot of literature exists on financial risk assessment in general, less work has been done in the area of financial risk modelling and analysis of chemical processes and specifically the mineral processing systems and associated mining operations.

Risk analysis in the mining industry has been studied in Heuberger (2005) and Chinbat (2011). These studies mainly focus on the perceptions and expert opinions of personnel associated with the mining sector and use a survey based technique to identify the risk vs. impact trade-off of specific components in the mining operation. However, the studies are more qualitative than quantitative and lack a strong mathematical background. The objective of the present thesis is to build and analyse models with some quantitative data from the existing mining and mineral processing

operations and build a consistent and reliable risk modelling methodology. In Chinbat and Takakuwa (2009) a simulation based approach has been outlined and an optimisation procedure has been developed to minimise the risk for the mining sector. This study is geared towards the project management risks involved in a mining operation and includes all possible tasks like exploration processes, ore estimation, planning of human resources, transportation, mine closure etc. among others. A heuristic risk ranking methodology has been proposed and a multiple objective linear programming model was used to find optimal number of engineers and workers for the mine. This kind of study is more on the lines of mine planning rather than component reliability based risk modelling to assess business interruption, which is the focus of the present thesis.

For the chemical process systems, only specific components are modelled for risk analysis and appropriate countermeasures are proposed to mitigate the same. For example, Lavaja and Bagajewicz (2004) analyse the risks of individual component failures, like the risks associated with the heat exchanger etc. Most other risk modelling techniques use process dynamics equations of mass and energy balance to analyse the effects of uncertainty or failures in certain components and arrive at a risk profile (Podofilini and Dang, 2012). These methods are effective in a data rich environment but would fail, or give highly inaccurate results, if precise and sufficiently representative data is not available. Generally component level data would be available to the process operator but they might not always be available to an external risk assessor or insurer.

Within the scope of the present study, the risk related to the mineral processing activities is characterised predominantly by the loss of production. Any shortfall in the output expected, according to the production schedule, is considered a loss. This loss can be due to various causes, such as mechanical breakdowns, unforeseen human factors, earthquakes etc. The following major types of risks are classified according to their origin:

- i. Catastrophic natural risks or extrinsic risks
 - a. Earthquakes
 - b. Floods
 - c. Typhoons

- d. Mudslides
- e. Wild fires
- f. Lightning strike
- ii. Operational failures or intrinsic risks
 - a. Mechanical breakdowns of different components in the process chain
 - b. Overheating of equipment and consequent fires
 - c. Reduction in machine life owing to dust, proximity to high temperature source etc.
 - d. Other operation specific risks
- iii. Human errors or other management decisions
 - a. Erroneous human operations
 - b. Housekeeping
 - c. Storage and use of combustible materials
 - d. Fire suppression systems and smoke alarms
 - e. Regular equipment monitoring and maintenance programs
 - f. Other strategic decision based risks

All these major categories of risk have been considered in the present study. The proposed systems model is driven by reliability data of different components and therefore considers the risks due to maintenance and breakdown of equipment. Bayesian networks have been integrated with the systems framework to model the catastrophic natural risks. Human errors and other management issues have been incorporated by coupling a questionnaire based risk modelling approach with the systems framework.

2.7.1 Catastrophic natural or environmental risks and their modelling

A site specific risk assessment index must be calculated to assess the vulnerability of the specific operation for business interruption. Various spatial features and vulnerability related indicators might be used to arrive at a risk index for the natural hazards afflicting a particular site. Specific existing modelling techniques for the

occurrences of these catastrophes might serve as background domain knowledge tools to arrive at a composite risk index.

In the case of earthquakes, the risk indicators can be the magnitude of previous earthquakes in surrounding areas, distance from seismic sources, peak ground acceleration etc. Some indicators which can classify the vulnerability of the site, like soil type, profile and response, geology of the area, ground declination etc. provide significant input to the risk index calculation and need to be used too. Various established methods exist in quantifying earthquakes. A magnitude recurrence relationship Gutenberg and Richter (1944) for the annual probabilities of various scales of earthquakes is very popular. Equations for estimating the horizontal response spectra and peak acceleration of seismic activities have been proposed in Boore et al. (1997). This can be incorporated in the present risk assessment framework. However, for a mining process, operations like blasting might result in induced seismicity and bring about catastrophic events like slope collapse. Therefore using models like Gutenberg and Richter (1944) would not be useful for considering these local effects.

In case of floods, the rainfall of the region along with its elevation, terrain, slope, soil conditions and vegetation can indicate the vulnerability of the site for these risks. The flooding cover and the water depth can be the indicators of the flood risk. Many models like the TOPMODEL (Ambroise et al., 1996), SHE (Abbott et al., 1986), DBSIM (Garrote and Bras, 1995) etc. have been proposed for predicting the relationship between rainfall and flow.

For the case of typhoons, the vulnerability indicators can be the meteorological states, elevation, terrain, slope, vegetation etc. The exposure of the site to the wind speeds and the route of the typhoon can serve as indices to identify associated the risk. The model in William and Arthur (1982) is widely used for the occurrence and reoccurrence of a typhoon along with the associated route it can take.

For assessing the risks of occurrence of mudslides, the volume of mud and the flow of sands must be considered. The rainfall, elevation, terrain, surface declination, soil characteristics, vegetation etc. can be taken as the vulnerability indicators. Jakob's equation as proposed in Jakob (1996) is popular in modelling the relationship between the volume of mudslide and the peak value of the flow.

All these models of flood, typhoons and mudslides can be incorporated in the systems risk modelling methodology as developed in this thesis. However, these models are not customised to reflect the effects of these hazards on disruption in mining activity (for example through slope collapse owing to heavy rain). Therefore a Bayesian network based formalism is adopted in this thesis to address the issue of catastrophic risks. Bayesian networks can be used to learn from existing datasets and also incorporate expert opinion in the form of causal dependence among variables and prior beliefs.

For catastrophic risks an area aggregate method (AAM) (Alexander, 1993) on coarse geological scales is used to assign a risk index. Recently, Geographic information systems (GIS) are widely used to assess catastrophic risks and assist in insurance pricing (Li et al., 2010). Other GIS based risk assessment methods rely on classical statistics or fuzzy comprehensive evaluation (Arnold, 2006; Hunsaker, 2001). Bayesian networks have been used in Korb and Nicholson (2003) for risk assessment and have the inherent capability to capture the uncertainty of natural disasters. However, precise spatial risk is generally proprietary to the insurance companies and is not publicly available. This serves as an obstacle towards novel techniques being developed, in order to improve the estimated risk (Hayes and Jacobson, 2001).

2.7.2 Operational failures or intrinsic risks

Operational failures refer to the breakdown of machinery like pressure vessel explosion, gear failure, generator breakdown, failure of drilling machines, trucks etc. Many publications look at specific mechanisms of failure and try to model these using different techniques. In Wu et al. (2012), corrosion based failure mechanisms for petrochemical plants are identified and a knowledge based reasoning model is developed for predicting the same. Structural safety assessments of building collapse in the context of reliability have been documented in (Raphael et al., 2011). Risk evaluation of failure modes for turbine rotor blades based on Dempster Schafer's theory have been attempted in Yang et al. (2011). However, these are very detailed component level reliability models to understand the root causes of failure of a very specific component. They do not incorporate the holistic picture of the whole process operation which is the focus of the present study. Stochastic and quantitative risk assessment techniques have been used in (Marhavidas and Koulouriotis, 2011) to model the worksites of an electric power provider. The use of Failure Mode Effect

and Criticality Analysis (FMECA) and Failure Time Modelling (FTM) for the case of a processing industry is documented in Ahmad et al. (2012). In general, the approach towards reliability modelling using statistical methods is to model the mean time between failures (MTBF) using some probability distribution function. The outage duration (OD) and maintenance cost associated with the breakdown can be suitably approximated using some function, based on historical data. The general philosophy of the reliability based risk modelling paradigm is adopted in the present thesis, but this is embedded in a systems framework to obtain the risk for the whole process operation.

2.7.3 Risks related to human errors or other management decisions

Significant risks can arise in the system operation if the human operator erroneously performs some activities. As a consequence of these errors, a loss of production can occur. Human Reliability Analysis (HRA) method (Dougherty, 1989) is a framework which is traditionally used as a probabilistic risk assessment methodology to quantify such risks. Many different models have been used in HRA and recent efforts have been focussed on benchmarking these techniques on a common platform (Boring et al., 2010). Broadly these diverse models for understanding the failure of human actions can be classified into three categories (Hollnagel, 1998). First of these, is the class of behavioural models which define various error modes in terms of omissions, inadvertent actions etc. The methods then try to arrive at a probability that a particular faulty action may occur. These are very simple cause-effect models and might be unrealistic in many practical settings. The second class of models focus on information processing, which aims to quantify the internal mechanisms for decision making or reasoning. The methods in this category try to explicate the flow of causes and effects through the models. The errors like lapses, slips, violations etc. are then seen as agents which cause a malfunction of the hypothetical information processing function. The third category is known as cognitive models which correlate the error modes with its various causes. These causes encompass the whole socio-technical environment. The cognitive models try to explicitly relate the context with the reason for efficient performance. The human operator is considered to both respond to events and act with an outlook of future events.

HRA for large scale engineering operations like chemical processing plants have been benchmarked in Lauridsen et al. (2002). Many techniques for identifying

hazards under a more general framework of quantitative risk assessment (QRA) have been used in process plants. Some of them are the Hazard and Operability (HAZOP) (Venkatasubramanian et al., 2000), Structured What-If Technique (SWIFT), Hazardous Scenario Analysis (HAZSCAN) (Lauridsen et al., 2001) etc.

In line with these analyses, a questionnaire based approach is developed, as a part of this work, for incorporating human factors which lead to elevated risk profiles for mining operations. It is developed by taking the inputs of veterans working in this field who have engineering experience and knowledge of the various risks associated with such operations. The questionnaire based assesement is then coupled with the systems level risk model for risk quantification.

2.8 Conclusion

The chapter introduced the mathematical notion of risk and reviewed the wide variety of techniques used in quantifying risk in different application areas. Many of the classical methods of financial risk quantification depend on various time series forecasting methods. However, these are only suitable if there is a large set of historical data and the dataset captures most of the possible excursions in profit or loss. The soft-computing or computational intelligence based approaches on the other hand are a promising alternative in cases where there is sparse data. These techniques can incorporate expert opinion in the risk modelling framework to arrive at a quantitative risk profile. Reliability based methods are especially good for engineering systems where the manufacturer of the equipment supplies the reliability data for each piece of equipment. Some of the techniques from soft computing (specifically fuzzy logic, multi-objective genetic algorithm and Bayesian networks) and reliability based methods (usage of probability distributions for time between failures and outage duration) have been incorporated in the risk models developed in this thesis.

The chapter also reviewed the current practices for risk quantification in the mining and the mineral processing industry. Most of the techniques are not mathematically rigorous and depend on the heuristic perceptions of the risk surveyor or the insurer. To address this issue, a new quantitative risk modelling framework is proposed in this thesis. The framework is very generic and is capable of integrating multiple modelling paradigms. The next chapter briefly introduces these different modelling

approaches which have been used as building blocks for constructing the risk modelling framework.

Chapter 3 Overview of design paradigms used in the risk model development

3.1 *Introduction*

The risk model developed in this research employs different modelling paradigms to cater to different design aspects including modelling flexibility, uncertainty in model parameters and source data, vagueness of information, incorporation of expert knowledge etc. This chapter provides a brief overview of the different modelling techniques used for the risk model development.

3.2 *Concept of systems thinking*

The traditional approach to modelling physical systems has largely been a reductionist one, which focuses on the individual parts of a system and aims to infer some kind of cause-effect relationship. The systems thinking framework, however, is a more holistic approach, which views a complex system as a composition of different inter-related subcomponents. The interactions among the different subcomponents often give rise to “emergent” properties which would not have been possible to detect by analysing single sub-systems in isolation. Figure 3.1 shows a schematic of the systems thinking framework.

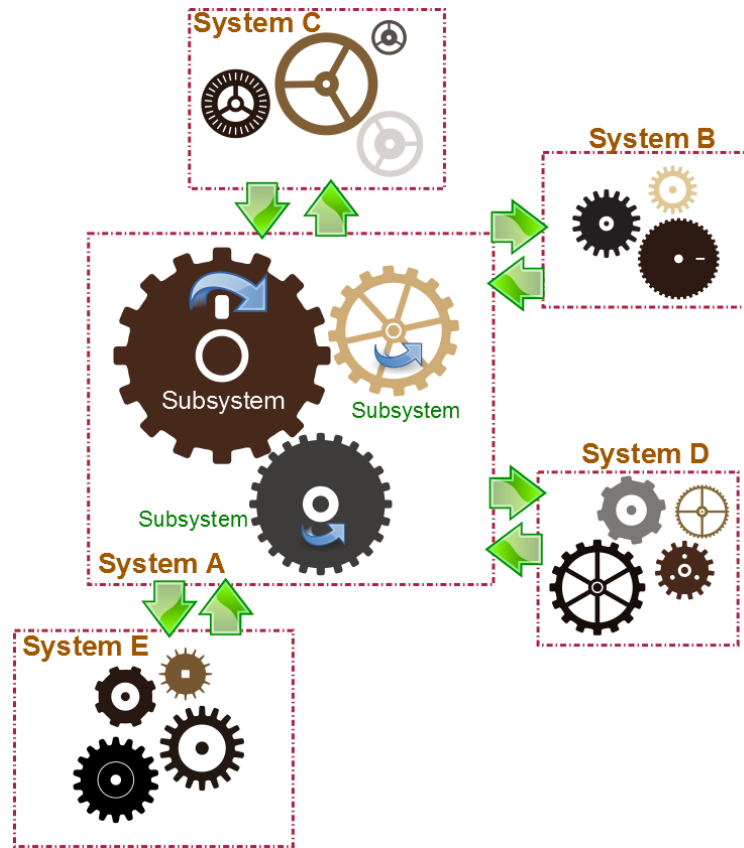


Figure 3.1: Concept of a systems thinking framework

The framework is very generic and has been adapted for modelling in diverse application areas like transportation systems, communication systems, integrated military platforms, space applications, supply chain management etc. (Checkland, 1999).

In a systems framework, each of the subsystems is composed of a number of units. Each unit has some inputs, outputs and comprises internal processes which map the inputs to the outputs. Each unit also has a well-defined boundary which serves as the interface amongst the units or between the units and the environment. The subsystems generally make up a nested hierarchy forming different system levels. Unlike traditional modelling approaches, the systems approach does not necessarily model mass or energy flows across the boundaries. The different flows across the boundaries might be information, management decisions, level of risk etc. As such, it is possible to integrate both the quantitative and the qualitative information within the modelling framework. One of the advantages of such a design paradigm is that it helps to take multi-stakeholder viewpoints into account and leverage the modelling skills of multi-disciplinary teams. In many practical applications, these different

viewpoints are contradictory in nature, i.e. an increase in performance metric from one view point would lead to a consequent deterioration in some other performance metric. For example, for a chemical process company, the primary objective of the management is to make profit which might be achieved through increased production. However, this might give rise to environmental concerns which affect the society at large. A systems level model which integrates the process operation with the interests of the environmental stakeholders would help in making effective management decisions for the overall system. This holistic framework is useful for system optimisation as well, since the system designer can choose the optimal trade-offs between the different conflicting system objectives. This is different from the traditional component level modelling and optimisation, since optimising individual components of the system might not lead to the best strategy for the overall system operation.

3.3 Finite state machines

A finite state machine is a computational model which consists of a discrete number of states and a mechanism for transition between these finite numbers of states. Figure 3.2 shows an example of a finite state machine which represents the condition of an equipment over time. There are two states – Healthy and Failure, and the system transits between these two states depending on the guard conditions. As soon as the system enters one of the states, the state actions are executed which assign the value of Health of the equipment. The guard conditions are functions of time, which might depend on the maintenance level of the equipment, age of the equipment etc.

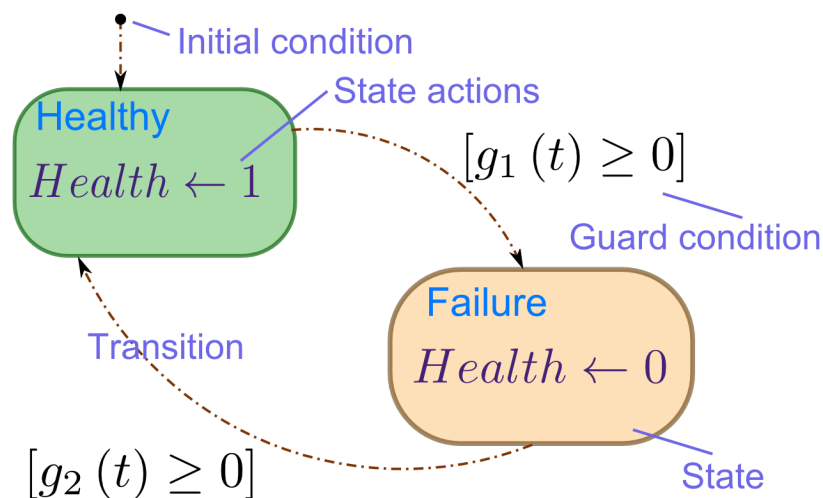


Figure 3.2: State chart of a finite state machine

Mathematically a finite state machine (FSM) can be represented by a triple $\Phi_f = (\Gamma_f, \gamma_f, \xi_f)$ where Γ_f and γ_f are non-empty finite sets and $\xi_f : \Gamma_f \times \gamma_f \rightarrow \Gamma_f$ is a function. The elements of Γ_f are known as the states of the system. The set γ_f is the input alphabet and its elements are called letters. ξ_f represents the state transition function. A finite state machine of this form can be pictorially represented by a Moore diagram which is essentially a directed graph with labelled vertices and edges which represent the states and the possible transition routes respectively. The FSM as described without any actions is known as a transition machine. It can be further augmented by different kinds of actions like an entry action, exit action, transition action and input action. An automaton is an FSM which consists of actions. FSMs can also be classified based on the type of actions that they use. A Moore machine uses only entry actions, whereas a Mealy machine uses only input actions. Both these representations are equivalent from the operation point of view. However, the Moore machines exhibit simple behaviour whereas Mealy machines can be designed to have less number of states.

3.4 Basics of hybrid dynamical systems

Hybrid systems are those systems which include both continuous and discrete dynamics. These are useful in capturing a wide range of dynamic phenomena, which have instantaneous transitions in between smooth evolution of state variables. A schematic of a hybrid automaton is shown in Figure 3.3. The different discrete states ($q = \{1, 2, 3\}$) are represented by the green oval shapes. Within each discrete state the continuous state variable (x) evolves according to a specific set of differential equation. The arrows represent the transition between the states and the guard conditions dictate when the transition would occur.

Mathematically a hybrid automaton (H) can be defined as a collection $H = (Q, X, Init, f, \psi, \zeta, \omega, \theta)$ where

- Q is a finite collection of discrete variables
- X is a finite collection of continuous variables with $X \in \mathfrak{R}^n$
- $Init \subseteq Q \times X$ is a set of initial states

- $f : Q \times X \rightarrow \mathbb{R}^n$ is a vector field
- $\psi : Q \rightarrow P(X)$ is the domain $\forall q \in Q$
- $\zeta \subset Q \times Q$ is a set of edges
- $\omega : \zeta \rightarrow P(X)$ is a guard of $\varepsilon \forall \varepsilon \in \zeta$
- $\theta : \zeta \times X \rightarrow P(X)$ is a reset map $\forall \varepsilon \in \zeta, x \in X$

The state of H is referred to by $(q, x) \in Q \times X$ (Johansson et al., 1999; Lygeros et al., 1999; Zhang et al., 2001).

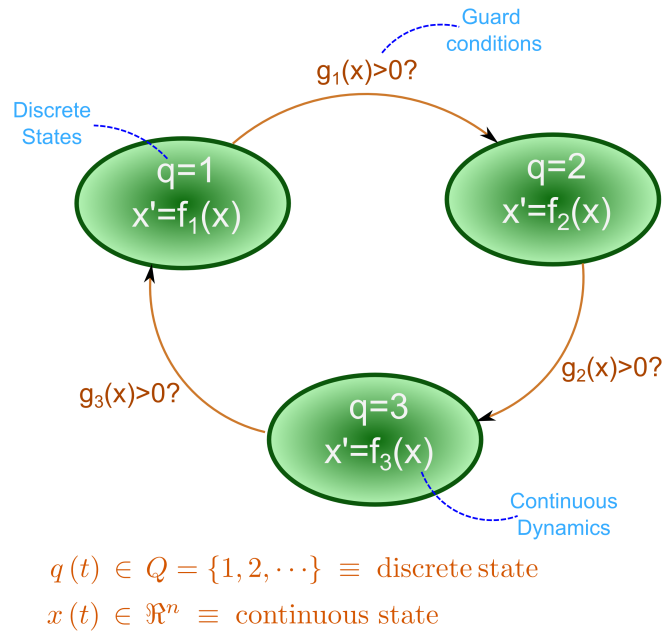


Figure 3.3: Schematic of a hybrid automaton

Figure 3.4 shows a schematic for the evolution of a switched system's state variable over time. The blue curves represent the evolution of the state trajectory due to some underlying set of differential equations. The red circles and the associated arrows indicate the switching of the state from one point to the other due to some switching logic. The switching is assumed to be instantaneous, i.e. the underlying temporal dynamics of the discrete switching logic are not considered.

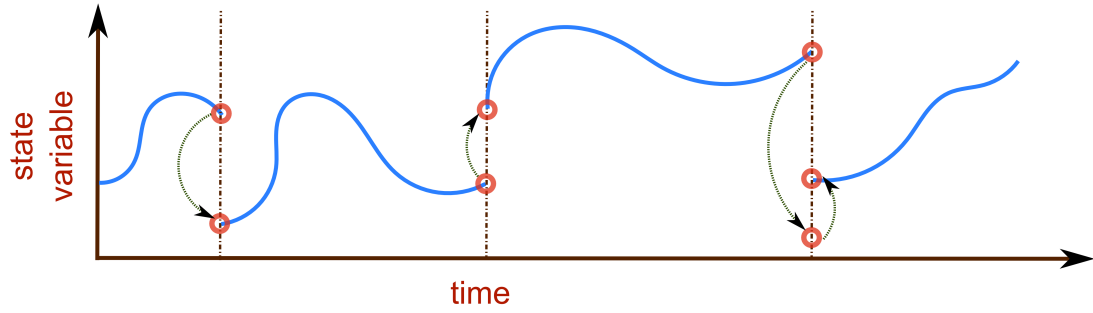


Figure 3.4: Schematic of continuous evolution of state variables with discrete switching instants

3.5 Discrete event simulation techniques

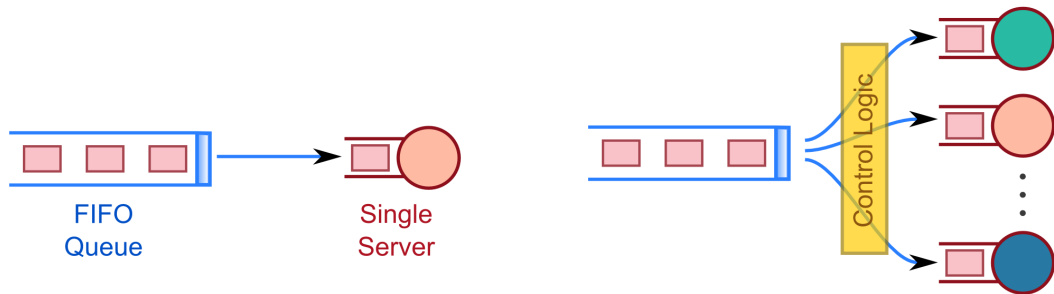
In general, simulation of a dynamical system proceeds by continuously calculating the states of the system over time. The resolution of each time step is generally small, to closely approximate the continuous time evolution of the states. This approach requires higher simulation times since all the intermediate time steps are required to be calculated. The discrete event simulation approach (DES), on the other hand, is an event based approach as opposed to a time based one (Cassandras and Lafortune, 2008). In this approach, the time is sliced up over different intervals and the states of the system are calculated based on the activities occurring in those intervals. All the time intervals do not need to be simulated in this approach and, therefore, DESs typically require less simulation time. A DES generally consists of system states, a clock, an event list, pseudo-random number generators, counters for calculating statistics, stopping criteria and a logic which defines how the system changes in response to an event. In a DES, events occur instantaneously. The clock keeps track of the simulation time. However, unlike continuous time simulators, DES skips the intermediate times between two consecutive events. The event list consists of the pending events which have arisen as a consequence of a previous event, but have not been simulated yet. All DES techniques use some pseudo-random numbers for generating the random variables in the system models. These pseudo-random number generators might be linear congruential generators, Mersenne twisters etc. and are useful for obtaining the same evolution of the states in multiple re-runs of the simulation. The DESs also typically employ some counters, which calculate different system statistics. The advantage of such a method is that it helps to statistically analyse complex system behaviour which might not be analytically tractable. These are obtained by multiple runs of the same simulation model and

confidence intervals are then ascribed to the variables of interest for effective system characterisation.

One of the basic models in DES is that of a queue in which there are different entities or customers arriving sequentially and are being serviced by a server. The next section outlines the basics of queuing theory and some of the basic queue-server configurations.

3.5.1 Basics of queuing theory

Figure 3.5 (a) shows a basic configuration for a First-In-First-Out (FIFO) queue with discrete entities which are served sequentially by a single server. For a generic case, the service time of the server, for each entity, is drawn from a prespecified probability distribution. The entity remains in a locked state with the server, for the duration of the service time. After this time elapses, the entity is forwarded by the server to the subsequent stages and the next entity in the queue is taken up by the server. The queue length can be finite or infinite and the entities can be added to the end of the queue at any time. In case the queue reaches the maximum number of entities it can hold, further entities are rejected, until the server processes an entity and frees up the queue.



(a) Single server with a single FIFO queue

(b) Parallel heterogeneous servers

Figure 3.5: Basic queue-server configurations

Figure 3.5 (b) shows one queue being serviced in parallel by different servers, each having its own range of service times. The control logic in this case, would try to assign one entity from the queue to that server which is free. Since each server has a different stochastic service rate, they would process a different number of entities by the end of a given time, i.e., the one with a lower service time would process more entities than the one with a higher service time.

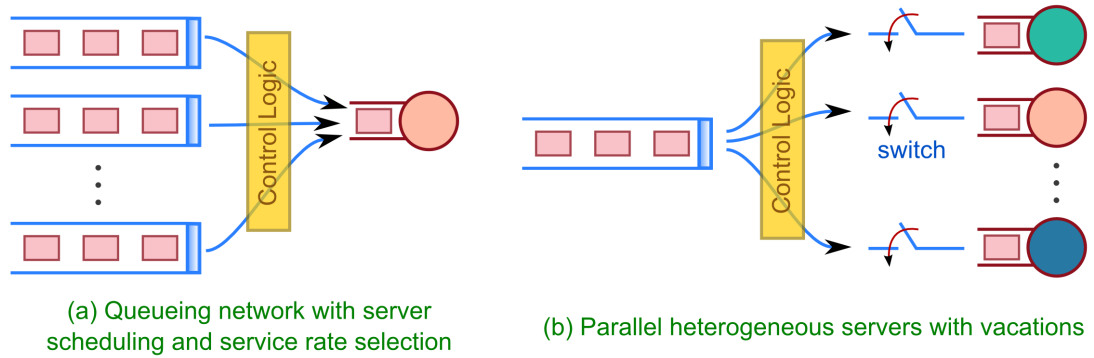
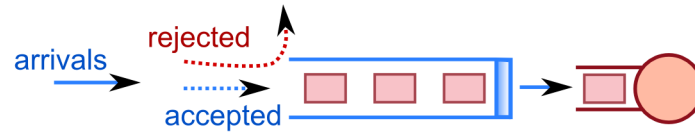


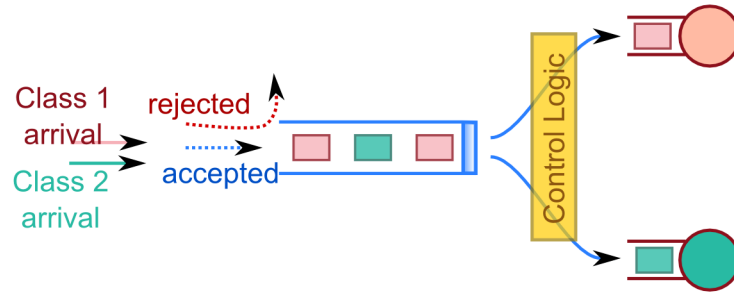
Figure 3.6: Queue-server configurations with scheduling mechanisms

Figure 3.6 (a) shows multiple queues being processed by a single server. The control algorithm in this case should have a scheduling mechanism and assign a service rate to each queue. If the objective is to minimise the average waiting time amongst all the queues, then the control logic might be designed to service more frequently the queues which have a higher queue length. In case there are some queues which are more important than the others, the control logic can be designed to service those queues more frequently than the others. Figure 3.6 (b) shows the case of parallel heterogeneous servers which can be switched on or off from the system. The control logic in this case must take account of the number of servers that are currently active at any given instant and decide the scheduling between the active servers.

Figure 3.7 (a) shows a single server with admission control. The input to the queue is controlled by some gating mechanism which selectively admits entities into the queue. This admission control may be based on some attributes of the entities themselves or might be based on some temporal logic. Figure 3.7 (b) shows a more complicated arrangement where there are two classes of arrival streams along with the admission control mechanism described in Figure 3.7 (a). Each class of the entity is serviced exclusively by a particular server and the control logic routes the entities to the appropriate server.



(a) Single server with admission control



(b) A parallel server queue with two classes of arrival streams and admission control

Figure 3.7: Queues with admission control

Figure 3.8 shows the time evolution of a special class of queues known as priority queues. There are two classes of entities which arrive at the queue. However, Class 1 entities have a priority over the Class 2 entities. Therefore, in case a Class 1 entity comes, it does not need to wait at the end of the queue. It is put in the beginning of the queue to be serviced first. In Figure 3.8, the queue is empty at time $t = 0$. A couple of Class 2 entities arrive and fill up the queue at $t = \{1, 2\}$. Then a Class 1 entity with higher priority arrives and is put in front of the queue instead of the end of the queue. The state of the queue at $t = 3$ shows this scenario.

Figure 3.9 shows a simple example of how different queues and servers can be coupled in series or parallel configurations to model a complex engineering process system. In a practical modelling scenario, there would be entity generation sources and entity sinks which are not shown in the figure. Also at each junction there would be some sort of control logic which would dictate the flow of the entities, e.g. how two or more streams of entities would be coupled into a single stream, if there are any priorities of the streams etc.

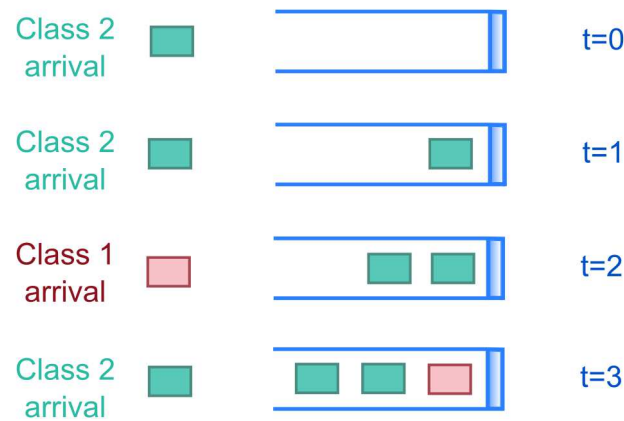


Figure 3.8: Time evolution of a priority queue

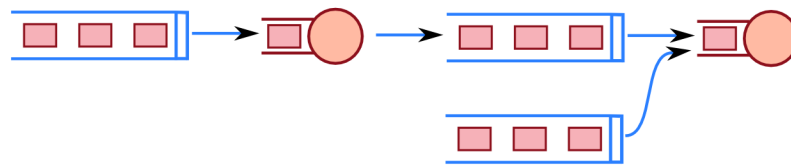


Figure 3.9: Series-parallel interconnections of servers and queues to model complex process systems

3.6 Basics of fuzzy logic

Fuzzy logic is a mathematical framework for approximate reasoning using multi-valued logic as opposed to conventional binary logic. In binary valued logic, an element can either be associated with a set or not associated with it, i.e., there are two states – true and false. In a multi-valued logic, there is a degree of membership with each set which varies between 0 and 1. Fuzzy logic has been applied to a wide variety of fields like control systems, machine learning, artificial intelligence, expert systems etc. (Ross, 2009). One of the advantages of using the fuzzy approach is that it helps to include the heuristic rules supplied by the experts, in the modelling framework. Figure 3.10 shows a schematic of a Fuzzy Inference System (FIS) with the fuzzification, inferencing and the de-fuzzification modules. The crisp inputs are taken by the FIS and are transformed into fuzzy variables by a set of linguistic rules. The inferencing module works on these fuzzified variables to produce an output set of fuzzy variables. The expert knowledge is encoded in the form of IF...THEN rules which is used in the inferencing process. The fuzzy output variables are then converted to crisp numbers using a de-fuzzification module.

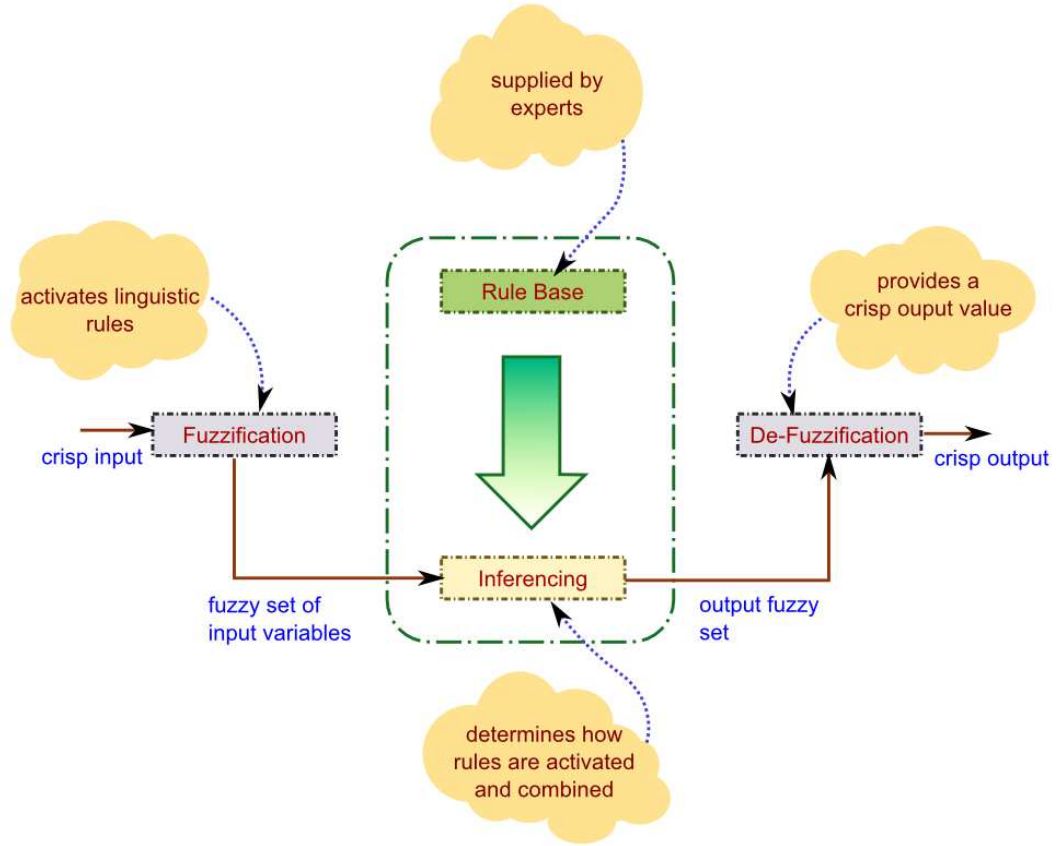


Figure 3.10: A fuzzy inference system

A fuzzy set comprises of a universe of discourse and a membership function that maps all elements in the universe of discourse to a membership value between 0 and 1. Mathematically, if an element is represented as $x \in X$ where X is the universe of discourse, the membership function of the fuzzy set $\tilde{A}$ is denoted by $\mu_{\tilde{A}}(x)$. Various fuzzy logic operations like AND, OR, NOT can be defined on the fuzzy sets and are used in the inferencing process. Some of them are described below (Jang et al. 1996), (adapted in Pan et al., 2012).

Zadeh fuzzy logic AND operator: $\mu_{\tilde{A} \cap \tilde{B}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$

Product fuzzy logic AND operator: $\mu_{\tilde{A} \cap \tilde{B}}(x) = \mu_{\tilde{A}}(x) \times \mu_{\tilde{B}}(x)$

Zadeh fuzzy logic OR operator: $\mu_{\tilde{A} \cup \tilde{B}}(x) = \max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$

Lukasiewicz fuzzy logic AND operator: $\mu_{\tilde{A} \cap \tilde{B}}(x) = \min(\mu_{\tilde{A}}(x) + \mu_{\tilde{B}}(x), 1)$

The main elements involved in fuzzy inference system include fuzzification, Mamdani or Takagi-Sugeno type rule base, fuzzy inferencing and defuzzification.

Fuzzification is a mathematical method for converting an element in the universe of discourse to the membership value of the fuzzy set. For example, for $x \in \tilde{A}$ defined on the universe of discourse $[a, b]$, the outcome of the fuzzification is $\mu_{\tilde{A}}(x)$.

The fuzzy rules are linguistic IF-THEN expressions, fuzzy sets, fuzzy logic and fuzzy inferencing. They are used to mimic the expert knowledge in the fuzzy inference systems and in linking the input variables of the fuzzy controllers to the output variables. A Mamdani type fuzzy rule involving three input variables and two output variables can be described as follows:

IF x_1 is $\tilde{A}$ AND x_2 is $\tilde{B}$ AND x_3 is $\tilde{C}$ THEN u_1 is $\tilde{D}$, u_2 is $\tilde{E}$

where, x_1, x_2, x_3 are the input variables and u_1, u_2 are the output variables.

Takagi-Sugeno (TS) fuzzy rules use functions of input variables as the rule consequent. The TS rule corresponding to the Mamdani rule may be expressed as:

IF x_1 is $\tilde{A}$ AND x_2 is $\tilde{B}$ AND x_3 is $\tilde{C}$ THEN $u_1 = f(x_1, x_2, x_3)$, $u_2 = g(x_1, x_2, x_3)$

where, $f(\cdot)$ and $g(\cdot)$ are two real functions.

The fuzzy inferencing is the approximate reasoning used to define the rule outcome from a given rule input information. The fuzzy inferencing thus takes the input variables in the rule antecedent and calculates the output variables in the rule consequent. Some of the popular fuzzy inferencing methods along with their mathematical model are defined as follows:

Mamdani minimum fuzzy inference: $\min(\mu, \mu_{\tilde{W}}(z)) \forall z$

Larsen product inference: $\mu \times \mu_{\tilde{W}}(z) \forall z$

Drastic product inference:
$$\begin{cases} \mu, & \text{for } \mu_{\tilde{W}}(z) = 1 \\ \mu_{\tilde{W}}(z), & \text{for } \mu = 1 \\ 0, & \text{for } \mu < 1 \text{ and } \mu_{\tilde{W}}(z) < 1 \end{cases}$$

Bounded product inference: $\max(\mu + \mu_{\tilde{W}}(z) - 1, 0)$

Here, $\mu_{\tilde{W}}(z)$ is the membership function of the fuzzy set $\tilde{W}$ representing $\tilde{W}_j$ in the rule consequent, whereas μ is the final membership yielded by the fuzzy logic AND operators in the rule antecedent.

Defuzzification is the process of converting fuzzy sets to a real number. Let the output variable of a fuzzy inference system be z . Let the evaluation of N Mamdani fuzzy rules using some inference method produce N membership values $\mu_1, \mu_2, \dots, \mu_N$, for N singleton output fuzzy sets in the rules. Let the fuzzy sets be nonzero only at $z = \beta_1, \dots, \beta_N$. The generalized defuzzifier gives the following output:

$$z = \frac{\sum_{k=1}^N \mu_k^\alpha \times \beta_k}{\sum_{k=1}^N \mu_k^\alpha} \quad (3.1)$$

Here α is a design parameter.

3.7 Overview of Bayesian networks

A Bayesian network or a probabilistic graphical model is a representation of a set of random variables in the form of a directed acyclic graph. The nodes of the graph represent the random variables and the edges represent the conditional dependencies amongst the different random variables. The random variables of the nodes might be a measurable quantity, latent variables or unknown parameters. Each of the nodes is directly influenced by its parent nodes. Each node consists of a probability function which takes a set of input variables from the parent nodes and computes the conditional probability of that node.

Mathematically a Bayesian network (BN) can be defined as follows (Koller and Friedman, 2009; Li et al., 2010):

For a set of random variables ξ_b , a BN is an ordered pair (ϕ_b, ϖ_b) with the following property—

$\phi_b = G(\xi_b, \psi_b)$ is a directed acyclic graph, where $\psi_b \in \xi_b \times \xi_b$ is the set of directed edges. These represent how the elements of ξ_b are conditionally dependent

on each other. The set ψ_b also satisfies the Markov property that there are no direct probabilistic dependencies in ϕ_b that are not expressed by the edges ψ_b .

$\varpi_b = \left\{ \lambda_q : \Lambda_q \times \Lambda_{\sigma_q} \rightarrow [0 \dots 1] \mid q \in \xi_b \right\}$ is a set of assessment functions, where the state space Λ_q is a finite set of values of q , σ_q denotes the set of parent nodes of q which can be obtained from the network structure in ϕ_b , $\Lambda_Y \forall Y \in \{q, \sigma_q\}$ is the Cartesian product of all state spaces of the variables in Y , λ_q uniquely defines the joint probability distribution $P(q \mid \sigma_q)$ of q , conditional on its parent set σ_q .

For a BN as defined above, it is possible to calculate the posterior probability or belief of a target variable t_v by computing the marginal probability

$$Bel(t_v) = \sum_{q_i \in \xi_b, q_i \neq t_v} P(q_1, q_2, \dots, t_v, \dots, q_n) \quad (3.2)$$

where $P(q_1, \dots, q_n) = \prod_{u_i \in \xi_b} P(q_i \mid \sigma_{q_i})$ is the joint probability over ξ_b .

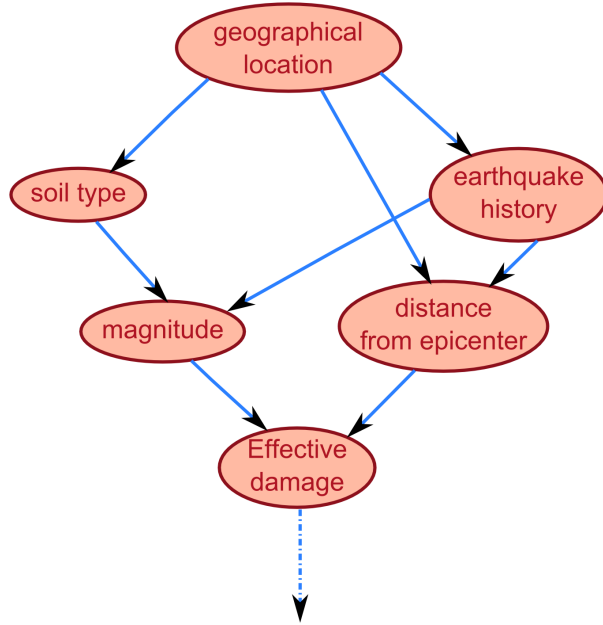


Figure 3.11: A Bayesian network structure for predicting the damage owing to an earthquake. Each of the nodes can take a set of finite values and have a conditional probability table associated with them. These are not shown in the schematic.

Bayesian networks have been successfully applied in many different fields like risk analysis, weather forecasting, financial systems, gene regulatory networks etc. (Koller and Friedman, 2009). Figure 3.11 show the graph based representation of a Bayesian network which can be used to predict the effective damage due to an earthquake. The inter-dependent random variables which help in computing the probability of the final output node (effective damage) is represented as labelled nodes in the graphical structure.

3.8 Questionnaire based expert approach

A traditional way of assessing risk in engineering sectors is to use a questionnaire based approach. An experienced risk surveyor inspects the engineering process operation and assigns a rating of the operation in specific sectors like level of safety procedures maintained, training of management staff, past history of events etc. The ratings from these different categories are then compiled into a composite risk index and compared with other similar operations with a known risk profile. This gives an indication of whether the operation is relatively more risky than the other contemporary ones. The advantage of this method is that it can use the expertise of the risk surveyor and capture the relative level of risk from soft information, which would otherwise be hard to quantify. However, on the other hand, this method suffers from the personal bias of the surveyor. There would also be an inherent human variability across different trials, which is not desirable. Considering that expert information can indeed be valuable in most cases, a provision is made in the risk modelling tool to adjust the probability distributions of the failure modes based on the expert inputs.

3.9 Brief outline of multi-objective evolutionary algorithm for optimisation

Evolutionary algorithms use the Darwinian principle of ‘survival of the fittest’ to evolve better solutions with respect to some fitness metric. The optimisation variables are encoded in the form of chromosomes and they exchange information among themselves by reproduction, crossover and mutation, to create new generations. Each individual created as a result of these stochastic operations are ranked according to their fitness which is given by an objective function. The individuals with better fitness are carried on to the next generation. This goes on

iteratively and the solutions get refined until some termination criterion is satisfied. This principle can be used to optimise either a single objective function or multiple contradictory objective functions to design a system.

A generalized multi-objective optimisation framework can be defined as follows (Coello, 2000; Fonseca and Fleming, 1995):

$$\begin{aligned} &\text{Minimize } F(x) = (f_1(x), f_2(x), \dots, f_m(x)) \\ &\text{subject to : } g_i(x) \leq 0, \forall i \in [1, p], \\ &\quad h_j(x) = 0, \forall j \in [1, q] \end{aligned} \quad (3.3)$$

such that $x \in \Omega$, where Ω is the decision space, $\mathfrak{R}^m$ is the objective space, $F : \Omega \rightarrow \mathfrak{R}^m$ consists of m real valued objective functions and $g_i(\bullet)$ and $h_j(\bullet)$ are the optional p number of inequality and q number of equality constraints on the problem respectively.

Let, $u = \{u_1, \dots, u_m\}$, $v = \{v_1, \dots, v_m\} \in \mathfrak{R}^m$ be two vectors. u is said to dominate v if $u_i < v_i \forall i \in \{1, 2, \dots, m\}$ and $u \neq v$. A point $x^* \in \Omega$ is called Pareto optimal if $\nexists x | x \in \Omega$ such that $F(x)$ dominates $F(x^*)$. The Pareto set consists of all the Pareto optimal points and is represented by PS . The Pareto Front (PF) is the set of all Pareto objective vectors and can be expressed mathematically as $PF = \{F(x) \in \mathfrak{R}^m, x \in PS\}$. In other words, there are no solutions which can simultaneously improve on all the objective functions. Any improvement in one objective function would lead to a consequent deterioration in another one or more contradictory objective functions.

There are many multi-objective evolutionary algorithm methods like the Non-dominated Sorting Genetic Algorithm (NSGA) (Deb et al., 2002), Niche-Pareto Genetic Algorithm (NPGA) (Horn et al., 1994), Pareto Archived Evolution Strategy (PAES) (Knowles and Corne, 2000), Strength Pareto Evolutionary Algorithm (SPEA) (Zitzler et al., 1998) MOEA/D (Zhang and Li, 2007) etc. The generic principles of all these algorithms are the same in that they use the concepts of selection, crossover and mutation to evolve future generations. They differ mainly in the way the ranking of the different individuals are done and the nature of selection pressure that is applied for evolving newer generations. Each algorithm has its own

advantages and disadvantages. The NSGA II algorithm (Deb et al., 2002) is one of the popular ones which work in a wide range of scenarios. It is also employed for optimisation of the copper processing system in the thesis.

The pseudo-code for the NSGA II algorithm is given as in Algorithm 1 (Deb et al., 2002, Fonseca and Fleming, 1995).

Algorithm 1: The NSGA-II Algorithm

Step 1. [Start] Randomly generate N chromosomes within feasible search space

Step 2. [Fitness] Evaluate multiple fitness of each solution in the population

Step 3. [Rank] Rank Population by

i. **[Domination rank]** Rank individuals using Algorithm 2

ii. **[Crowding Distances]** Calculate using Algorithm 3

Step 4. [New Population] Create New Population through

i. **[Selection]** Select parent chromosomes from previous population using crowding selection operator in Algorithm 4

ii. **[Crossover]** Using specified probability, do crossover of parents to form new offsprings.

iii. **[Mutation]** Using specified probability, mutate chromosomes to form new offsprings.

iv. **[Accept]** Place new offsprings in the new population

Step 5. [Replace] Replace old population by new and continue

Step 6. [Test] If termination criteria is satisfied, return Pareto set of solutions from current population and stop

Step 7. [Loop] Go to **Step 2.**

The domination rank assignment, crowding distance assignment and the crowding selection operator are illustrated in Algorithms 2, 3 and 4 respectively (Coello, 2000, adapted in Pan and Das, 2013).

Algorithm 2: Domination rank assignment

Step 1. Rank counter $r \leftarrow 0$

Step 2. $r \leftarrow r + 1$

Step 3. Find non-dominated individuals from population P

Step 4. Assign rank r to these individuals

Step 5. Remove these individuals from P and continue

Step 6. If P is empty then stop, else Go to **Step 2.**

Algorithm 3: Crowding Distance assignment

Step 1. Let $d_i \leftarrow 0$ for $i = 1, 2, \dots, \Psi$

Step 2. For each objective function $f_k, k = 1, 2, \dots, \Gamma$
 sort the set in ascending order

Step 3. Let d_1 and d_Ψ be the maximum values, e.g. $d_1 = d_\Psi = \infty$

Step 4. For $j=2$ to $(\Psi-1)$, set $d_j \leftarrow d_j + (f_{k_{j+1}} - f_{k_{j-1}})$

Algorithm 4: Crowding Selection Operator

$x \succ y$ iff $r_x < r_y$

or $r_x == r_y$ AND $d_x > d_y$

3.9.1 Concept of stochastic objective function and its optimisation

In some optimisation problems, like the one used in this thesis, the objective function does not give a unique value for each particular set of optimisation variables. Every time it is evaluated, a different outcome occurs due to the stochastic nature of the underlying problem. Figure 3.12 shows the concept of a stochastic objective function.

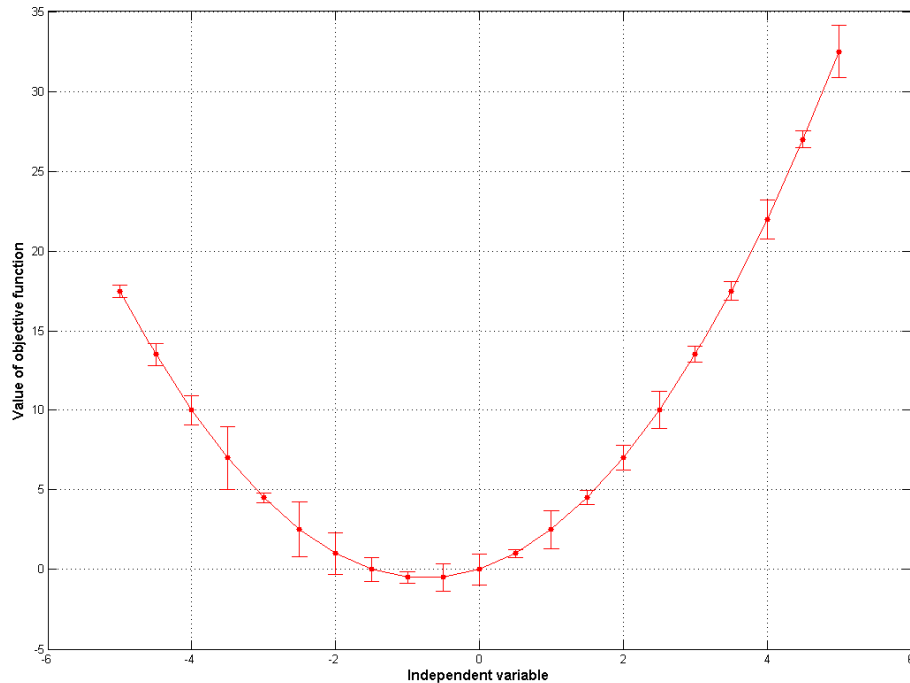


Figure 3.12: Concept of a stochastic objective function

The abscissa represents the independent variable which needs to be optimally chosen by an algorithm to yield an expected minimum of the objective function on the ordinate. The error bars associated with each point in the figure, show the maximum and the minimum values of the objective function when evaluated multiple times at that same point. Also it can be seen that the deviation of the solution at some points is much greater than at other points.

These kinds of problems cannot be handled by conventional gradient descent methods which rely on derivative information at a point. However, evolutionary algorithms (EAs) are expedient in optimising these kinds of functions. In evolutionary computation literature these kinds of problems are known as noisy or uncertain function optimisation problems (Jin and Branke, 2005). There are many research articles which propose improvements over existing evolutionary algorithms to tackle these kind of problems better (Beyer, 2000). The simplest technique is to evaluate the function multiple times at the same point and calculate the expected value of the objective function at that point. The EA ranks the solutions based on this expected value and subsequently performs selection, crossover and mutation. The self-averaging nature of EAs (Tsutsui and Ghosh, 1997) helps in the convergence of solutions, i.e. the solution vectors which have good fitness values survive through the generations and remain in subsequent populations. However, the convergence of the algorithm is more time consuming than the simple GA. This method is adopted in the present study for its simplicity. The optimisation algorithm is used to find out different modifications of the process system which would have different reliability vs. capital cost tradeoff. Since the underlying risk model is itself probabilistic, it gives rise to a stochastic objective function and therefore this method is suitable for optimisation.

3.10 Conclusion

This chapter introduced the elements of the systems framework which is the core of the proposed modelling methodology and is delineated in the next chapter. The finite state machines and the hybrid systems introduced in this chapter are flexible modelling tools which are used to formulate the logic of the systems based risk model. The fuzzy logic and Bayesian network based mechanisms are used to incorporate expert knowledge in the risk model. The next chapter proposes the systems based risk modelling framework and validates its working with a

hypothetical risk item. Subsequent chapters build on this basic model and incorporate fuzzy, Bayesian and queuing theory techniques to address different modelling issues.

Chapter 4 A systems approach for risk modelling

4.1 *Introduction*

This chapter proposes a new risk framework using the systems approach methodology and modelling paradigms outlined in the previous chapter. The risk model uses finite state machines to incorporate the logical elements within it. A theoretical risk item is used to illustrate the model functionality. Specific improvements are proposed which can handle storage operation, redundancy schemes, maintenance schedules etc. in the scheme. These are useful for modelling actual engineering process operations which often employ such measures. The chapter also highlights the relative merits and the potential shortcomings of the proposed method.

4.2 *Overview of the proposed systems framework*

For large scale process systems, most analytical methods are intractable and often a simulation based method must be resorted to. A new systems based modelling approach is proposed in this section and the risks associated with mineral processing are modelled along the lines of this new paradigm.

The general structure of the model is composed of several layers of nested subsystems as shown in Figure 4.1. Each component in the practical engineering system can be abstracted to represent a risk item with given inputs and outputs. The

“in” and “out” flows can be manipulated suitably by some flow controls. Other associated components which affect the risk item are the lifeing engine, health of the specific risk item etc. The detail of abstraction would depend on the available data and the impact of each component on the overall risk profile. Most cases would include constituents where a large failure or breakdown would lead to significant financial losses. These failures might be owing to unforeseen events like flooding, earthquake etc. or owing to mechanical breakdowns. The risk modeller can apply his/her own discretion in more involved situations. For example some components might be known to fail regularly and hence the engineering team has a specific preventive maintenance strategy in place to mitigate the risk of major breakdowns. Therefore in such situations, the standard failure distributions would not be applicable to those components, as the maintenance strategy ensures that the item is operational almost all of the time. Depending on this, the modeller may or may not include this item in the risk calculation model.

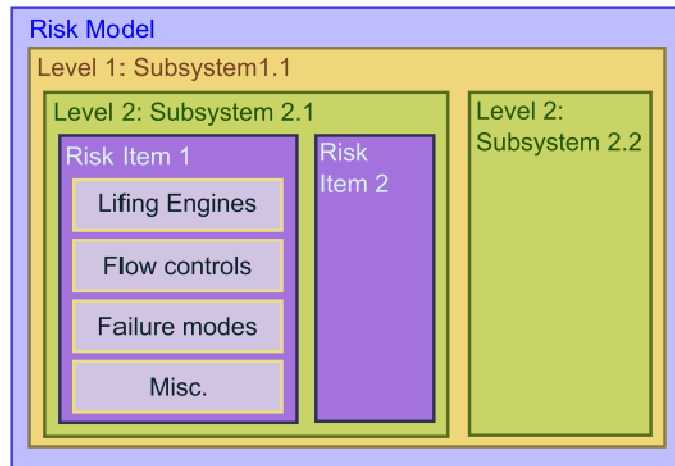


Figure 4.1: Hierarchical scheme of the risk model

Figure 4.2 shows the inter-connections of the sub items in each risk item. In the context of a mineral processing plant, risk items can be a single unit process, such as the grinding, flotation, filtering process etc. Each would have a set of inputs which might affect the risk profile of the item. However, there need not be any specific mass balance for the processes and the inputs and outputs can be of different nature. Of course, the input has to be mapped to the output using some functional relationship. This is specifically done to facilitate the modelling process in the event where there are unknown process variables or where the detailed knowledge of the processing plant at each stage is difficult to obtain. For the grinding process, a set of

inputs can be the raw unground ore, water, energy etc. whereas, the outputs might be small sized ore, waste water etc.

Each risk item can have multiple failure modes. For example the risk item for the grinding process can have a failure mode owing to mechanical breakdown of the mill, severe earthquakes, accidental fire etc. For multiple inputs and outputs (MIMO) in each risk item, the flow controller is basically a set of discrete logical conditions which map the inputs to the corresponding outputs. The health of each risk item is a variable in the range $[0,1]$ for the duration of simulation and is affected by multiple factors. A value of 0 would indicate that the system is in a failed state and a value of 1 would indicate that the system is in its state of maximum health. Any value in between would represent other intermediate conditions for the health. Each failure mode can independently affect the health of the system. Thus, for example, if there is a fire in the milling plant, then one of the failure modes would become active and reduce the health of the risk item (grinding mill) to be 0. It would then be non-functional for a specific duration of time. In the same time interval, an unlikely event of flooding might occur and thus failure mode 2 would become active to make the health of the milling state to 0. The mill will come online, only when both the failure modes have reverted back from their failed states to their non-failed states.

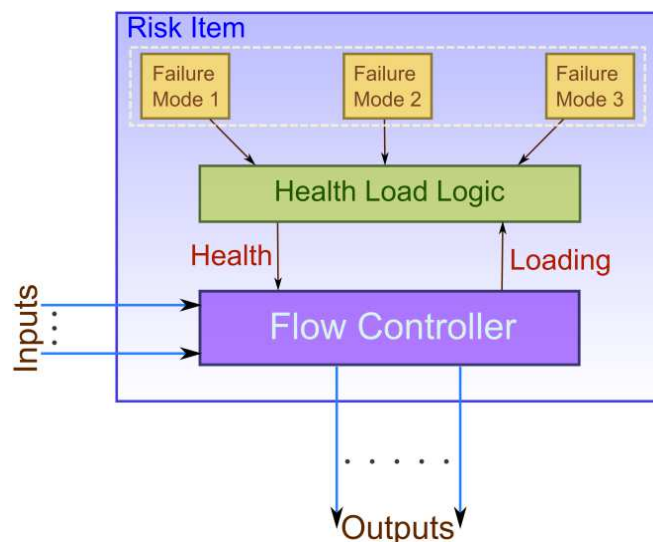


Figure 4.2: Components of each risk item

A variable identified as the “Loading”, affects the health of a risk item and impacts the life consumption of the item. For example there might be more inflows than the

rated capacity and the physical processing equipment might have an accelerated usage leading to a quicker mechanical breakdown. In such cases, depending on the incorporated logic, the loading variable affects the health of the system adversely over the simulated time period. The health of the risk item as calculated from these complex interactions, manipulates the outputs from the flow controller. For the simplest case, the outputs are multiplied element-wise, by the health variable. This would imply that as the health of the risk item decreases, there would be a corresponding decrease in the output flows from the item. Thus, for the milling plant this might translate to lesser amount of finely ground ore coming out because of some faults in the ball mills and a corresponding decreased efficiency of processing the same amount of input.

Each risk item is characterised by three variables which are based on historical data of process operation by the industry and data of past insurance claims. These variables are the Time Between Failures (TBF), Outage Duration (OD) and the Maintenance & Breakdown Costs (MB Costs). The TBF, MB Costs and the OD are represented by probability distribution curves with parametric representation.

There might be situations where there is missing data or no information is supplied for a specific case. In such situations failure data from a standard component can be included in the model and modelling would not be hampered owing to lack of data for a specific case. The operator would generally have detailed information about the reliability of the specific machine components. However, this full dataset might not be available to the external risk modeller or insurer. This is the missing information that is referred to in this context. Also in most cases, all the plant specific information is a trade secret and is not shown to the external risk assessor. Therefore, even though the reliability data of each component is logged by the process operator, they would only convey qualitative information to the risk assessor and the risk quantification has to be done based on such heuristic inputs.

Another situation can arise where there is no information about some component and there are two or more likely candidate solutions. For example, the manufacturer of the ball mill might not be known, but in the market there might be three or four large conglomerates which manufacture this kind of product. Each ballmill would have a different failure probability and repair time associated with it. So, assuming that the ball mill is manufactured by one of them, the composite risk profile of the overall

system can be calculated by doing Monte Carlo (MC) runs taking each of these three or four different makes with equal probabilities in each simulation. If one of the makes is more commonly installed than the others, then it can be drawn more often (i.e. assigned higher probability than the others) from the set for the MC runs. This is an additional advantage of using this modelling methodology over other mass balance based methodologies.

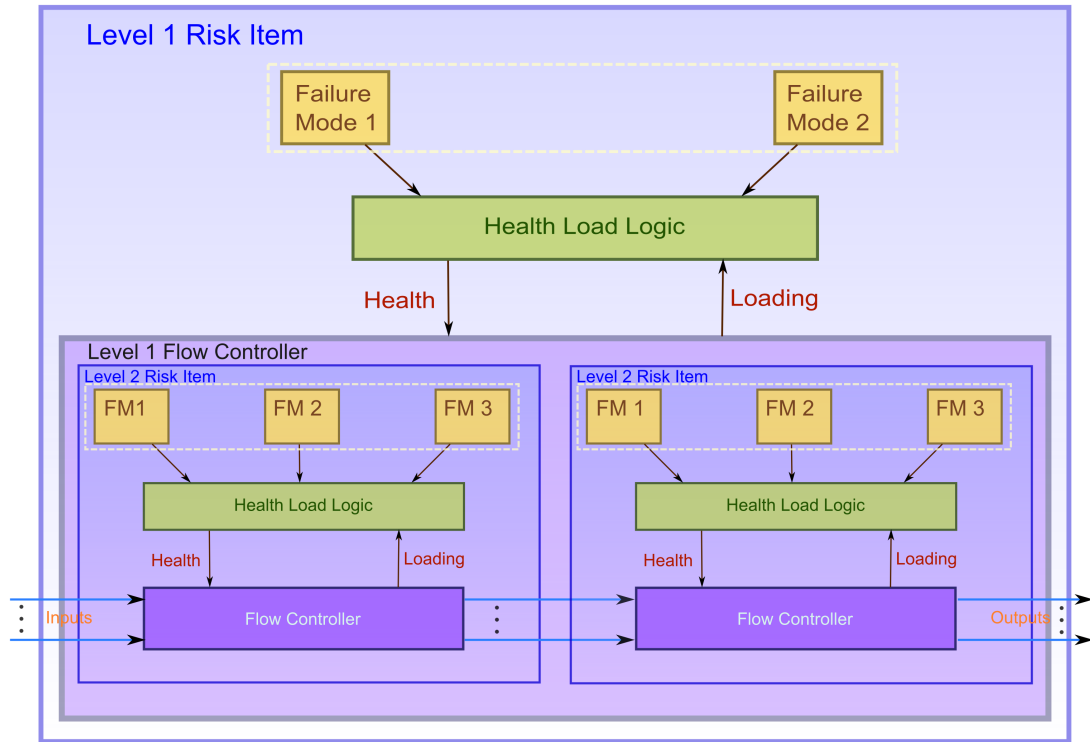


Figure 4.3: Modelling risks in the present framework that can affect multiple risk items

Due to the hierarchical structure of the model, certain risks which can affect a large number of equipment at once can be easily incorporated. Figure 4.3 shows such a scheme within the present modelling framework. The Level 1 risk item subsumes the two smaller Level 2 risk items. Essentially the Level 2 risk items are placed in the flow controller of the Level 1 risk item with the inputs and the outputs providing the necessary link between the two levels. The Level 1 risk item might be catastrophic risks due to natural calamities which affect the whole plant operation. For example, severe earthquakes would affect the whole operation and not any one equipment in particular.

4.3 Modelling of a hypothetical risk item

Figure 4.4 shows the statechart diagram of a failure mode. It has two main states: the NoFailure State and the Failure State as shown by the enclosed boxes. The Failure State has two further sub-states LeadTime Sub-state and Maintenance Sub-state. The former indicates the time when the component sits idle after failure and before going for maintenance. This provision is made to account for additional costs incurred by the company during this period, if any. The arrows with a solid dot indicate a default transition. The arguments in the curly braces denote an assignment operation. Thus, when the model is simulated for the first time, it would be in the NoFailure State with Health=1. In the NoFailure State, a TBF (Time Between Failures) value is generated randomly from a pre-specified probability distribution. In a practical setting, the historical data of the failure of an equipment can be taken into account and then a probability distribution curve can be fitted into the data to be incorporated into the model. The transition from one state to the other is indicated by arrows again and these are governed by conditional statements in square braces. These conditional statements are a function of time (more specifically, it is a random draw from the TBF distribution which is done everytime the system enters into the NoFailure state) The system in Figure 4.4 would shift from the NoFailure state to the Failure state after a random number of time steps as governed by the function $g_1(t)$ have elapsed.

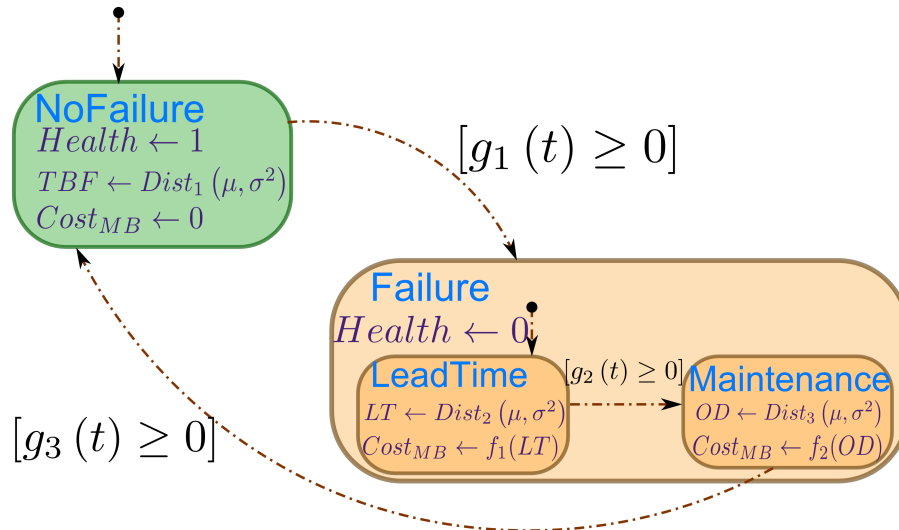


Figure 4.4: State chart schematic of a failure mode

On entering the failure state it would go into the LeadTime sub-state as this is the default transition. After a random number of time steps given by the function $g_2(t)$

(which depends on the randomly generated LT (Lead Time) variable) have elapsed, the system would transit to the Maintenance sub-state. The Maintenance sub-state generates an OD (Outage Duration) variable from another fitted probability distribution curve. Both in the LeadTime and the Maintenance sub-states, the associated costs are calculated through the functions f_1 and f_2 respectively. These functions depend on the time duration that the system is in that state. The system transits to the NoFailure state from the Maintenance sub-state after a specified number of simulation time-steps that have elapsed, as dictated by the function $g_3(t)$ (which depends on the OD variable). There is a hierarchical dependence of the Failure state with its substates (LeadTime and Maintenance). This is done as it is conceptually easier to envisage and the Health variable which is assigned to zero in the parent Failure state is automatically inherited in the two sub-states of LeadTime and Maintenance. The implementation could also have been done without the hierarchy and considering three different states (NoFailure, LeadTime and Maintenance). In such a case, the Health variable would have to be assigned to zero in each of the two states of LeadTime and Maintenance.

The probability distributions are continuous and hence the number sampled from them is a real number. This essentially implies that the simulation is done in continuous time and the minimum time-step is limited by the capacity of the finite bit size of the computer, or the user can specify a small time step depending on the desired resolution of the model.

Figure 4.5 shows the state chart schematic for the implementation of a Health-Load Logic. It has two states: Healthy and Faulty. The system transits from the Healthy state to the Faulty state if the input ($Health_{in}$) from the FailureMode is 0 and vice-versa, if the input is 1. In the Healthy state, the output Health is given as some linear/non-linear function (Ψ) of the input health and the loading from the flow controller.

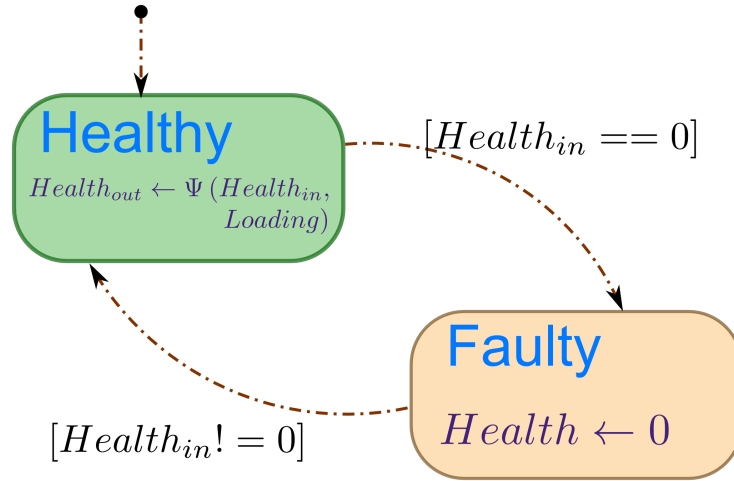


Figure 4.5: State chart based schematic of Health-Loading logic

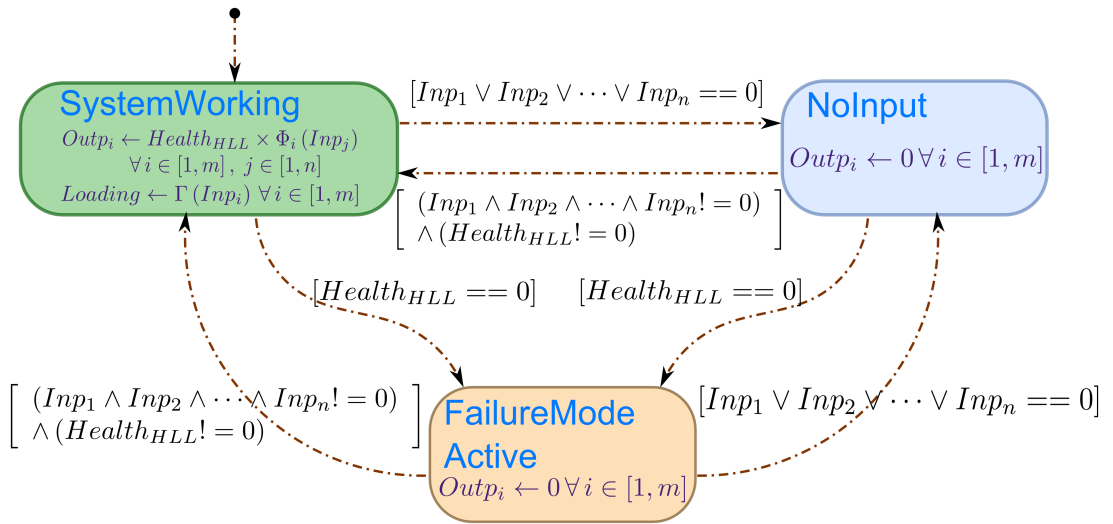


Figure 4.6: State chart schematic for flow controller logic

Figure 4.6 shows the State chart based representation of the FlowController consisting of three states: SystemWorking, NoInput and FailureModeActive. The transitions from the different states are governed by the conditions of the system. The outputs ($Outp_i$) are some functions (Φ_i) of the inputs (Inp_i) multiplied by the Health of the system. The loading is calculated in the SystemWorking state using some user defined function (Γ) which depends upon the inputs to the risk item.

4.3.1 Validation of the model for a risk item with a single failure mode

A hypothetical risk item with a single failure mode is simulated next to verify that the model is working. It has two inputs and one output. The simulation is done for 100 days and the resulting curves are reported in Figure 4.7.

The following values are considered for the simulation. For the failure mode as in Figure 4.4 the state transition functions are given by

$$g_1(t) := t_{state-entry} - TBF \quad (4.1)$$

$$g_2(t) := t_{state-entry} - LT \quad (4.2)$$

$$g_3(t) := t_{state-entry} - OD \quad (4.3)$$

where $t_{state-entry}$ is the time calculated from the instant that the system enters a particular state/sub-state. TBF, LT, OD refers to Time Between Failures, Lead Time and Outage Duration respectively. These are randomly drawn from the probability distributions given by Equations (4.4)-(4.6) respectively.

$$Dist_1 \sim 30 \ln \mathcal{N}(0.4, 0.25) \quad (4.4)$$

$$Dist_2 \sim \mathcal{U}(0, 5) \quad (4.5)$$

$$Dist_1 \sim 10 \ln \mathcal{N}(0.4, 0.25) \quad (4.6)$$

These types of curves are generally used to model failures of different components in reliability based designs. A typical characteristic of this curve is that it is asymmetric and has significant values in the tail of the distribution. The random numbers occasionally generated from the tail portion of the curve can thus model extreme events like large catastrophic failures.

The $Cost_{MB}$ of the system in the LeadTime and Maintenance sub-states in Figure 4.4 is considered as

$$f_1 := LT \times MB_{costfactor_{LT}} \quad (4.7)$$

$$f_2 := OD \times MB_{costfactor_{OD}} \quad (4.8)$$

where $MB_{costfactor_{LT}}$ and $MB_{costfactor_{OD}}$ are constant values for the maintenance and breakdown cost expressed in units of thousand £/day. The duration for which the equipment breaks down, multiplied by this factor gives the total costs incurred for the failure. These are taken as 10 and 5 thousand £/day respectively, for illustrative purposes.

The function (Ψ) in the Health-Loading logic in Figure 4.5 is taken as

$$\Psi := Health_{in} - 0.1 \times Loading \quad (4.9)$$

The function (Φ) which relates the inputs to the outputs in Figure 4.6 is taken as

$$\Phi_1 = 5 \times Inp_1 + 10 \times Inp_2 \quad (4.10)$$

The loading (Γ) is defined as the sum of the ratios of each input (Inp_i) to its rated input ($Inp_{i_{Rated}}$)

$$\Gamma = \sum_{i=1}^n \frac{Inp_i}{Inp_{i_{Rated}}} \quad (4.11)$$

Figure 4.7 (a) and Figure 4.7 (b) show the variation of the two inputs over time. It can be seen that the first input is unavailable from the 25th day to the 45th day and the second input is unavailable from the 60th to the 70th day.

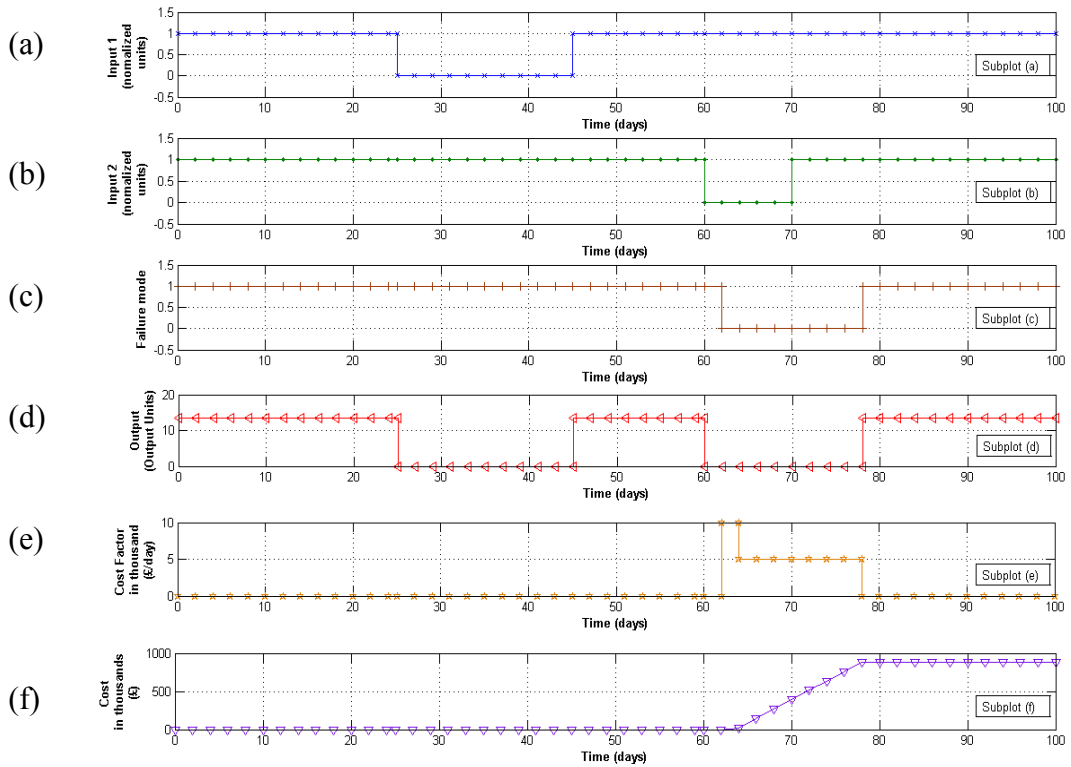


Figure 4.7: Simulation of a sample risk item with one failure mode. The simulation time-step is taken as one day. The granularity of the time-step can be set according to the application requirements.

Figure 4.7 (c) shows that a failure occurs sometime after 60 days and the system recovers from failure before the 80th day. Figure 4.7 (d) in the red curve shows the output over time. As can be seen, for the periods during which there were no inputs,

the output was zero. Also during the failure mode there is no output. Figure 4.7 (e) and Figure 4.7 (f) show the cost factor and the total cumulative cost over time for the maintenance of the particular risk item. As can be seen, the cost factor curve takes non-zero values only during the period when the failure has actually occurred. Also, during the failure period it takes one value for the lead time and the other for the actual maintenance as discussed previously. The last curve shows the cumulative cost of maintenance over time and the final value represents the total cost until that time.

4.4 Model improvements

The model is extended to incorporate a few more features so that it may be applied to a real engineering process system.

4.4.1 Modelling of a storage element

In most engineering processes there are storage elements at different stages. An example might be the stockpile of the input raw ore to a mineral processing plant. This kind of provision acts as a buffer against any upstream faults that might affect the production of the downstream components, albeit for a short time. Therefore, in a case where the mine fails to produce raw ore for a couple of days, the processing plant can still be run using ore from the stockpile. Later when the mine recovers and comes online, it can accelerate its production over a certain period of time, to replenish the stockpile. This in effect acts as a natural hedge against some of the risks which might affect the process operation.

In the present modelling philosophy the storage element is modelled as a tank which has a maximum and minimum level. An analogy can be drawn with a physical water tank, which has a certain capacity and which can deliver a specified outflow rate without having any inflow as long as there is sufficient material in the tank (i.e. the level of the material does not fall below the lower level of the tank). The tank can be modelled on the same lines as the previous risk items with suitable modifications to the flow controller logic. The tank would have failure modes because of large catastrophic events like floods (which can inundate the storage area) or earthquakes (which can damage the structural components of the storage space). The health associated with the tank would be influenced by these failure modes. Depending on the specific application area, there may be other events which affect the normal

operation of these storage elements. For example, the tank might be physically a large open area of land used for storing intermediate products of the process. If they are precious metal ores or ones that can be easily consumed in the household, there is a possibility of theft. Another example can be that of leakage of a tank, assuming that it contains liquid products. In the context of the present work, these events cannot be classified as failure modes as they do not abruptly stop the operation of the tank (although they can be technically classified as functional failures). These would essentially affect the health (which is a time varying quantity) or instantaneous storage level of the tank. To incorporate these effects into the risk model, the functional mapping in the Health-Loading-Logic must be appropriately modified.

A sample schematic of the tank element is shown in Figure 4.8.

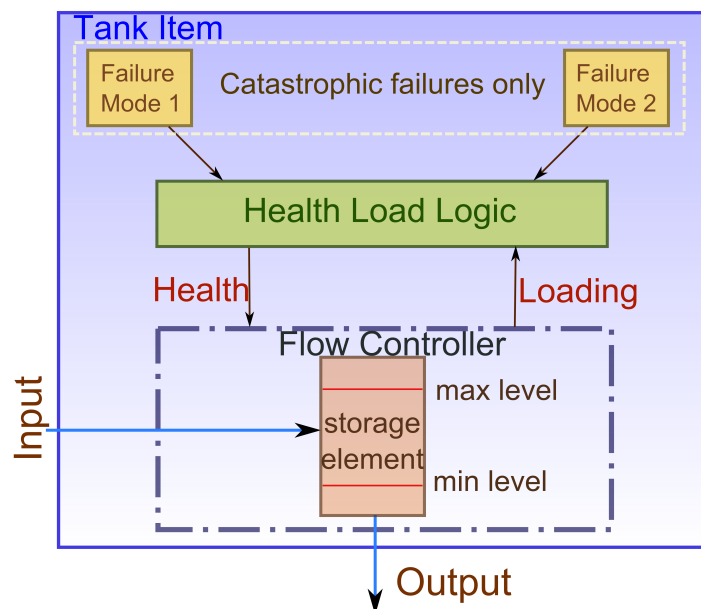


Figure 4.8: Schematic of the tank element

The state chart implementation of this is along the lines of the original modelling paradigm as described earlier. The flow controller is modified to include a minimum and a maximum tank level along with the current level of the tank. The current level of the tank is manipulated based on user specified logic. Another important difference of this item from the previous risk items is that the loading component which goes from the flow controller to the HLL is set to zero. This is because of the assumption that there are no moving mechanical components and, therefore, the degradation in the health of the tank is not dependent on the amount of loading that is impressed upon the tank at any given time.

Figure 4.9 shows the time evolution of the various parameters during the simulation of one run of a hypothetical tank item. Figure 4.9(a) shows the input to the tank over a period of 100 days. The input falls from 1 to 0 on a couple of occasions, between the 5th and the 10th day and between the 40th and the 42nd day. The rated input is one, which is also the rated output. Between the 25th and the 28th day, the input is double the rated input. This would go towards filling up of the tank in case its level is less than the maximum level. There is one catastrophic failure mode which affects the tank. The plot in Figure 4.9(b) shows how the health is affected because of the single failure mode of the tank. Figure 4.9(c) shows the instantaneous level of the tank. The maximum level of the tank is set to 5 and the minimum to 0. The initial level of the tank is set to 2. The simulation starts with this initial level of the tank. Then on the 5th day the input goes to zero. However, the tank still goes on producing the output for a couple of days from its own reserve until its level falls to the minimum level. After 2 days the tank capacity exhausts and there is still no input. At this time the output becomes zero. Again between the 25th and the 28th days when the output becomes 2, the tank supplies the rated output, but replenishes its reserves. As a consequence the tank level rises to indicate that the tank is filling up. When the input goes to zero again between the 40th and the 42nd day, the output is supplied by the tank since the tank has enough capacity to provide output for this short interruption.

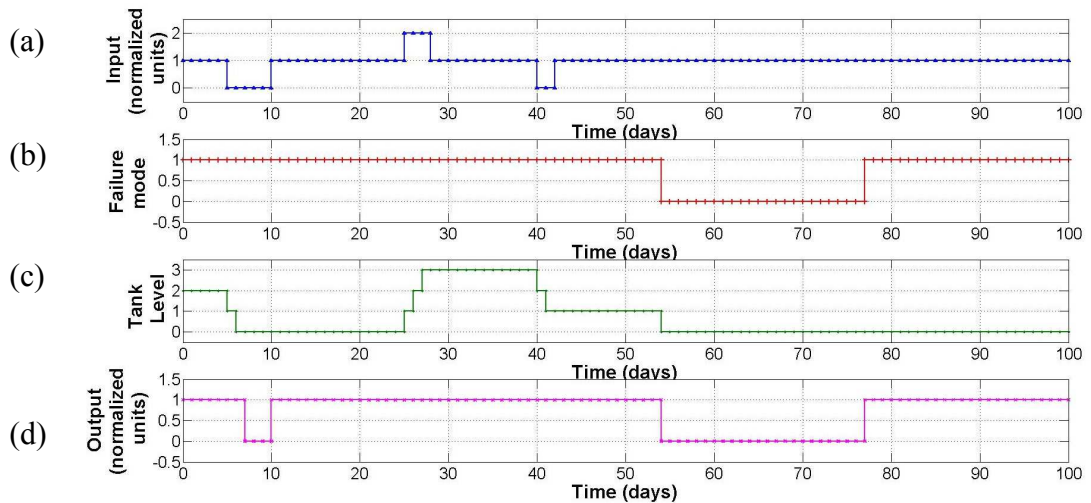


Figure 4.9: Validation of the tank model

The effect of the tank acting as a buffer mechanism to prevent any losses in the output is evident here. Finally when the failure mode becomes active, the tank is affected and it loses whatever level of material that it had. The output is also zero

during this time as the tank is out of operation. This shows that the implementation of the tank model is working as expected.

For demonstrating the tank model, the simulation time step has been taken as 1 day and therefore the obtained tank levels seem to increase or decrease in a discrete, stepwise fashion. However, as discussed previously, the time step can be made smaller depending on the desired resolution of the model, to have a better approximation of a continuous time process.

4.4.2 Modelling of redundancy schemes

Figure 4.10 shows a supervisory logic and switching scheme, which can be used for modelling redundant components in the process flow model to improve system reliability. Many process systems have backup equipment which come online if the main equipment fails. This does not hamper the process operation and gives sufficient time to the maintenance team to repair the failed component.

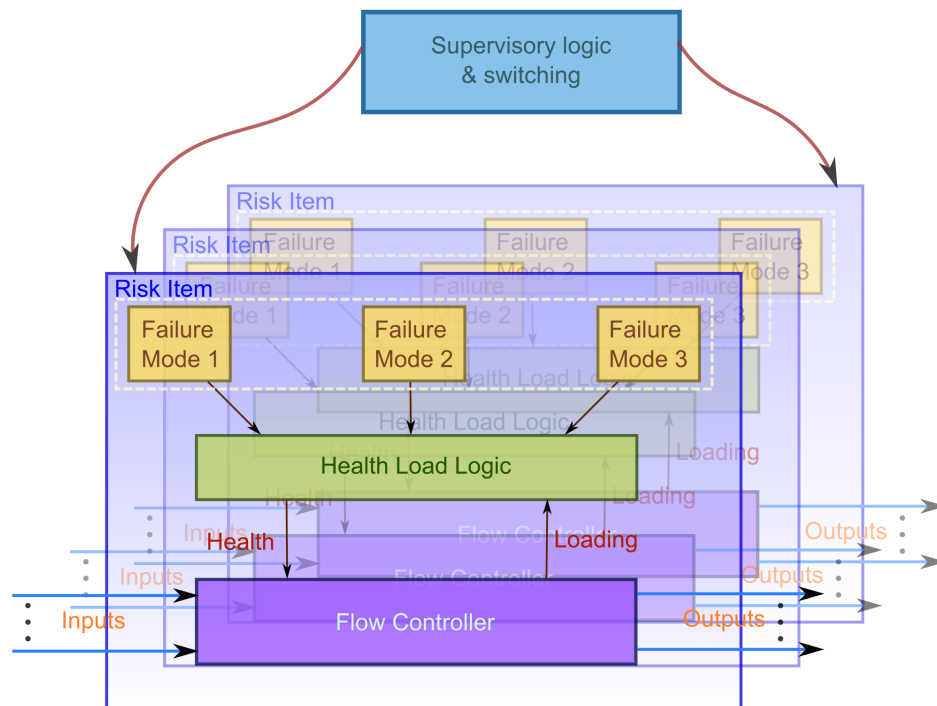


Figure 4.10: Modelling a redundancy scheme

The scheme employs two backup items in the event of failure of one item. The first backup item B comes online if the main item A fails. If both items A and B fail, then item C comes online. The actual logic of this switching between different backup components is achieved by the supervisory logic and switching block. Figure 4.11 illustrates the state chart schematic.

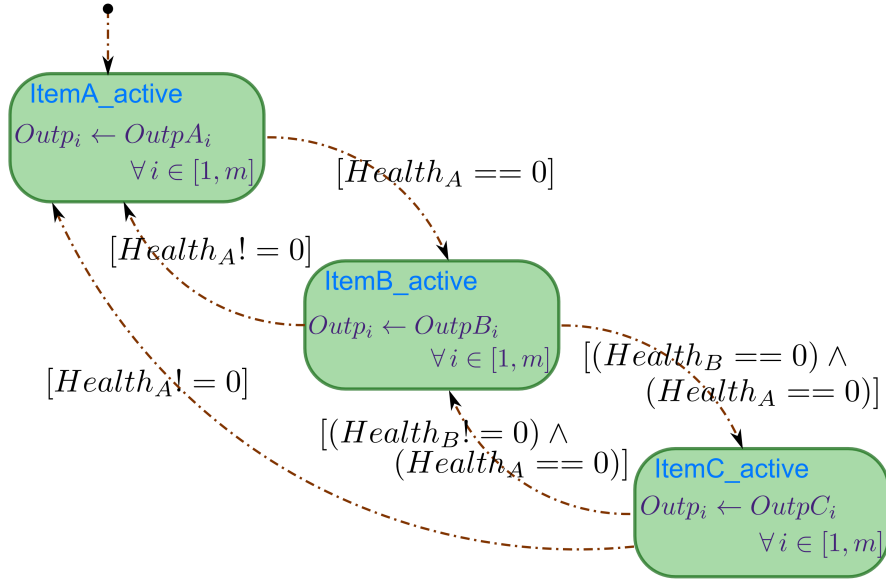


Figure 4.11: State chart schematic of supervisory logic for the redundancy scheme

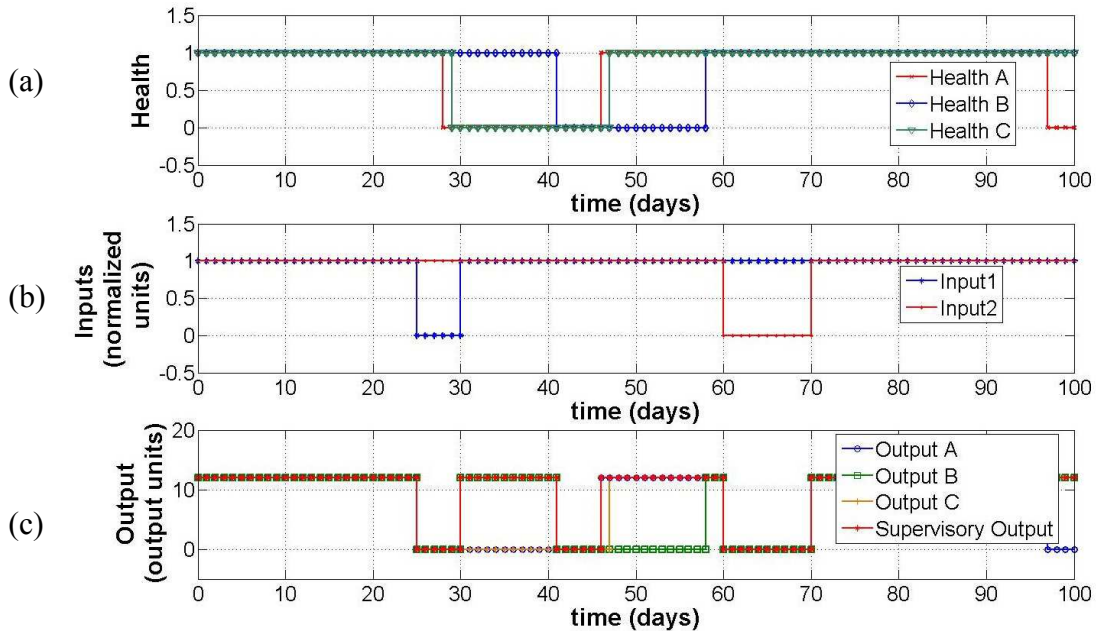


Figure 4.12: Validation of the redundancy scheme modelling

Figure 4.12 illustrates the effective working of the redundancy scheme. Figure 4.12 (a) shows how the Health variable of the three systems evolves over time. All the three systems are connected in parallel driven by the same input and evolve in time simultaneously. The supervisory logic decides which of the three systems should be chosen for the output. All three systems are out of operation (zero health) for specified time intervals. Between the 41st and the 46th day, all of the three systems are concurrently out of operation. Therefore, not only the main equipment, but also both the backup equipment fail during this time. As expected the output curve as

shown in Figure 4.12 (c) is zero during this time. The only other two times when the output curve is zero is when either of the two inputs become zero. Apart from these three time periods, the system is able to maintain some output as at least one of the three risk items in the redundancy scheme is working.

4.4.3 Incorporating maintenance schedules

Most risk or reliability models consider constant maintenance and repair schedules which make it easier to calculate the reliability in an analytical or a semi-analytical framework. However, this grossly oversimplifies reality, as different maintenance schedules may be followed. Therefore, it is important to capture the effects of these schemes in the risk modelling methodology as these would significantly affect the end result. Since there is a health parameter in the proposed modelling scheme and a corresponding loading logic to manipulate it, a wide variety of maintenance scenarios can be modelled by the tool and their effects on the overall reliability of production can be studied. An interesting trade-off that can be inferred from these is whether it is more efficient and cost effective to have storage units after different stages of operation or whether a carefully designed maintenance plan, with additional labour costs, is better in the long run. This trade-off is not always trivial and would depend on the type of intermediate products in the process, availability of cheap labour etc. For example, if the intermediate products are solids, then some space can be allocated for their storage and the capital costs needed for this will not be too high. However, if the intermediate products are in the form of emulsions or other fluids then a proper leak-proof pressure vessel needs to be constructed and the capital costs might turn out to be prohibitive. In such cases availability of cheap labour costs would mitigate the reliability issue to some extent, by enforcing strong maintenance policies in the process plant.

Figure 4.13 shows the schematic of a couple of commonly used maintenance strategies which are incorporated in the present modelling framework. Owing to the generic nature of the tool, other more complicated maintenance strategies can also be simulated and their effect on the overall system reliability with the associated expenses can be studied. In the first figure, the maintenance team does a periodic maintenance at a short specified constant time interval. The green circles on the time axis indicate the instants where the maintenance has taken place. Owing to maintenance, the Health variable of the system goes up every time. In the second

figure, maintenance is carried out only after there is a breakdown. The red circles on the time axis indicate a breakdown where the health of the system goes to zero. The system regains its health (the Health variable goes up) after there is maintenance as indicated by the green circle.

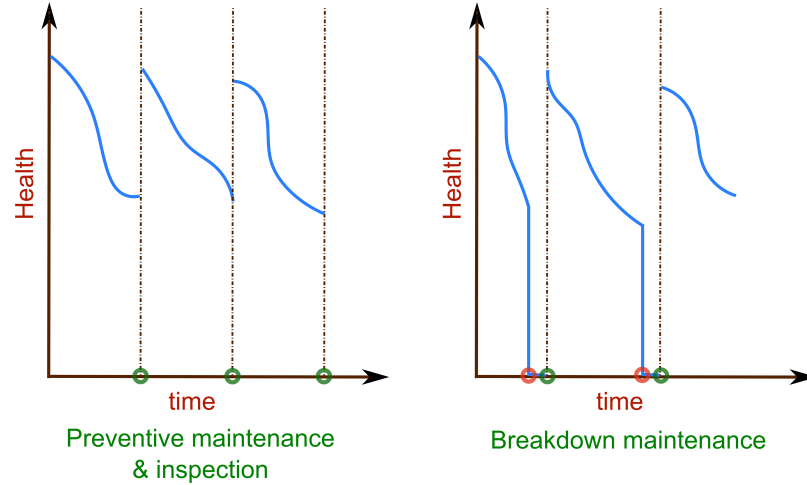


Figure 4.13: Different maintenance strategies

Figure 4.14 and Figure 4.15 show two sample cases where there is a periodic maintenance at a constant interval of 50 and 20 days respectively. Owing to this maintenance schedule the probability of failure of the equipment is delayed in time. Therefore, even though the equipment has the same failure probability, it does not fail immediately after maintenance. However, as is clear from Figure 4.14 and Figure 4.15, a longer maintenance period might result in the equipment failing in between the maintenance intervals, whereas a shorter maintenance period would prevent the equipment from going into the failure states. In both the figures, the output goes to zero on two instances when there are no inputs. Additionally, in the case of longer in-between maintenance times in Figure 4.14, the health goes down to zero during the length of time there is no maintenance. In case the equipment would not have had a breakdown, the preventive maintenance policy would have been applied on the 50th day and the next failure would have been delayed in time. However, since the breakdown has already taken place, the usual maintenance schedule is implemented to bring the equipment back online.

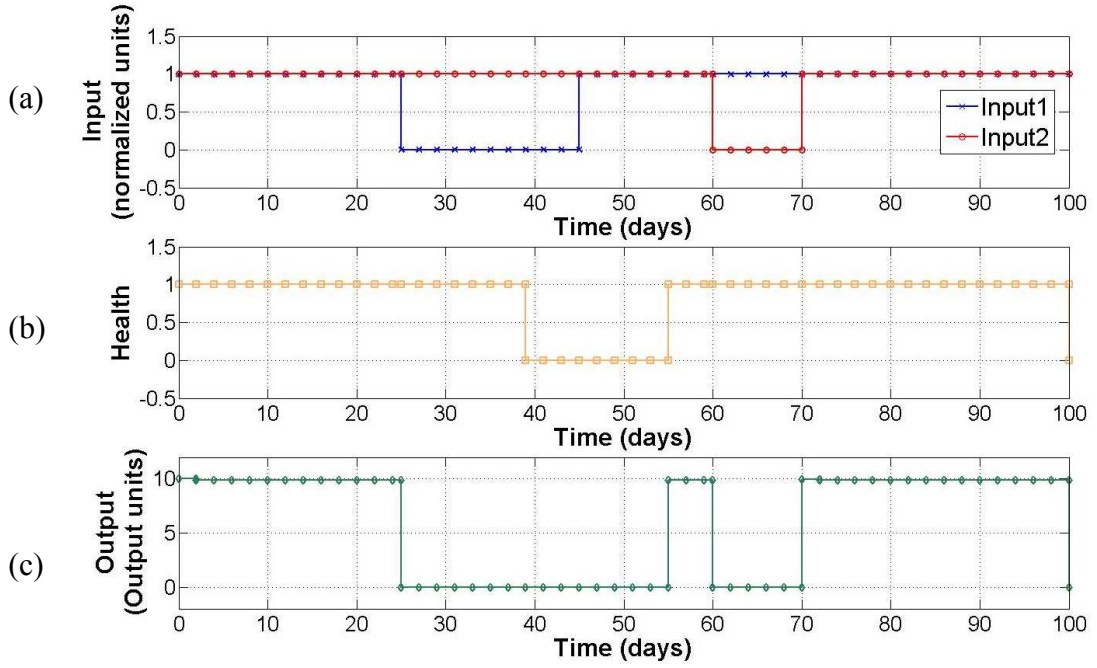


Figure 4.14: Constant interval maintenance with a period of 50 days

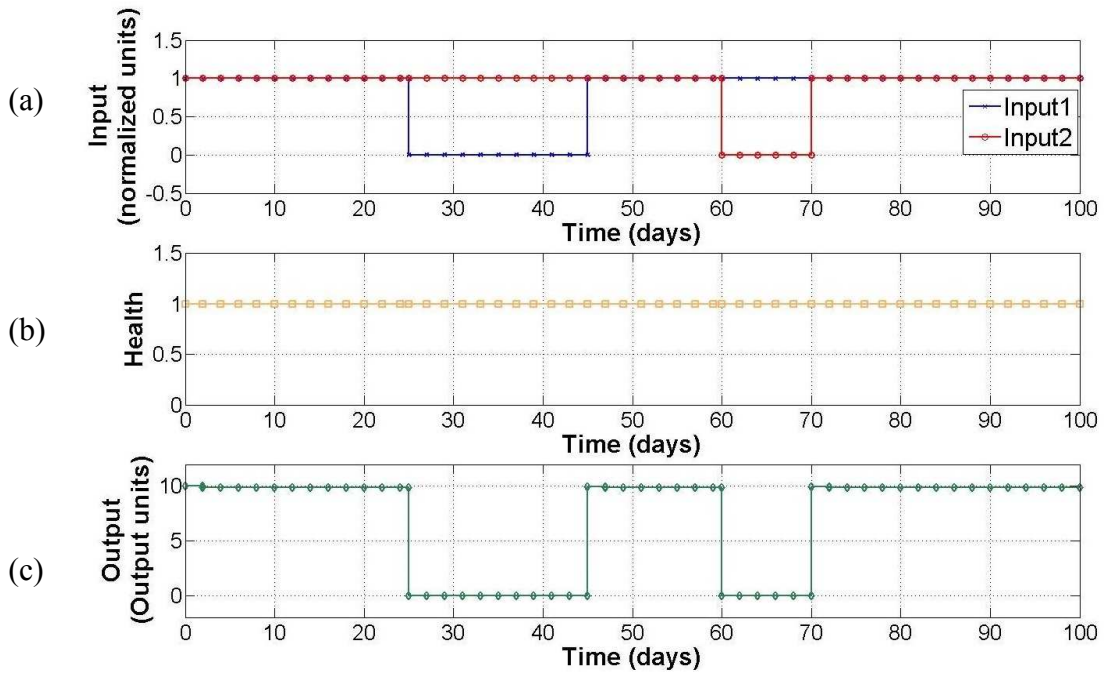


Figure 4.15: Constant interval maintenance with a period of 20 days

4.4.4 Modelling hybrid systems

In the model presented here, the health of a risk item is modelled by a differential equation and coupled with the switching logic of the failure mode to form a hybrid system. A first order differential equation model of the health is considered as in

Equation (4.12). The value of the constant k can be suitably chosen to fit some empirical data and it controls the rate at which the health deteriorates over time.

$$\frac{dH(t)}{dt} = -kH(t) \quad (4.12)$$

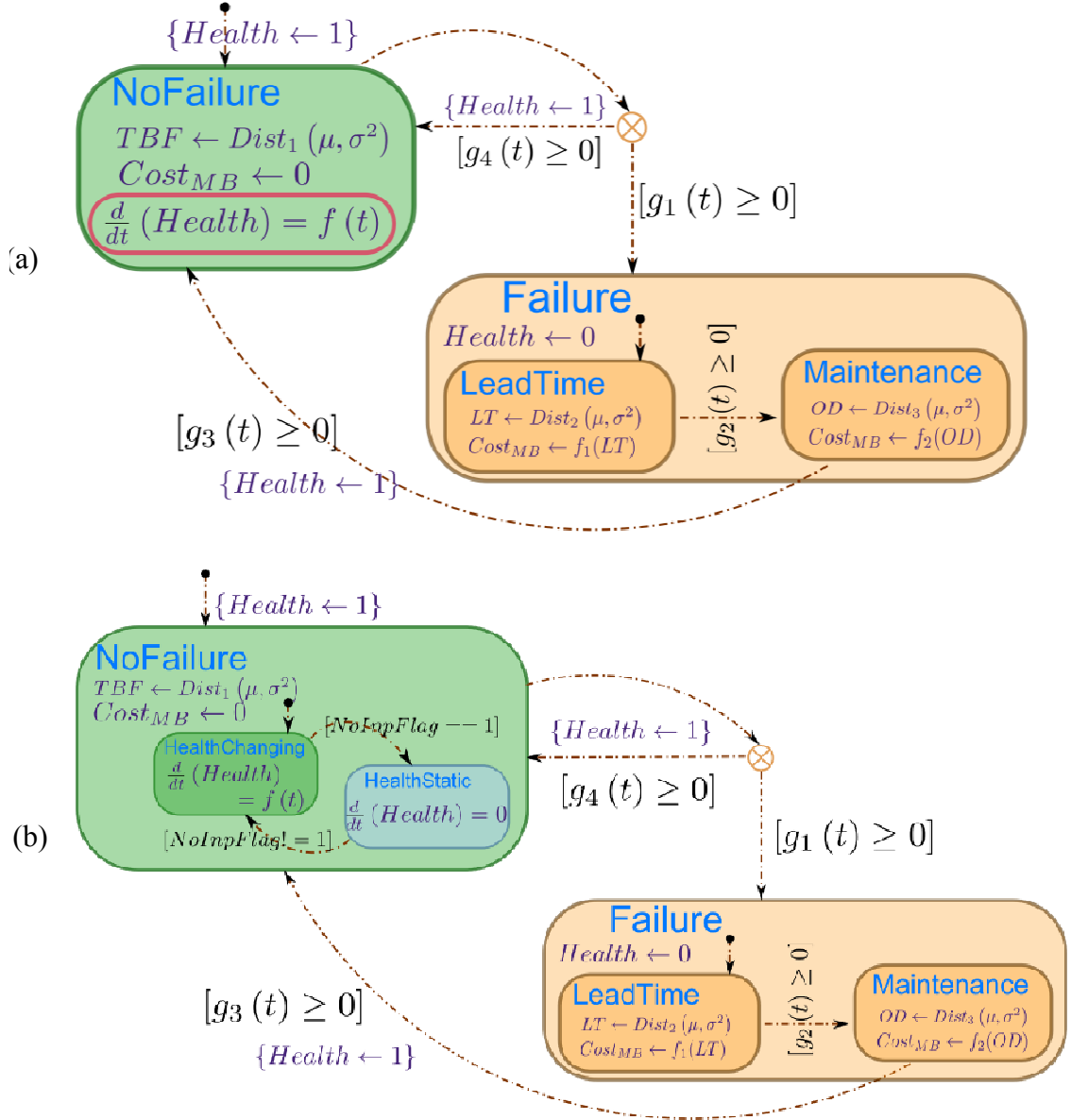


Figure 4.16: (a) State chart schematic of a hybrid system and (b) implementation for simulation

Figure 4.16 shows a state chart schematic of the hybrid system in the failure mode block of a risk item. The encircled red portion is the code for the continuous time

differential equation of the health variable. This coupling between the discrete and the continuous dynamics is achieved by declaring continuous variables in Matlab Stateflow (Matlab Inc., 2010) so that it tracks the derivative of the state. Also a zero crossing needs to be enabled, to accurately calculate the exact values of the variables at points of sudden discrete switching transition. This zero crossing method enables the solver to ‘go back in time’ in the event of a discrete transition and march forward in time again with smaller time steps, to capture the dynamics at the transition point.

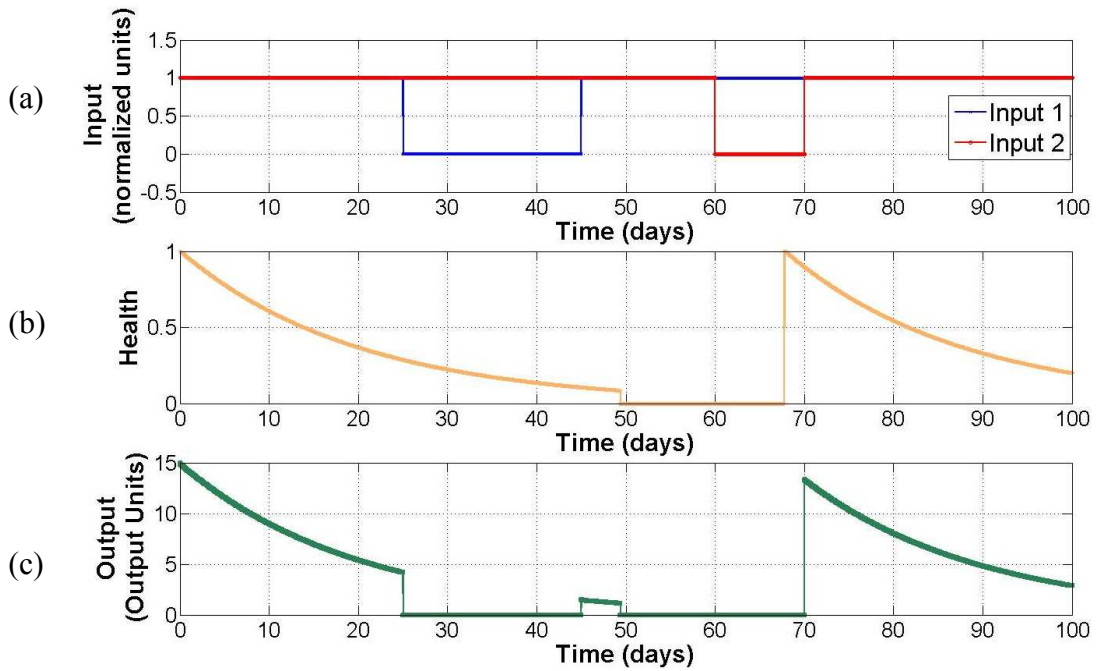


Figure 4.17: Risk item with health modelled by differential equation and switching logic with long maintenance interval

Figure 4.17 and Figure 4.18 show the simulation results with this kind of hybrid modelling. In Figure 4.17 the maintenance interval is 50 days and in Figure 4.18 the maintenance interval is 20 days. As can be seen, there are no inputs during two time intervals. During these two intervals the system output is zero. At the other time periods, the health decreases from its maximum value. Since the output is affected by the health, the nature of the output curve is also similar to that of a scaled version of the health curve. This in effect models the decreasing efficiency of the equipment over time. After maintenance, the health variable shoots up instantaneously and then again follows Equation (4.12) to represent the degradation from the time after maintenance. In Figure 4.17, there is a failure of the system between the 49th and the

67th day. However, owing to the shorter maintenance time in Figure 4.18, the system does not go into the failure mode and its health is propped up after every maintenance. Hence, apart from the time that there is no input, the system produces some output depending on its health and the cumulative output quantity is higher than that in the previous simulation. This also matches qualitatively with the logic that some sort of preventive maintenance at short regular intervals would help to prevent frequent failures of the equipment.

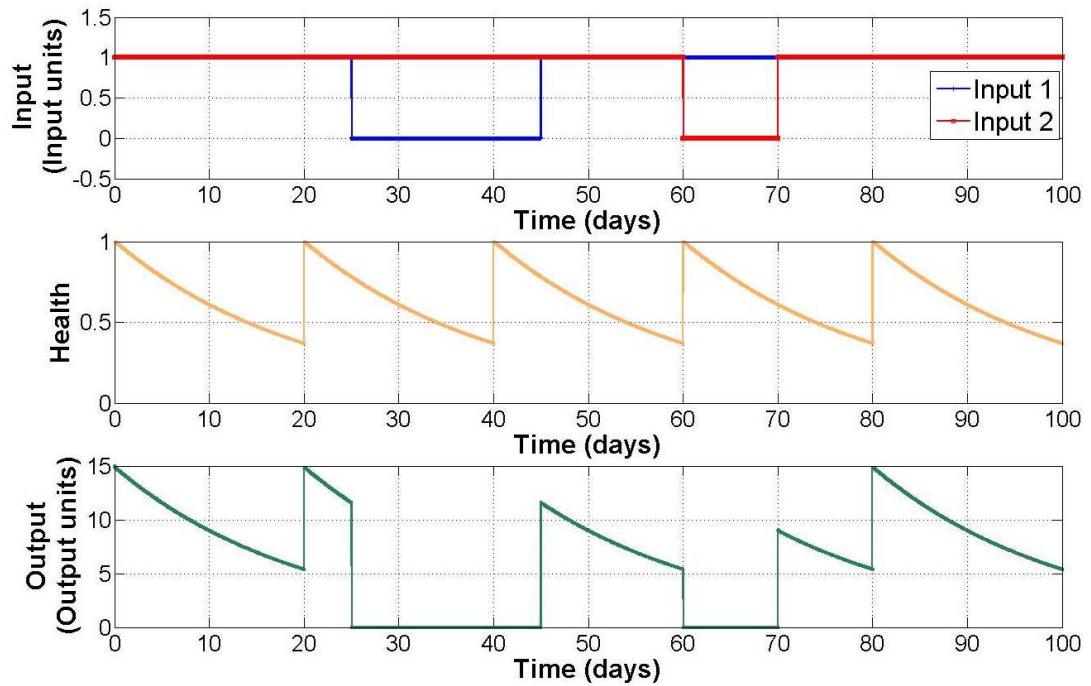


Figure 4.18: Risk item with health modelled by differential equation and switching logic with short maintenance interval

4.5 Relative merits and limitations of the modelling paradigm

The ability of the model to handle incomplete information in the process chain is especially useful from a practical standpoint. The systems level model also gives flexibility to the designer, to refine models in places where necessary, and use coarser top level models where appropriate. The modelling philosophy is not limited to any particular class of systems and can be used to model risks for other large scale systems as well. Unlike other risk modelling methods presented by contemporary researchers, this is a new and different view of risk modelling, which is accomplished by abstracting all the process system components in a systems level modelling framework. The user should be aware of both the advantages and the short

comings of the model to get a better insight into the risk profile for his particular application domain. The next few paragraphs provide a summary of these.

One of the advantages of using this method is that it can be easily extended to incorporate a hybrid systems based modelling (where both continuous and discrete dynamics co-exist). In the present simulation study this is illustrated by considering a differential equation model for the health of a system. However, this need not be limited to the abstractions of the model only. It might well extend to modelling the physical dynamics of the process equipment. These might be useful in circumstances where more detailed modelling of one of the components is desired and the other items can be modelled at a much coarser level of abstraction. For example, the smelter is very expensive and it generally does not have a standby. On the other hand, the grinding mills are comparatively less expensive and might have a standby. Therefore, the failures associated with the smelter would act as a bottleneck and disrupt the production schedule. In such cases it might be advantageous to capture the dynamics of the smelter using underlying physics based differential equations and integrate them into the risk model in this kind of systems framework. This also opens up possibilities of using Dynamic Risk Assessment methodologies as proposed in recent literatures (Podofillini et al., 2010, Podofillini and Dang, 2012).

A few important pitfalls with the use of hybrid systems are that, under some conditions, they might not be well posed and might have issues with uniqueness and existence of solutions (Goebel et al., 2009). Hybrid verification techniques are often used to track whether the system goes into unsafe states during execution (Tomlin et al., 2003). Another issue with hybrid systems is the problem of zeno executions (Goebel et al., 2009). This undesirable phenomenon occurs owing to the system having infinitely many discrete transitions in a finite time (Zhang et al., 2001). Zeno hybrid systems bring in imprecision in the simulation procedure and make the computation time consuming (Zhang et al., 2001).

The model presented here does not take into account the failures owing to human behaviour, management decisions, environmental risks etc. However, the systems framework is very adaptable and these kind of risks can be easily included within the modelling paradigm. A Bayesian network based analysis for environmental and management risks is developed in subsequent chapters and integrated within the systems risk modelling framework. Also a fuzzy logic based mechanism is

developed in later chapters, which helps in tuning the parameters of the risk model based on expert opinion.

4.6 Conclusions

This chapter established the basic structure of the risk modelling framework and validated it with a hypothetical risk item. Various modelling improvements were proposed and the relative merits and limitations of the method were highlighted. However, this model is useful mainly for engineering systems with continuous process variables (e.g. mineral processing). For processes which can be naturally modelled using discrete event systems and do not have continuous system variables (e.g. a mining operation), the risk modelling methodology needs to be suitably modified. In the next chapter, queuing theory based modelling is integrated within this systems framework and validated using a suitable example.

Chapter 5 Queuing theory integration for risk modelling

5.1 *Introduction*

This chapter adapts the systems based risk modelling framework for process operations which can be conceived as a natural flow of discrete entities at specific instants of time. It leverages on concepts of the queuing theory, to represent the underlying process operation using an interconnected system of servers and queues with stochastic generation and service times. This kind of queuing theory based modelling can be used to represent many types of systems like those in production engineering, supply chain management, manufacturing engineering, logistics etc. It is expected that the risk model developed here can easily be adapted to different disciplines with small modifications to serve the specifics of the application. A sample example is shown to illustrate the implementation of the proposed methodology.

5.2 *Integration of systems risk model with queuing theory concepts*

Discrete event simulation techniques are established methods for modelling the flow of discrete entities through various configurations of servers and queues representing a complex process operation. One of the advantages of using such

models is that they are analytically tractable for simple configurations. Therefore, it is easy to obtain mathematical bounds on the average waiting time, queue length etc. However, for more complicated process operations, analytical methods are not feasible and simulation techniques are used to obtain these parameters through bootstrapping or other methods. The idea behind coupling the queuing theory framework with the risk model is that it is possible to retain all the analysis techniques of the queuing theory models and work out the effect of the risk items on the individual subsystems, as explained next.

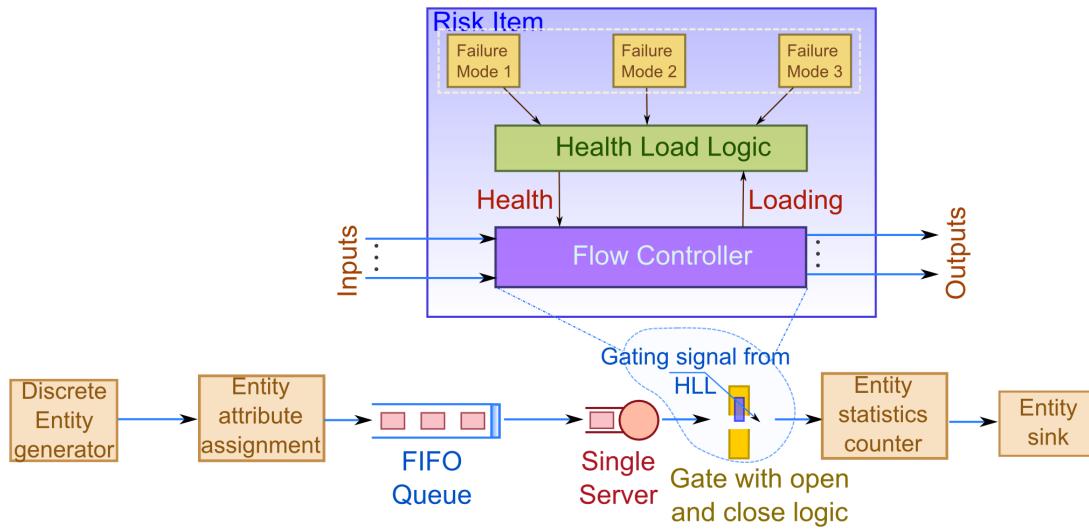


Figure 5.1: Systems approach integrated with queuing theory for risk modeling

The engineering process is modelled using the queuing theory framework and the risk items are modelled in the systems framework. The integration between these two is implemented through the flow controller which regulates a gating signal to control the flow of entities through the gate. Each component of the engineering process system is modelled with servers and queues as shown in Figure 5.1. The discrete entities might represent physical quantities like trucks, excavators etc. or might represent abstract quantities like the decision to drill, blast etc. Each server has a specified service time which represents the time it takes to complete the operation. The server can output the entities only if the gate in front of it is open. The gate receives the signal from a systems based risk model and remains open or closed depending on whether a failure has occurred in the system or not. The risk model can be nested in structure or might affect multiple gates (which represents failures that affect multiple components).

The risk model is composed of various failure modes which signify the different causes which might result in the component failure. It also has a health-load logic which reflects the system health and a flow controller which regulates the flow of the output from the input. In this case, the input is a constant open signal to the gate and the output is the gating signal modified by the different failure modes and the health load logic.

5.3 Validation for a single risk item

For validating the integrated model, a single queue server configuration with a single risk item, as shown in Figure 5.1, is chosen. Three different cases are demonstrated to show that the model is working as expected. The total time for simulation is the insurance period for the risk model which is considered as 100 days for all cases.

5.3.1 Case A: Service time is same as the entity generation time and risk item is absent

For the first case (Case A), the entity generation rate is the same as the service time of the server and there are no risk items considered. This implies that the gate after the server, which is otherwise controlled by the risk item, would always remain open. The intergeneration time of the entity generator is 5 days and the service time is also 5 days. Each of the entity has a data attribute value which encodes the time that the entity was generated. This is essentially a time stamping mechanism and helps in understanding the order in which the discrete entities arrived at the sink and also the total time it has taken to traverse through the system. The system is simulated for a period of 100 days.

Figure 5.2 shows the statistics of the queue over time. Subplot (a) shows that there is a steady departure of the entities from the queue to the server at a regular interval of 5 days. Subplot (b) shows that the number of entities in the queue never exceeds one. This is because the rate of servicing of the entities by the server is the same as that of the generator. So the entities do not need to wait in the queue and are serviced as soon as they arrive. This is also indicated by Subplots (c) and (d) where the average wait time and the average queue length are both zero.

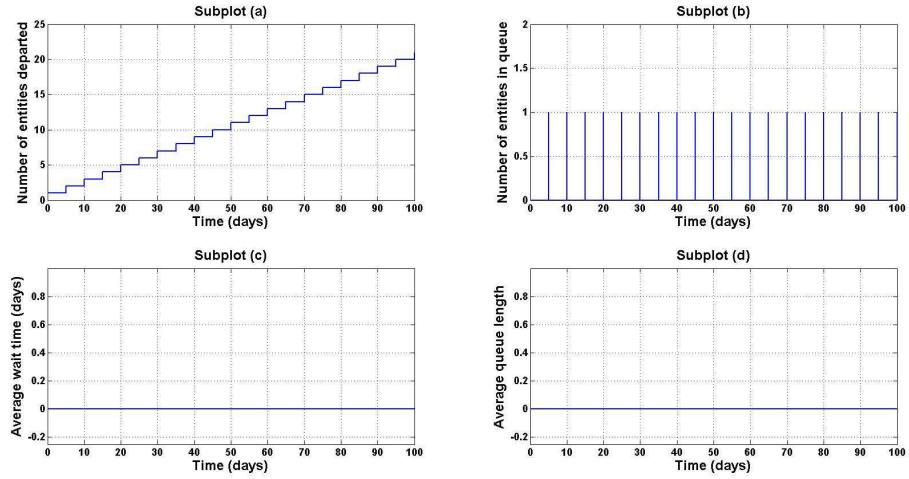


Figure 5.2: Temporal evolution of the queue statistics

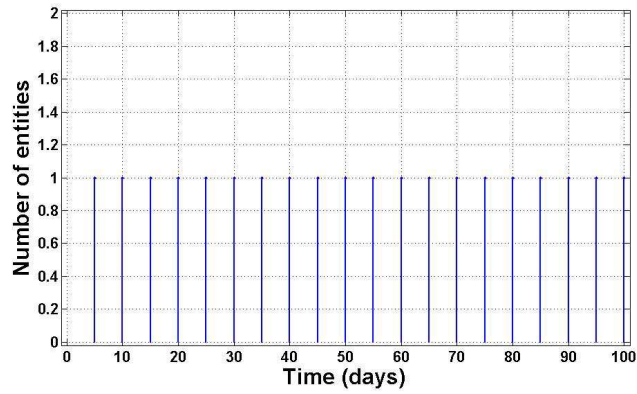


Figure 5.3: Output of the entity counter after the gate

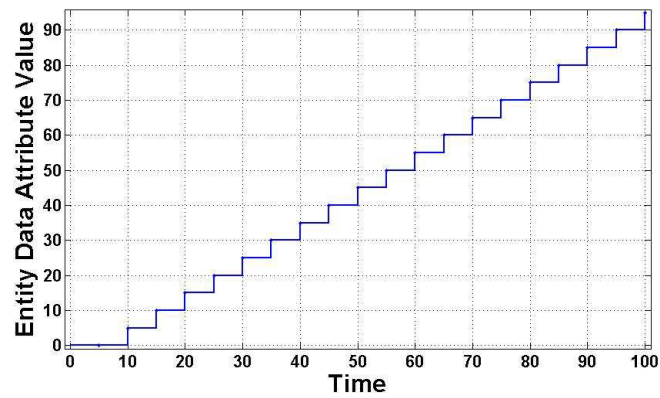


Figure 5.4: Data carried by each entity which comes out of the gate

Figure 5.3 shows the output of the entity counter which keeps track of the entities coming out of the gating mechanism. It can be observed that all the entities come out

of the gate at a fixed interval of 5 days with the starting entity from time instant 5, instead of zero and the last at 105 days. This is because each of the entity is delayed by 5 days owing to the server. This fact can also be corroborated by Figure 5.4, which shows the time stamping of the arriving entities.

5.3.2 Case B: Service time is more than inter-generation time and risk item is absent

For Case B, the service time of the server is more than the inter-generation time. The inter-generation time is 5 days, as before, and the service time is a stochastic number between 5 and 8 days. Owing to this, the entities are generated at a faster rate than what can be processed by the server. This is reflected in the temporal statistics of the queue in Figure 5.5. Subplot (b) of Figure 5.5, shows that the number of entities in the queue progressively increases with time. Subplots (c) and (d) show that the average wait time and the average queue length increase over time. This implies that the entities coming into the queue at later stages need to wait even longer than the ones which came towards the beginning of the simulation.

Figure 5.6 shows the arrival of the entities which pass through the gating mechanism. There is a slight variability in the arrival times of these, owing to the stochastic service time of the server. Figure 5.7 shows the corresponding time stamps carried by these entities. As can be observed from Figure 5.8, the last entity that arrived on day 100, has a time stamp of 60 (was generated on the 60th day). This implies that the other entities which were generated between 60 to 100 days are still in the queue, waiting to be serviced by the server.

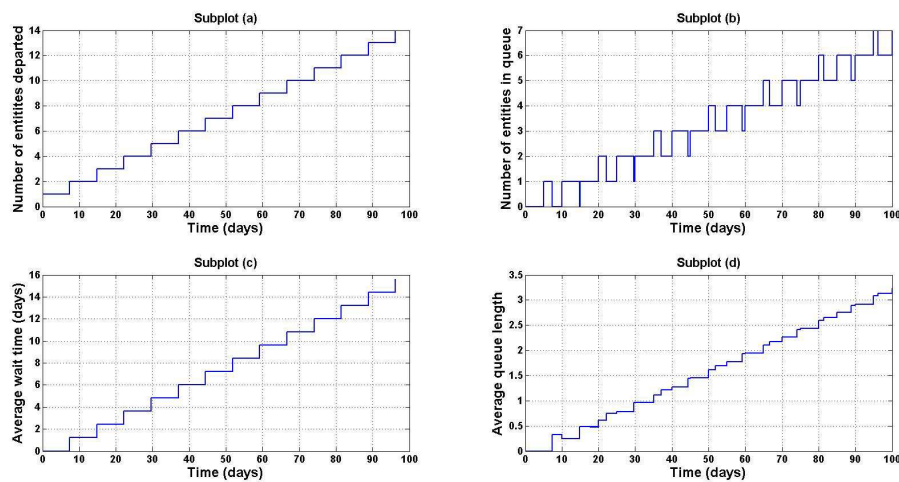


Figure 5.5: Temporal evolution of the queue statistics

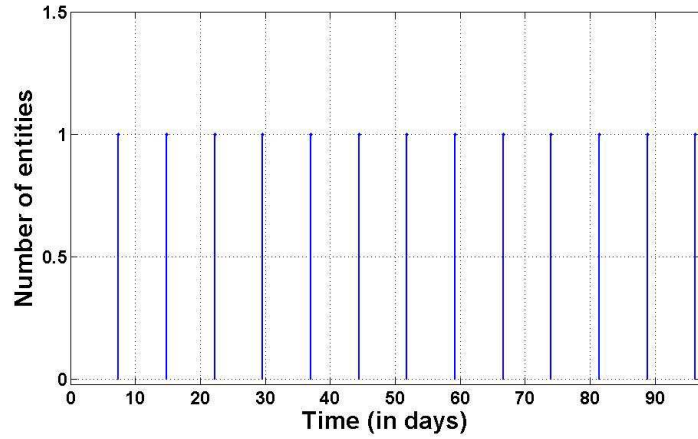


Figure 5.6: Output of the entity counter after the gate

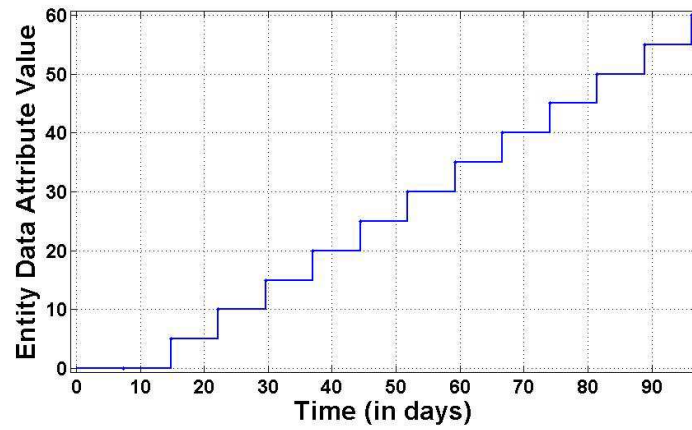


Figure 5.7: Data carried by each entity which comes out of the gate

5.3.3 Case C: Risk item present

Case C considers the same queuing constructs as in Case B, but has a single risk item, which turns the gating mechanism on or off depending on the health of the system.

Figure 5.8 shows the arrival instants of the entities, which pass through the gating mechanism, with blue spikes. Superimposed on these spikes is the gate opening and closing signal (as obtained from the risk item) in red. It can be observed that the entities come out of the gate only during the times when the gate is open (i.e. the red signal has a value of one). Figure 5.9 shows the corresponding time-stamps of the entities. It can be observed that the last entity which comes out on the 100th day, was generated on the 40th day. Compared to Case B, this implies that more entities are in the queue. This also goes with our intuitive understanding, since in this case, the gate was closed at certain instants of time, so less number of entities could pass through it

at the end of 100 days. This can be further supported by looking at the temporal evolution of the queue statistics in Figure 5.10. The number of entities that have departed, as shown in Subplot (a), is much lower than that in Case B (Figure 5.5a). This is because, even though the server has the same stochastic service time in both cases B and C, the entity is retained until the next time when the gate is open. Therefore, no other new entities are taken in for servicing during this time. As can be seen from Subplot (b), the number of entities in the queue grows higher as compared to Case B. Also Subplots (c) and (d) show that the average waiting time and the average queue length increases more rapidly than in Case B.

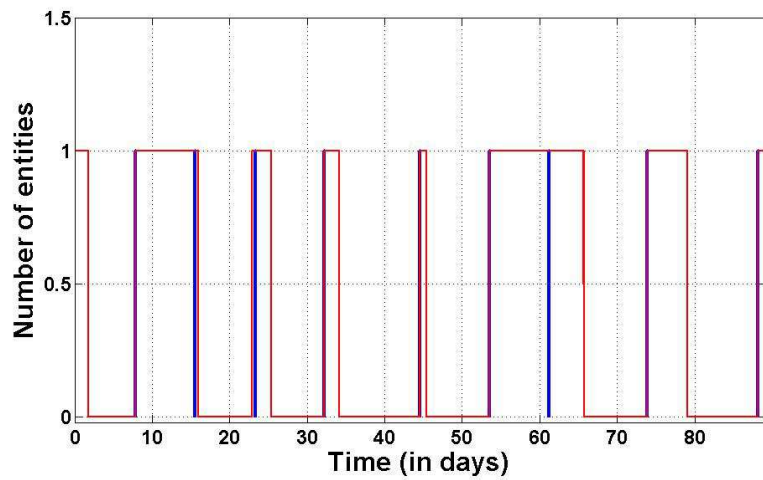


Figure 5.8: Output of the entity counter after the gate with health of the risk item superimposed

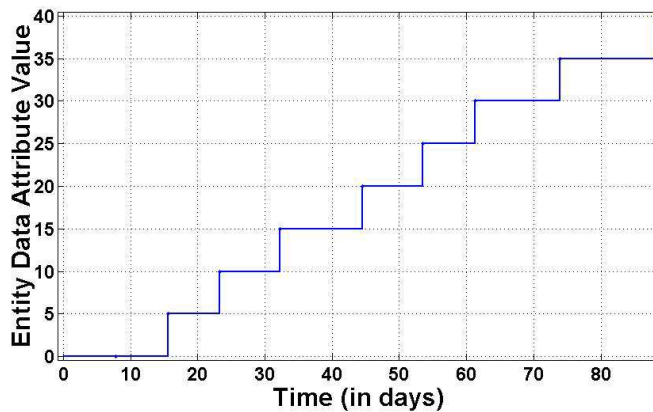


Figure 5.9: Data carried by each entity which comes out of the gate

The simulations show that the risk model was appropriately integrated with the queuing theory based modelling approach. The results are in accordance with our intuitive understanding of the system providing a validation for the modelling paradigm.

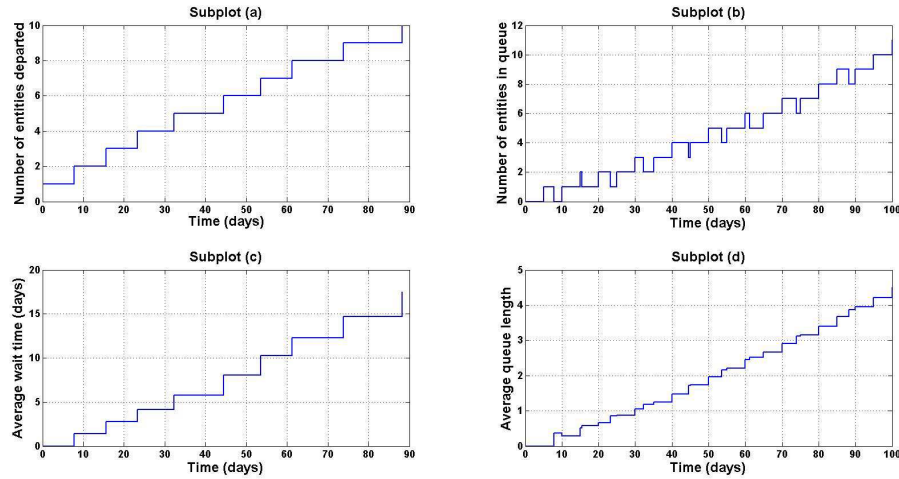


Figure 5.10: Temporal evolution of the queue statistics

5.4 Conclusions

In this chapter, the systems based risk modelling framework was extended to incorporate engineering operations which can be modelled using queuing theory concepts. This is especially beneficial for risk modelling of an open pit mining operation, as this process system can be easily modelled using discrete event simulation techniques. A validation of the methodology is carried out using a single server and queue configuration with different generation rates of the entities.

The systems framework can be designed to fit the system boundaries and have a coarse model in case there is less availability of the underlying process data, or higher resolution where more data is available. However, this still does not address the issue of uncertainty in the parameters of the risk model. The next chapter tackles this issue by introducing a fuzzy logic based method for incorporating expert opinion to handle the uncertainty in the risk model parameters. This uses a fuzzy rule base formed from the domain knowledge of the experts and arrives at a quantitative risk profile. A base case is defined where all the process parameters are known and then

the fuzzy inference system is used to analyse other process systems (with unknown model parameters) which have similar configuration as the base case, but different risk profiles.

Chapter 6 Fuzzy logic based method to incorporate expert opinion in risk modelling

6.1 *Introduction*

The previous chapters introduced the systems modelling framework and extended it to the case for discrete event systems using queuing theoretic formalisms. The systems framework has the capability to group sub-systems together and represent them as one combined entity in the risk model. This is a flexibility in cases where the system architecture is partially known or the reliability parameters of the constituent sub-systems are unknown. In order to address the issue of uncertainty of the model parameters in each of the risk items, the fuzzy logic based mechanism is proposed in this chapter. The implementation methodology is also validated using a simple example.

6.2 *Need for uncertainty quantification of model parameters*

The systems level risk model as described in the preceding chapters is appropriate for calculating financial risk in the case where maintenance data for the failure modes of the various components are available for the given process system. However, for an external risk assessor, like an insurance underwriter, there is a requirement to calculate the risk of a large number of similar processes with very

limited data about each individual process. These risk calculations are especially important in initial surveying actions, when the insurance company needs to get a rough quantitative estimate of the underlying risk, so as to come to a decision regarding whether it is wise to insure the company and the premium that needs to be charged.

Usually, detailed process parameters for a similar operation can be collected in advance by the company and a base case may be defined. Data from various manufacturer catalogues that have maintenance and repair costs for each equipment, can also be leveraged on. The short time-scales in which these insurance decisions need to be made prohibits extensive historical data collection for the new operation. Additionally, the maintenance data and other parameters are often trade secrets and not given out to a prospective insurance company. Generally, a team of underwriters visit the operation site for a period of time and are provided with important parameters: like the capacity of the plant, process layout, redundancy etc. In addition other qualitative information referring to the management practices of the staff, accident records, the safety standards implemented, any known risks (like severe rainfall), the effect of power outages which might damage specific equipment, reasons for delay to expected production schedules, the expected lifetime of the operation etc. The opinions and ratings of the underwriters can be different and to arrive at a quantitative risk profile from these multiple sets of qualitative information, it is important to fuse this information in a framework suitable for the risk model as defined in the previous sections. This is achieved by the fuzzy inference mechanism.

6.2.1 Fuzzy logic as a pre-processing tool of the risk model

Figure 6.1 shows the schematic of how a Fuzzy Inference System (FIS) is integrated with the risk model. This enables the model to take advantage of the heuristic domain knowledge of the experts by encoding it in the form of a rule base. The fuzzy inference engine consists of fuzzification, inferencing and de-fuzzification modules which work together to update the parameters of the risk model. The risk modeling is initially validated for sample process systems where the components and parameters are relatively well known and data is available. If the financial risk profile needs to be estimated for another process system of similar configuration, where only general site specific information is available (e.g. plant capacity, name and rating of the

operating company, geographical location etc.), the heuristic opinion of industry experts can be used. The fuzzy logic method expresses this heuristic expert knowledge in the form of a rule base and adjusts the model parameters accordingly. The original model of the base case process system is run with these modified parameter values to arrive at an appropriate risk profile. This is especially advantageous from the perspective of insurance underwriters who need to quantitatively assess the risk involved in a given operation with very little specific information, utilizing expert knowledge available for the industry.

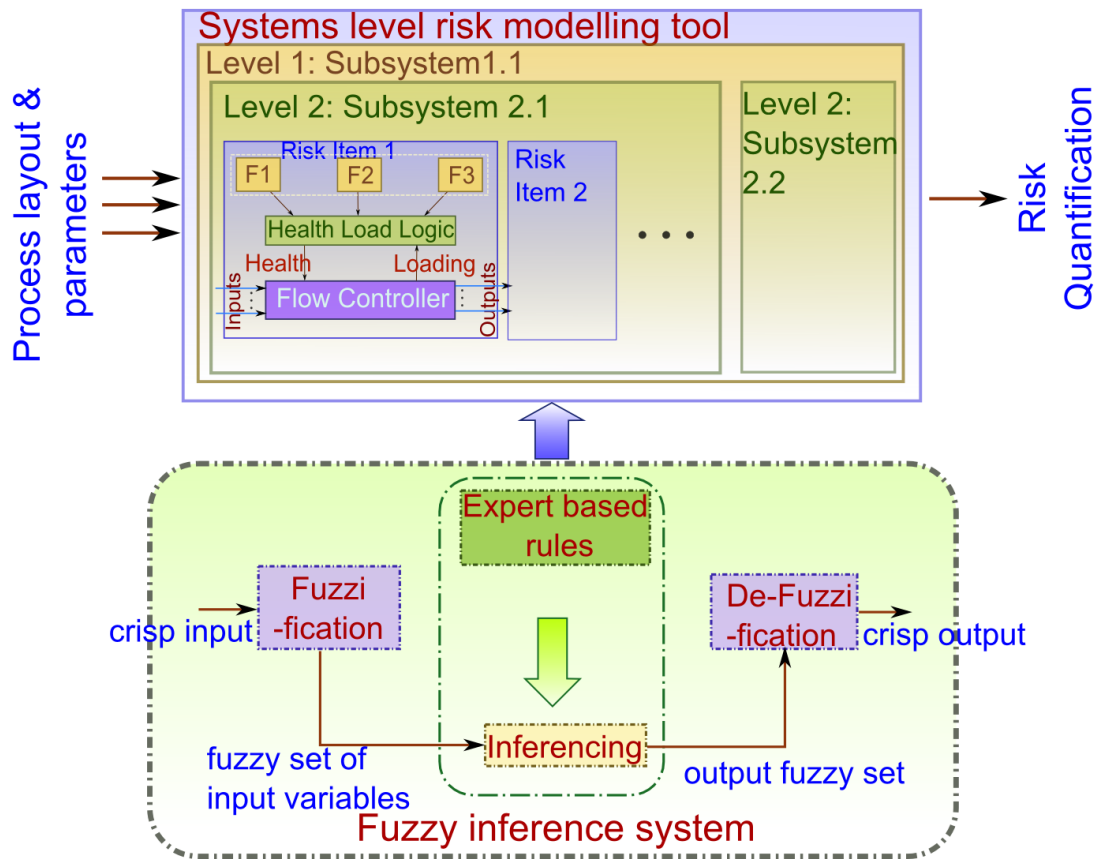


Figure 6.1: A systems level risk model augmented with fuzzy inference

6.3 Fuzzy inference system for incorporating expert domain knowledge

Some gross qualitative indicators which can distinguish the relative level of risk with respect to the base process system operation, as perceived by the insurance underwriters need to be identified. For example, for a mineral processing plant, this can be the company rating, number of years the plant is in operation, location

specific risks that affect the operation etc. Depending on the management practices, the event of failure in the process operation would be much less if it is, say, an ISO 9001 certified company. Again, if the plant is in operation for a longer period of time then the rates of failures are expected to increase (assuming no replacement of machine components) since machinery has a particular lifespan and its frequency of failures would increase over time. Also, the age of the plant might be an influencing factor in cases where the operation is at the end of its lifetime (e.g. a mine in the last few years of its projected life). In such cases, even if some of the equipments break, then the management would be reluctant to make huge capital investment to replace it as the return on investment might not be financially viable. They might rather fix it temporarily and it might work with reduced efficiency for the rest of its life time.

Fuzzy membership functions are allocated to each of these key parameters and a fuzzy rule base is constructed from expert opinions about how the system would qualitatively behave with variation in each of these key parameters. The output of the rule base then manipulates the probability distributions of the OD and TBF of each risk model in the systems level risk modelling tool, so that the risk profile would reflect this deviation from the base case as perceived by the experts. In essence, the fuzzy inference mechanism can be seen as a pre-processor to the risk model. Therefore, while predicting the quantitative risk profiles, the risk model is able to leverage on the prior expert knowledge and the qualitative information supplied by the insurance underwriters in their field visits. This is useful in the decision making process of the insurance company as, in spite of very limited knowledge about the process operation, and considerable uncertainty in the choice of model parameters, a quantitative assessment of the risk profile can be obtained. Detailed explanation on the workings of the fuzzy inference system and its coupling with the systems level risk model is illustrated through an application example in the next section. The way the systems level risk model is updated using the fuzzy inference system, implies that the process flow of the whole operation must be very similar to that of the base case so that the base case model is still an appropriate representation of the new process system with different parameter values for the OD and TBF distributions.

6.4 Application example

The fuzzy inference system as shown in Figure 6.1 is used to adjust the parameters in the systems level risk model and encodes a predefined expert rule base as shown in

Table 6.1. The expert knowledge is encoded in the form of ‘IF...THEN...’ rules and the arguments in the IF statement is denoted by the antecedents and those in the THEN statement as consequents. Three fuzzy variables viz., company rating, location and age of plant are chosen as antecedents and the $OD_{coeff-fuzzy}$ and $TBF_{coeff-fuzzy}$ which are proportional to the adjustment factors of the OD and TBF distributions are chosen as the consequents. All the input and output variables are normalized in the range of $[0,1]$.

Table 6.1: Fuzzy rule base used in the simulation example

Rule no.	IF			THEN	
	Company Rating	Location	Age of Plant	$OD_{coeff-fuzzy}$	$TBF_{coeff-fuzzy}$
1	BAD	LO RISK	YOUNG	MED	LO
2	BAD	LO RISK	MIDDLE	MED	LO
3	BAD	LO RISK	OLD	MED	LO
4	BAD	MED RISK	YOUNG	MED	LO
5	BAD	MED RISK	MIDDLE	HI	LO
6	BAD	MED RISK	OLD	HI	LO
7	BAD	HIGH RISK	YOUNG	HI	LO
8	BAD	HIGH RISK	MIDDLE	HI	LO
9	BAD	HIGH RISK	OLD	HI	LO
10	AVG	LO RISK	YOUNG	LO	MED
11	AVG	LO RISK	MIDDLE	LO	MED
12	AVG	LO RISK	OLD	MED	MED
13	AVG	MED RISK	YOUNG	MED	MED
14	AVG	MED RISK	MIDDLE	MED	MED
15	AVG	MED RISK	OLD	MED	MED
16	AVG	HIGH RISK	YOUNG	MED	MED
17	AVG	HIGH RISK	MIDDLE	HI	MED
18	AVG	HIGH RISK	OLD	HI	MED
19	GOOD	LO RISK	YOUNG	LO	HI
20	GOOD	LO RISK	MIDDLE	LO	HI
21	GOOD	LO RISK	OLD	LO	HI
22	GOOD	MED RISK	YOUNG	LO	HI
23	GOOD	MED RISK	MIDDLE	MED	HI
24	GOOD	MED RISK	OLD	MED	HI
25	GOOD	HIGH RISK	YOUNG	MED	HI
26	GOOD	HIGH RISK	MIDDLE	MED	HI
27	GOOD	HIGH RISK	OLD	HI	HI

The membership functions chosen for the five fuzzy variables are of trapezoidal nature and divided into three sets each. For the company rating, the membership

functions are Bad, Average and Good; for the Location, they are Low Risk, Medium Risk and High Risk; for the age of the plant, they are Young, Middle and Old; for the $OD_{coeff-fuzzy}$ and the $TBF_{coeff-fuzzy}$ they are Low, Medium and High. A sample membership function for the company rating is shown in Figure 6.2.

The fuzzy rule base along with the membership functions can be represented by the non-linear rule surfaces as shown in Figure 6.3 for the time between the failures (subplot a) and outage duration (subplot b). These essentially represent how the different variables of the antecedent part of the rule base correspond to the consequent part in the fuzzy inference space.

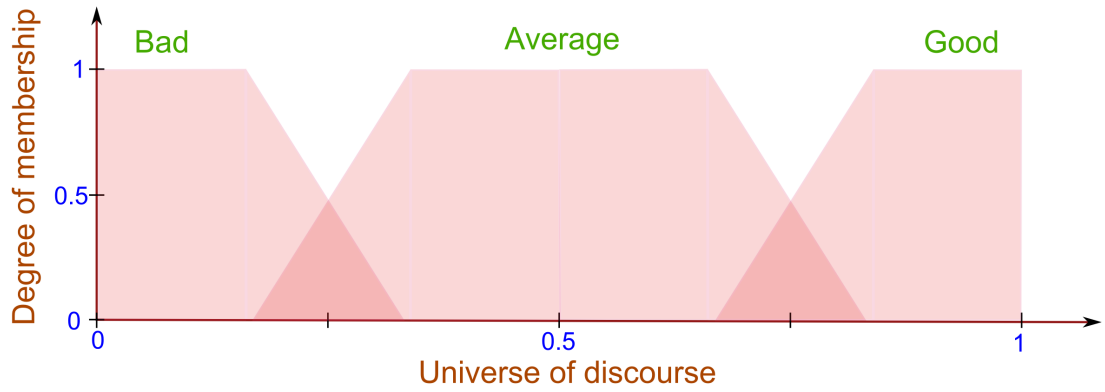


Figure 6.2: Membership functions for Company rating

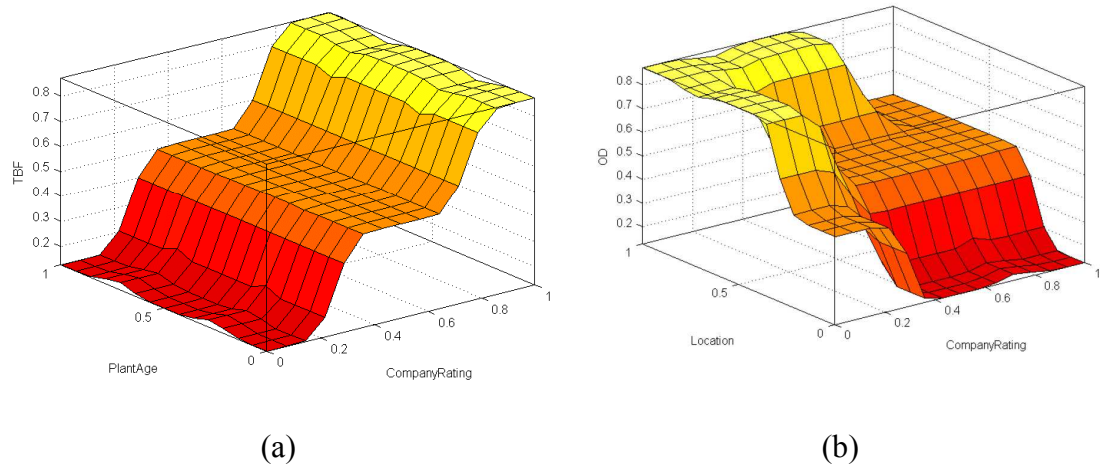


Figure 6.3: Fuzzy rule surface for TBF and OD

The update of the OD and TBF probability distributions of the risk model from the outputs of the fuzzy inference system is achieved by multiplying these distributions with the adjustment factors α_1 and α_2 respectively, which are defined as follows:

$$\alpha_1 = \frac{OD_{coeff-fuzzy}}{OD_{coeff-base}}, \quad \alpha_2 = \frac{TBF_{coeff-fuzzy}}{TBF_{coeff-base}} \quad (6.1)$$

The $OD_{coeff-base}$ and the $TBF_{coeff-base}$ represent the base case values which are taken as 0.5 for both. Therefore, the adjustment factors α_1 and α_2 represent the relative change of the probability distribution values over the base case. The reason for updating the probability distributions using adjustment factors in this fashion is that it is independent of the mathematical form of the probability distribution and is therefore generic in nature. This allows for flexibility in the systems based risk model as the shape of the underlying distributions can be changed to a different one (e.g. from a lognormal to a mixture of lognormal distribution) to incorporate additional data which might be available later. Typical values of α_1 and α_2 would be slightly greater than or less than one. If α_1 is less than one, it implies that the outage durations of the specific risk modelling case is lower than the base case, i.e. risk associated with the plant is lower and vice-versa. If α_2 is less than one, it implies that the time between failures for the specific risk modelling case is lower than the base case, i.e. the plant fails more frequently and therefore has a higher level of risk than the base case.

The reason for constructing the rule base in such a manner can be inferred by reasoning through each of the rules individually. Consider rule no. 9 in the rule base table to understand the logic behind its construction. The general principle of this rule corresponds to the intuitive understanding that if the management practices of the company is poor (i.e. Company Rating = BAD), the plant is in a high risk zone (i.e. Location = HIGH RISK) and is near the end of its lifetime (i.e. Age of plant = OLD), then the plant is more likely to suffer from multiple failures and maintenance and breakdown issues. Therefore in such a case, the time between failures would be much smaller (i.e. failures occur very frequently) and the outage durations would be much higher (i.e. it takes longer times to get the equipments fixed, once a breakdown occurs). Hence the probability distribution of the outage durations for such a plant would have a much wider spread (i.e. frequently result in larger outage durations) than that of the base case (i.e. a plant with Company Rating=AVG, Location=MED RISK, Age of Plant=MIDDLE as can be represented by rule no. 14). Therefore the fuzzy inference system should give an output of $OD_{coeff-fuzzy}=HI$ (as compared to the

base case in rule no. 14 which would give a medium value for this coefficient, $OD_{\text{coeff-base}}=\text{MED}$). Now since α_1 as given in Eqn. 6.1 is the ratio of the two coefficients, it would be greater than one (since there is a high value in the numerator and a medium value in the denominator). As discussed before, the probability distribution of the outages for such a plant would be α_1 multiplied by the probability distribution of the outages for the base case. Therefore, the outage durations sampled from such a distribution would have larger values as compared to the ones sampled from the distribution of the base case, implying that there is a higher level of risk associated with the plant.

For this specific rule, the rule base also outputs the $TBF_{\text{coeff-fuzzy}}=\text{LO}$ (Low). The corresponding value for the base case in rule no. 14 is MED (medium). Therefore α_2 as given in Eqn. 6.1 is less than one (a low value in the numerator divided by a medium value in the denominator). Hence the distribution of the TBF for the plant (which is α_2 multiplied by the TBF distribution for the base case) would result in smaller values than the base case. This implies that the plant is more prone to failures since it has smaller values of the TBFs.

Rule number 9 is one of the worst case scenarios since the company rating, location and age of plant all indicate a high risk operation which is prone to failures. Another extreme case is rule number 19 where all the three inputs indicate that the plant has very low risk associated with it (i.e. Company Rating =GOOD, Location=LOW RISK and Age of Plant=YOUNG). The corresponding values of the $OD_{\text{coeff-fuzzy}}$ and the $TBF_{\text{coeff-fuzzy}}$ is LO (Low) and HI (High) respectively. By similar argument as done for rule 9, one can see that the values of α_1 and α_2 would be less than one and greater than one respectively. Therefore, this plant would have lower outage duration values and longer time between failures than the base case plant, implying that it is more reliable, does not have significant downtimes and is less risky than the base case.

All the other rules lie in between these two extremes (rule numbers 9 and 19) and modify the distributions of the ODs and the TBFs to varying degrees depending on the level of risk indicated by the input variables (Company Rating, Location and Age of Plant).

The rule base only assigns linguistic variables in a logical fashion to reflect the intuitive opinion of the experts. However, this has to be mapped into numeric values in order to update the probability distributions of the OD and the TBF in the systems based risk model. This is done by assigning fuzzy membership functions to the input and output variables of the rule base and using a fuzzy inference scheme to map the inputs to the outputs.

Figure 6.2 shows the membership functions for the company rating which has three levels of discretization (Bad, Average and Good) with a trapezoidal shape. The abscissa (universe of discourse) refers to the values the company rating can take and the ordinate refers to the degree of membership of the input value to any of the three sets (Bad, Average and Good). Now, in the case of binary logic, the input value can either belong to one of the sets (Bad, Average and Good) or not belong to a set (i.e. can take only True or False values). However, in the fuzzy paradigm (which extends binary logic to multi-valued logic), the input value can belong to different sets with different degrees of membership or association. Therefore the trapezoids have partial overlap with each other in Fig. 6.2. To clarify this concept further, consider that the company rating on the abscissa has a value of 0.5. From Fig. 6.2 it can be observed that it is associated with the Average membership function with a degree of membership 1. On the other hand, a value of 0.25 falls in the overlapping region between the Average and the Bad membership functions. Therefore it is associated with both sets with a degree of membership of 0.5 each (as can be seen from the ordinate). Mapping the linguistic variables in this fuzzy way leads to powerful inference tools and allows for incorporating imprecise information in a mathematically tractable framework. Since there is an overlap of the membership function, multiple rules might get activated to different degrees and would be combined by the fuzzy inference system to produce the final output.

The shape of the membership function is purposefully chosen as trapezoidal instead of others like triangular or exponential, to suit the application at hand. For example, consider the trapezoidal membership function for Average. It has a flat region around 0.5. This implies that the experts believe that if the company rating is near about 0.5 then it still fully maps to the Average membership function with a degree of association of 1. As the company rating goes further above or below 0.5, the degree of association falls off linearly. At intermediate values around 0.25 and 0.75, the

input are associated with two sets simultaneously (Average-Bad and Average-Good respectively). At values near zero or one, the input is associated with the Bad or the Good membership function only. A triangular membership function (say with its peak at 0.5) would imply that the input can only have a membership of one for the Average membership function at the exact value of 0.5. If it slightly deviates from 0.5, the degree of association falls off quickly. This allows for incorporating linguistic imprecision in a limited sense and can be useful in other applications such as process control, but is not very representative of the present case of risk assessment and is therefore not considered in the thesis.

Another issue is the choice of the number of membership functions for each input or the levels of discretisation (e.g. three levels for company rating, viz. Bad, Average and Good). Increasing the number of discretization levels results in more granularity and would help in changing the OD and TBF distributions with more precision. However, this also increases the complexity of the rule base and the associated fuzzy inference system. For example, if five membership functions are considered for each of the input variables, then the number of rules becomes 125 ($5 \times 5 \times 5$) as opposed to the present case of 27 ($3 \times 3 \times 3$), which makes it difficult for the expert to come up with such a large set of rules. Moreover, since this inference is used in a risk assessment framework to modify probability distributions themselves, a finer granularity does not improve much in terms of the final outputs of the risk quantification. This is because, the probability distributions incorporate the uncertainty in the OD and the TBF parameters and a slight change in the probability distributions would have little effect on the range of samples drawn from them.

6.5 Testing with a hypothetical risk item

In order to test the fuzzy inference system, a hypothetical risk item which could represent a chemical process plant, is chosen which outputs a single material over time. It has a single failure mode which controls the rate of the output depending on a health variable. The fuzzy inference system adjusts the distributions of the outage duration and the time between failures through the adjustment factors α_1 and α_2 .

Table 6.2 shows the different cases that are simulated to test the applicability of the fuzzy based methodology. A simple case of three input and two output parameters are defined using the company rating, location and age of the plant (in this case, that

of the hypothetical risk item) as inputs to the FIS and the coefficients of the OD and TBF as the outputs from the fuzzy system.

The different scenarios in Table 6.2 are designed so that they can serve as representative cases to quantitatively rank process operations with a wide spectrum of risk levels. All the inputs to the fuzzy system are normalised between [0,1]. Therefore, values near zero for the three inputs imply that the company rating is bad, but the location is in a low risk zone and the age of the plant is relatively low (Young). Values near one imply that the company rating is good, but the location of the plant is in a high risk zone and the plant is near the end of its life (Age of the plant is old). The base case in Table 6.2 is defined with all the input values as 0.5, implying an average company rating with the plant located in a medium risk zone and operating at the middle of its total expected lifetime. Case I is a highly reliable plant where the company rating is good (near to one), the risks owing to the location is low (near zero) and the plant is relatively young (age near zero). Case II is the opposite of this and is a highly risky plant with unreliable operation.

Even though case I and II can be intuitively ranked as the best and the worst, it is difficult to quantitatively estimate where cases III and IV lie. The fuzzy inference scheme overcomes this issue as it can adjust the OD and TBF distributions in the systems model by combining all these inputs and taking the expert rule base into consideration. Running the systems model with these modified probability distributions then gives the risk profile of the plant under consideration.

Table 6.2: Different cases with inputs and outputs from the fuzzy inference system

Case	Company rating	Location	Age of plant	OD coeff	TBF coeff
Base	0.5	0.5	0.5	0.5	0.5
Case I	0.9	0.1	0.2	0.1412	0.8598
Case II	0.1	0.95	0.9	0.8727	0.1285
Case III	0.7	0.8	0.8	0.8505	0.5729
Case IV	0.3	0.4	0.6	0.575	0.425

The corresponding adjustment factors α_1 and α_2 are calculated from the outputs of Table 6.2 and the OD and TBF distributions of the systems level risk model are updated using these factors. The systems level risk model is run for each of these different cases in a Monte Carlo fashion for 1,000 times (this implies that the OD

and TBF distributions are updated for each case before running the simulations). The corresponding mean and standard deviations for the different cases, as obtained from the systems level risk model, are reported in Table 6.3.

Table 6.3: Statistics of 1,000 Monte Carlo simulations of the different cases

Case	Material Output (tons)		Cost (£)	
	Mean	Std. Dev.	Mean	Std. Dev.
Base	5,849.76	385.88	174,326.95	51,486.31
Case I	6,822.70	196.55	12,496.74	5,316.61
Case II	2,994.10	398.96	985,374.41	126,754.60
Case III	5,503.13	501.60	398,583.83	124,202.17
Case IV	5,531.15	434.15	250,543.05	68,073.10

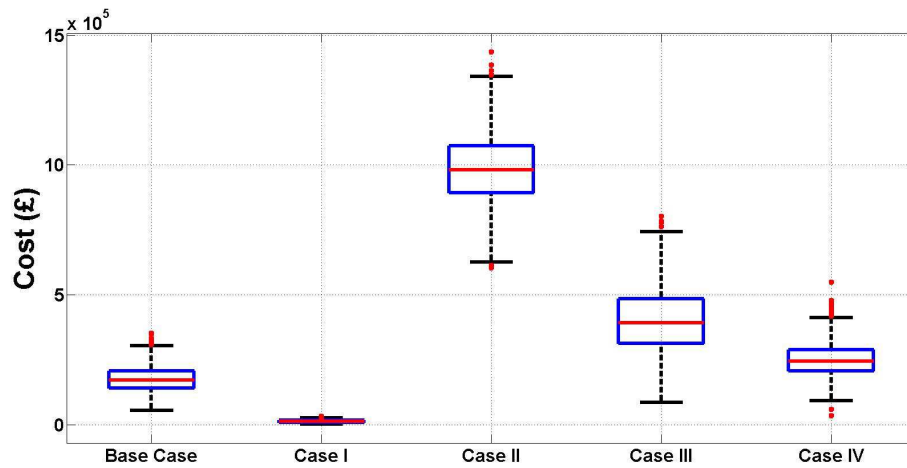


Figure 6.4: Box plots of maintenance and breakdown costs for 1,000 Monte Carlo simulations of different cases

Figure 6.4 and Figure 6.5 show the box and whisker plots for the cost and the material output respectively. Compared to the base case, the maintenance costs is lower and the material output is higher for Case I. This implies better performance of the process operation over the base case. This is in line with the expert inputs in Table 6.2, where, for case I, the company rating is higher, the plant is at a low risk location and its age is relatively new as compared to the base case. This is the best case among all the cases considered in this study.

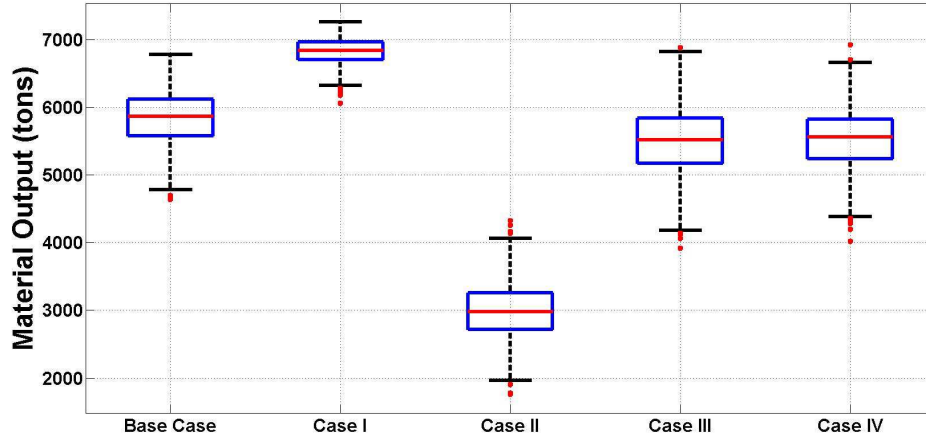


Figure 6.5: Box plots of material output for 1,000 Monte Carlo simulations of different cases

For Case II, the maintenance and breakdown costs are the highest with a large variance. The corresponding material output is also the lowest, indicating that it is the worst performing process among all the cases considered. This is also supported by the expert inputs in Table 6.2, where it has a very low company rating, is situated in a high risk zone and has a relatively older age of operation.

The other two cases (Case III and IV) lie in between these two extremes. Even though they have comparable material outputs, the maintenance and breakdown costs have a higher mean and standard deviation for Case III than for Case IV. This is also supported by the expert inputs in Table 6.1, where the age of the plant for Case III is higher than that of Case IV. The corresponding outage duration coefficient, which is calculated from the FIS is also higher.

6.6 Conclusions

This chapter augmented the systems based risk modelling framework with a fuzzy inference engine, to handle the uncertainties associated with the model parameters. The methodology was tested with a hypothetical risk item and the quantitative results matched with our intuitive understanding of the problem. Encoding the expert opinion in the form of such a fuzzy rule base is important since it helps in eliminating the personal bias and trial to trial variability as commonly found in the heuristic methods. The next chapter introduces a probabilistic Bayesian framework

which incorporates the expert opinion in the form of cause-effect relationships in graphical models. This is more geared towards quantifying the catastrophic environmental risks that have not been taken into account, in the previous approaches.

Chapter 7 Bayesian networks for risk inference

7.1 *Introduction*

The fuzzy logic method, as introduced in the previous chapter, can incorporate expert inputs to infer unknown parameters of the risk model. However, this method cannot deal with partially available information. In other words, it cannot handle situations where only some of the inputs of the fuzzy inference system are specified and the others are missing. This issue can be circumvented by using a probabilistic framework. A Bayesian network (BN) is a suitable candidate for this purpose. This flexibility is desirable from a practical point of view as well, since it implies that the model can make the best possible use of the data to refine the risk predictions. This chapter applies the Bayesian networks for the case of catastrophic risk modelling (slope collapse and flood risks) and integrates it with the previously developed systems based risk modelling framework. Model testing is conducted using a hypothetical risk item and the results are found to be in agreement with our intuitive understanding of the problem.

7.2 *Integration of Bayesian networks within the systems based risk modelling paradigm*

The Bayesian networks integration with the developed risk model is designed to take into account the catastrophic risks. Figure 7.1 shows a systems level risk model which is coupled with a Bayesian network for incorporating the losses associated with risks like flood and related potential slope collapse. It has three separate components, namely, the systems level risk model (which simulates the maintenance and breakdown risks), the loss distribution based catastrophic risk model (which simulates losses owing to catastrophic events with low, medium and high impacts) and a Bayesian network which dictates the relative proportions of the different catastrophic risk impacts in the final output. The Bayesian network therefore essentially acts as a modifier on the loss distribution based catastrophic risk model.

The final risk quantification output which comes out of the tool, consists of maintenance and breakdown losses from the systems risk model (which are large in number, but has lower costs associated with them) and catastrophic events with different impact levels (which are smaller in number, but have higher costs associated with them).

The loss distribution risk model is essentially a set of probability distribution curves which are assigned to the different impact categories of risk. For example, low impact risks might have higher frequencies but less costs associated with them. In the case of medium or high impact risks, the events would be less frequent but would have higher costs associated with them.

The reason for choosing a loss distribution model for such catastrophic risk is that these risks are not very frequent like those owing to equipment breakdown and therefore detailed data would not be available. Also there are many different routes through which a catastrophic risk can occur and it is difficult to model them directly using the systems framework. Therefore a coarse model in the form of probability distributions which can differentiate between different categories of catastrophic losses (low, medium and high impact) is appropriate for such a case. This approach cannot account for the detailed causes of the catastrophic risks but can represent their aggregate effect on the output which is essential to make an informed judgement on the overall risk profile of the system.

When the system is run in a Monte Carlo fashion, in some instances the systems based risk model is simulated and at other instances the loss distribution based risk model is simulated. This is dictated by a fixed probability (P_{cat}) which is assigned to determine the relative occurrence of the catastrophic risk events with respect to total events including the normal maintenance and breakdown events.

The total losses are aggregated to present a multi-modal distribution to represent the losses associated with the system for all the different kinds of risks. Whenever the algorithm simulates the loss distribution based catastrophic risk model, one of the probability distributions is chosen (i.e. either of the low, medium or high impact) and a number is drawn from it to represent the loss. The Bayesian network helps in guiding the algorithm to choose the probability distribution based on the partially available evidence. In other words, if some of the evidence indicates that the impact of the catastrophic risk is low, then the BN will give more preference to the probability distribution associated with the low impact risks and consequently result in a smaller loss value.

The consideration of the fixed probability P_{cat} is justified from an insurance perspective as well. Generally a risk survey report classifies catastrophic risks like hurricanes, floods etc. based on the likelihood of occurrence (e.g. a one in ten year event). From the perspective of the model, this implies that if the model is run ten times, then in nine cases, the systems based risk model should be used to produce the output (which would only result in small maintenance and breakdown events) and in the last remaining case, the loss distribution based catastrophic risk model should be used to produce the output (which would result in a large financial loss). The final aggregate output (of ten simulation runs in this case) would therefore consist of a large number of small maintenance and breakdown events (which is nine in this case) along with a small number of high impact catastrophic risk events (which is only one in this case).

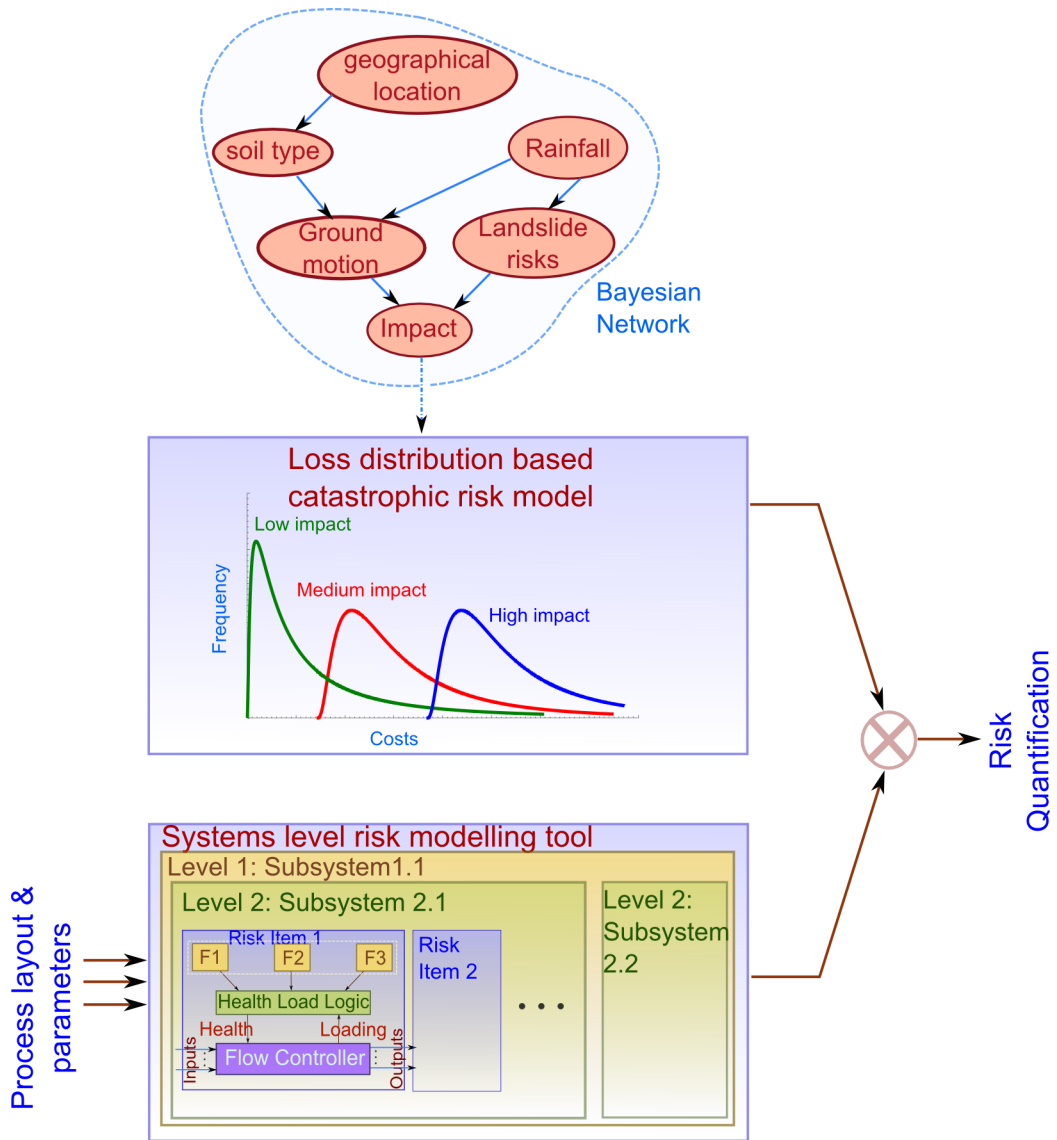


Figure 7.1: A systems level risk model augmented with Bayesian networks

The advantage of such a tool is that it helps to quantitatively distinguish the different levels of catastrophic risks (i.e. within the category of catastrophic risks how many of them would be low/medium/high impact and their associated financial costs). This helps to make an informed decision with respect to insuring the operation. For example, if the simulation results indicate that there is a chance of very high impact catastrophic risks, then a re-insurance should be done by the insurance company itself, so that in the event that the insurance company has to payout for the calamity, it does not become bankrupt.

7.3 Testing of the methodology

In order to test the modelling methodology, a hypothetical risk item with a single failure mode is considered for the systems level risk model. This failure mode

represents the costs owing to maintenance and breakdown. For the catastrophic risk component of the tool in Figure 7.1, loss distributions (with different levels of impact) are constructed for a slope collapse risk. A BN for slope collapse risk is constructed using different environmental variables like rainfall, ground motion, soil type etc. as shown in Figure 7.2. The construction and analysis of these BNs are done using the SAMIAM (Sensitivity Analysis, Modelling, Inference and More) tool (Darwiche et. al., 2001). Each node in the BN (depicted by the oval boxes), represents a random variable which influences the final risk output. The interdependencies between the random variables are indicated using arrows. The text boxes next to the nodes indicate the number of discrete states in the node and the probabilities (expressed as a percentage) associated with each of the nodes. For the case of Figure 7.2 all these probabilities are indicated in green which implies that these are prior probabilities from expert inputs. In cases where evidence of one of the variables is fully known (e.g. in Figure 7.3 it is known with certainty that the site is close to faults), the influencing variables in the oval boxes are indicated in red and the corresponding state has a 100% probability (expressed as a percentage). Based on this evidence, the posterior probabilities of the different variables in the network is recalculated using Bayesian inference. Therefore the posterior distribution of the other unobserved nodes (indicated by the text boxes having green coloured representation of the states in Figure 7.3 and 7.4) has changed from their previous values in Figure 7.2 to incorporate the effect of the evidence.

The structure and the variables to be included in the BN is adapted from Li et. al. (2012) with suitable modifications, to adjust the network for the specific case of slope collapse, for a mining operation. The final output from the BN is the probability distribution over the level of impact of the catastrophic risk.

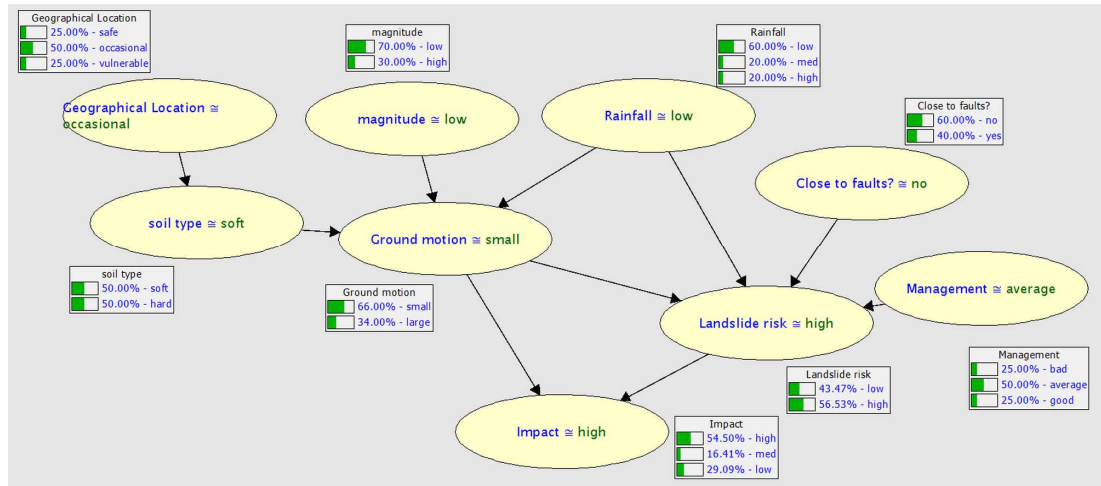


Figure 7.2: Case B: A BN with no evidential knowledge

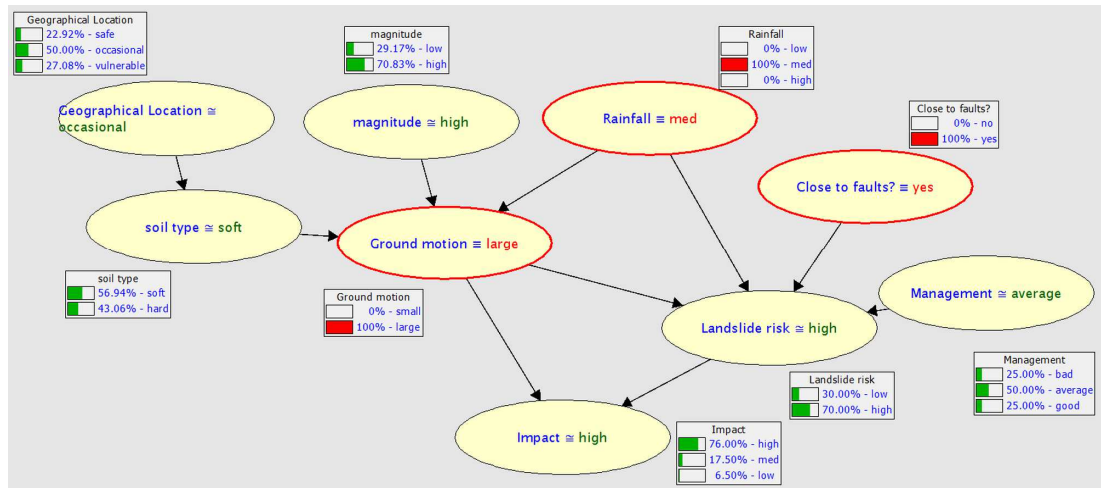


Figure 7.3: Case C: A BN with evidence about high impact risks

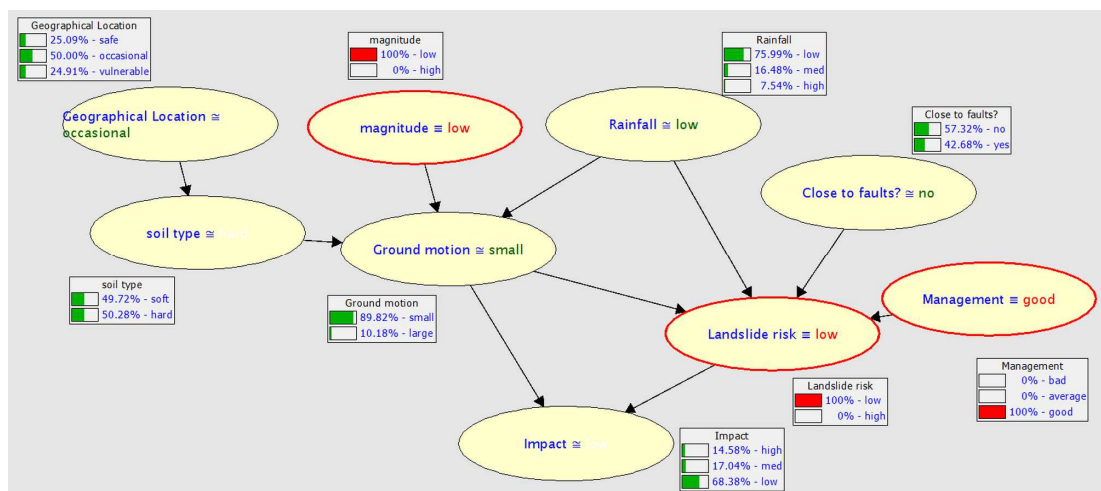


Figure 7.4: Case D: A BN with evidence about low impact risks

Four different cases A, B, C and D are considered for validation. Case A represents the situation where the risk model is run without integrating the Bayesian

framework. So the only losses associated with it are the maintenance and breakdown costs. For Case B, the Bayesian network for slope collapse risk is integrated with the systems based risk modelling approach. However, as shown in Figure 7.2, it is assumed that all the nodes of the BN are uncertain. In other words there is no evidential knowledge of the definitive state of any node. For both cases C and D, there is evidential knowledge about some of the nodes in the BN. The evidence for case C, in Figure 7.3, indicates more probability of high impact risks whereas the evidence for case D, in Figure 7.4, indicates more probability of low impact risks. Table 7.1 enumerates the different probabilities of the impacts for the three cases where the BN is used.

In this study, all the priors in the BN are based on heuristic expert opinion. The influence of the prior distributions, of the different nodes, on the final output would depend on the strength of the beliefs which are encoded in the form of priors. For example, in Figure 7.2, the Rainfall node is assigned a low state with 60% probability. This implies the belief that without any available information, it is appropriate to assume that the operation is in a region with relatively low rainfall. This follows from intuitive understanding as well, since surface mining is generally not done in areas of very heavy rainfall as there are risks of pit flooding, slope collapse and interruption of work due to heavy rains. In cases where it is difficult to form a generalised opinion about a particular variable in the BN, non-informative priors may be assigned. For example, this is done for the soil type in Figure 7.2. Mining can be undertaken in both cases of soft and hard soil type. Therefore it is difficult to make statements that most mining areas have soft soil type. Hence the prior belief is that both of the states (soft/hard) are equally likely (i.e. the prior does not give any additional information and is termed as non-informative). A strong prior belief would require a large number of observations or evidences to change the belief to a contrary one in the posterior. On the other hand, if a small number of observations indicate that the prior belief was right, the posterior would reflect this reinforced belief.

Table 7.1: Probabilities of different impacts expressed as a percentage for the various cases

Cases	Impact		
	Low	Medium	High
B	29.09%	16.41%	54.50%
C	6.50%	17.50%	76%
D	68.38%	17.04%	14.58%

For all the four different cases, the risk model is simulated for 1,000 times in a Monte Carlo fashion. The value of P_{cat} used in the simulation is 0.1. Three uniform probability distributions are constructed for the loss distribution based risk model. The distributions for the low, medium and high risk events are in the range of $\pounds [2 \times 10^6 - 5 \times 10^6]$, $[4 \times 10^7 - 8 \times 10^7]$ and $[3 \times 10^8 - 5 \times 10^8]$ respectively. The ranges are intentionally chosen to be non-overlapping and have an order of magnitude difference between them so that it is easy to distinguish their effects in the final simulation results and ensure that the model is working as expected. In real situations however, the ranges might be overlapping and the difference between high impact and low impact catastrophic risks would be less than two orders of magnitude. Nevertheless, the tool is generic and would work in such situations as well since it is not dependent on the ranges of the loss distributions.

The summary statistics for the costs is presented in the form of box plots in Figure 7.5. It can be observed that, in all the four cases, the majority of the losses are due to the maintenance and breakdown costs associated with the single risk item. This is because the probability of the catastrophic risk is much lower than those associated with maintenance and breakdown. These costs associated with the catastrophic risks also manifest as outliers in the box plots of Figure 7.5. Another important fact is that, the outliers cluster themselves and vary by different orders of magnitude. These order of magnitude differences, gives the financial impact of the various outcomes of the catastrophic risks. In Figure 7.5, Case A does not have outliers, which are orders of magnitude more, owing to the fact that it is not coupled with any BN and predicts only the losses associated with maintenance and breakdown.

To look at these extreme outliers in more detail, they are represented by stem plots in Figure 7.6-Figure 7.8 corresponding to cases B to D. As can be observed from Figure 7.7, Case C has more high impact events than Case B and most of the losses are in

the range of $\pounds 10^8 - 10^9$. A comparison with case D in Figure 7.8 shows that most of the catastrophic risks are low impact events and the losses are in the range of $\pounds 10^6 - 10^7$, which is an order of magnitude less than those in Case C.

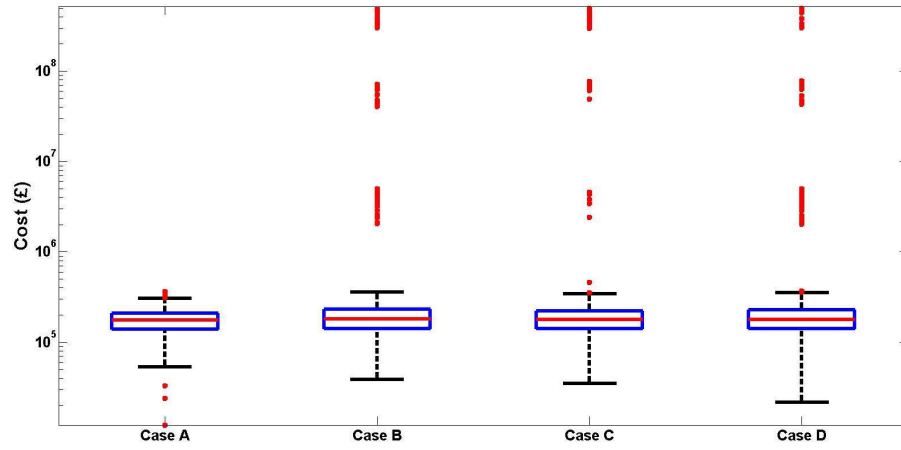


Figure 7.5: Costs due to breakdown and environmental risks for the four cases

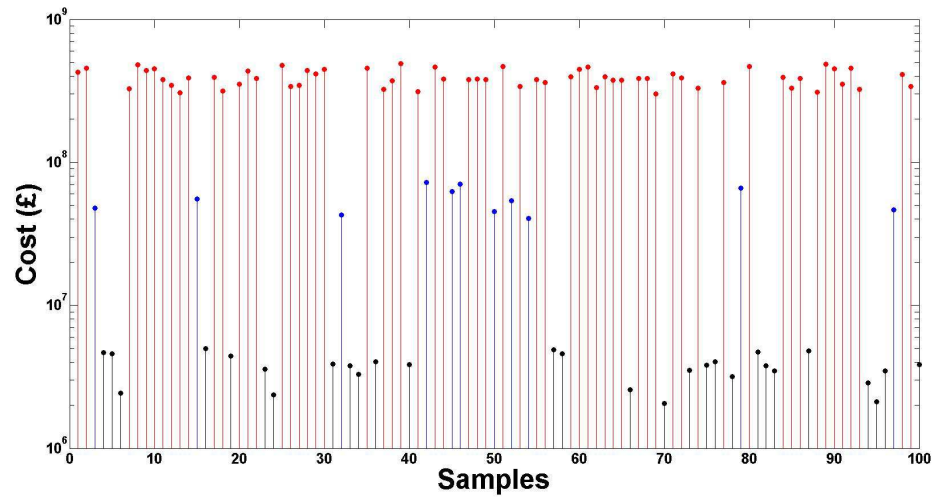


Figure 7.6: Stem plot of the costs of catastrophic risks for Case B

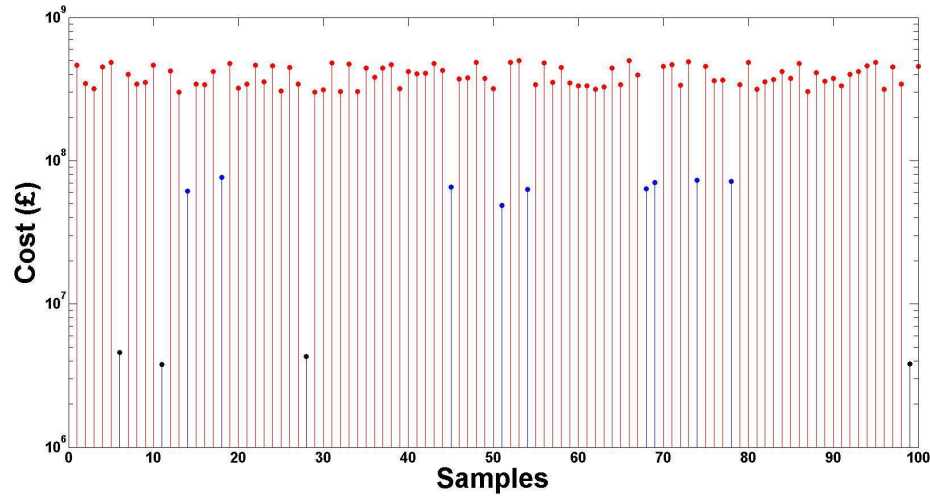


Figure 7.7: Stem plot of the costs of catastrophic risks for Case C

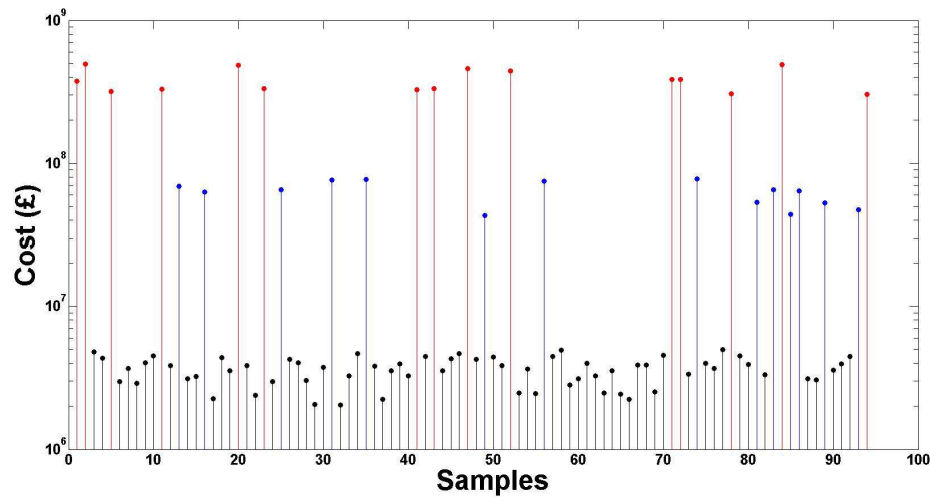


Figure 7.8: Stem plot of the costs of catastrophic risks for Case D

7.4 Discussions

The Bayesian network based risk modelling framework, as presented in this chapter, is generic in nature and is not necessarily limited to catastrophic risks only. Other types of risks including those that involve management, extrinsic events (financial market risks, political instability or worker agitation leading to prolonged shut down etc.) etc. can be easily subsumed in this framework. Even though the quality of the management has been considered as a single node in the BN, this can be extended to

include additional dependencies (like rate of employee retention, worker satisfaction etc.) if more information is available.

In the present work, the structure is pre-specified from the expert judgement of the modeller. But in cases where there is sufficient data available, both the structure and the function can be learnt from the data so that more accurate predictions can be made and less human intervention is required. One option, which is commonly pursued by contemporary researchers, is to include remote sensing and geospatial data for ascertaining the risks related to environmental events like flood, earthquake etc. For example, such studies have been undertaken using spatial analysis and BNs for insurance pricing and catastrophic risks in Li et al. (2010a). This would definitely improve the modelling results, but it is not very attractive from the insurance perspective since such detailed terrain data would not always be available for all cases. Also a model built on one terrain data is unlikely to work for another case with a completely different geological profile. Moreover, insurance decisions are generally taken in large numbers within a very short time span and therefore it is not possible to make such customised risk models on a case by case basis.

Another advantage of the Bayesian network based risk prediction is that the subsequent predictions can be fine-tuned based on the outputs of the previous predictions. Essentially this implies that the network would become more accurate at predicting the risk profile, as the probabilities are updated on a case by case basis.

7.5 Conclusions

This chapter proposed the use of Bayesian networks for catastrophic risk modelling. This approach leverages on partially available information which is commonly encountered in many insurance application scenarios. A BN was constructed for slope failure and was validated on a hypothetical risk model with different evidence scenarios. All the preceding chapters including this one, proposed different modelling methodologies and validated it with hypothetical risk items. In the next two chapters these methodologies are applied to the case of risk quantification for a mineral processing operation and a mining operation respectively.

Chapter 8 Systems risk modelling case study: Copper processing and metal production operation

8.1 *Introduction*

This chapter applies the systems based risk modelling framework to the case of a copper processing and metal production unit. The engineering process flow of the operation is briefly introduced, the individual risk sub-systems are identified and the risk profile of the system is characterised. The fuzzy inference mechanism, as proposed in the preceeding chapters, is used to characterise other process systems of similar configuration but with different levels of risk. This is useful for analysing the risk of process operations with similar layouts but for which the qualitative expert opinions differ. System optimisation is also carried out using multi-objective evolutionary algorithms. This results in a Pareto optimal set of solutions with cost versus reliability on the two axes. This can be useful especially in business planning or consulting scenarios where various risk reduction methodologies along with their capital investment can be proposed to the management. The tool can be used to quantify the reliability gained from implementing different alternative risk reduction measures.

8.2 Brief outline of the copper processing and metal production unit

Mineral processing for extraction of different minerals would involve different process operations and it is difficult to come up with a generic template which covers all of these diverse pathways under one umbrella. This chapter focusses on the specific case of copper extraction and processing and builds risk models for it. Though the methodology is applicable to other mineral processing operations as well, the underlying process systems model has to be re-constructed on a case by case basis to reflect the specific of the engineering system examined.

8.2.1 Copper ore production and other mineral processing operations

Surface mining is a exploitation method of mineral resources. In U.S., about 85% of all minerals (leaving out petroleum and natural gas) are produced using surface mining (Hartman and Mutmanský, 2002). 98% of all metallic ores, 97% of non-metallic ores and 61% of coal in the U.S. are mined using surface methods (Hartman and Mutmanský, 2002).

The case studies in this chapter focus on a copper ore production system. Most copper is mined or extracted from copper-iron-sulphide and copper sulphide minerals e.g. chalcopyrite, bornite etc. from large open pit mines in porphyry copper deposits. The average grade of copper in the ores obtained from surface mining methods is approximately 0.5% and for underground mining methods it is between 1-2% (Davenport and Biswas, 2002). The extraction of pure copper from the ore is achieved through flotation, concentration, smelting and electro-refining process. The extraction of copper from the less common oxidised minerals like carbonates, sulphates, oxides etc. is carried out using hydrometallurgical methods. However about 80% of copper comes from Cu-Fe-S ores which are difficult to dissolve in aqueous solution. Hence, pyro-metallurgical techniques are used for these cases.

Figure 8.1 shows a sample schematic of a copper processing unit. Different component flows through the process are indicated by different coloured arrows and corresponding labels. Operations which have a similar nature or function are identified by similar coloured blocks. The red dotted lines encircling one or more blocks indicate that they are grouped together and are considered as one composite unit for the risk model case study carried out.

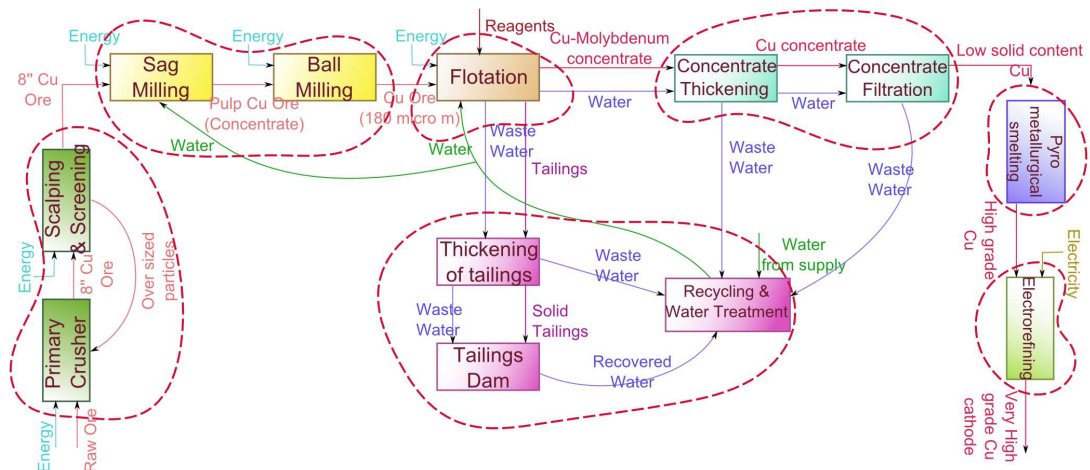


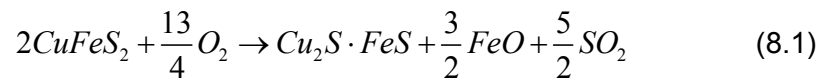
Figure 8.1: Schematic of a copper processing unit

Mineral processing is a wide discipline and numerous chemical processes are used in the conversion of an ore to the final metal product. Each of these chemical processes is specific to the particular ore or mineral and, hence, there is no generic template for all the mineral processing operations.

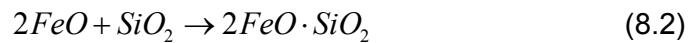
For the case of copper, the first stage in the process is the crushing of the raw ore using the primary crusher. The screening process then takes in crushed particles of various sizes and classifies them into different groups. The milling process then converts the crushed ore to finer particles by attrition and compressive forces. Different types of milling equipment like Ball mill, SAG (Semi Autogenous Grinding) mill, Autogenous mill, Pebble mill etc. might be used. In the case presented here, the Ball mill and the SAG mill are considered. The SAG mill is the generally used as the first stage of the grinding process. The SAG milling is similar to the autogenous milling where there is self-grinding of the ore. The crushed rock from the primary crushing is rotated in a cylindrical drum which causes the larger rock fragments to break on impact and grind the finer rock fragments using compressive forces. The only difference between the autogenous mill and the SAG mill is that the latter also uses balls to help in the grinding process. The Ball mills generally have a horizontal rotating cylindrical structure supplied with metal balls which move freely inside them. The particles inside the cylinder are converted to finer size by friction and impact with the balls. This finely ground copper ore is then fed into the flotation cell. The ground copper ore cannot be directly fed into the smelter as the grade of copper (0.5-2%) is too low. Hence the Cu-Fe-S and Cu-S

minerals are separated from the ore by the froth flotation process. In the froth flotation process, the sulphide ore and water mixture is supplied with reagents and surfactants, then air bubbles are passed through the solution. The Cu bearing minerals in the slurry are rendered hydrophobic and are selectively attached to the air bubbles of the vessel. The air bubbles which are injected by the air dispersal system at the bottom, rise upwards taking the copper with them. The purpose of the reagents is to render the Cu minerals water repellent while maintaining the other waste minerals 'wetted'. The minerals that float and collect at the top of the chamber are concentrated in copper (30% Cu). The minerals that do not float, known as tailings, are low in copper content and are sent to the tailings dam before extracting some waste water for recycling. Some water is also recovered from the tailings dam, and pumped back to the flotation cells and the milling processes. The flotation concentrate is further thickened and filtered to extract more water, and is then fed into the pyro-metallurgical smelting unit.

The smelting unit can be divided into several sub-processes as shown in Figure 8.2. The Matte smelting process is used which oxidises the S and Fe minerals from the Cu-Fe-S concentrate to generate matte (Cu enriched molten sulphide, $Cu_2S \cdot FeS$). Generally, air with enhanced oxygen content is used as the oxidant. The sample reactions can be summarised as



The iron oxide formed in this reaction then reacts with the Silica flux (SiO_2) to form molten slag ($FeO \cdot SiO_2$). This slag is then treated for some Cu recovery. The corresponding chemical reactions are given by



Sulphur dioxide, which is a harmful to the environment, is also generated in this process. Thus, the SO_2 is treated in the sulphur recovery system as shown in Figure 8.2 and converted to sulphuric acid. Finally, only a small percentage of SO_x is released into the atmosphere.

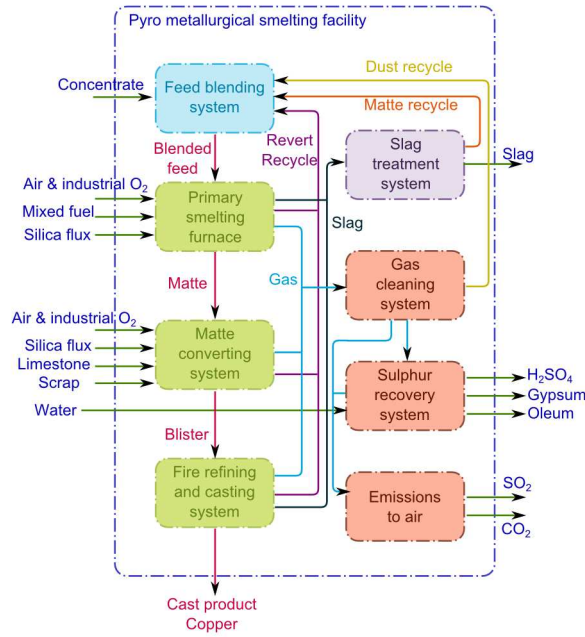


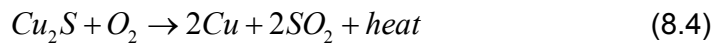
Figure 8.2: Pyro-metallurgical smelting of copper (adapted from Goonan, 2005).

The Matte then undergoes a process of converting, which is basically the air oxidation of the molten matte. This process removes Fe and S from the matte to produce crude molten copper (99% Cu). The conversion takes place in two sequential stages:

- i. Fe-S elimination or slag forming stage



- ii. Blister copper forming stage



This form of copper known as blister copper is then sent for electro-refining to obtain very high grade copper (less than 20 ppm impurities). The copper electro-refining process uses an aqueous solution of H_2SO_4 and $CuSO_4$ as the electrolyte. The blister copper with the impurities acts as the anode and pure copper is electrochemically plated into copper cathodes which can then be washed and sold as commercial commodities in the market.

8.3 Modelling of financial risks owing to operational hazards in copper processing and metal production

In this section the financial losses only owing to mechanical breakdowns of the copper processing and metal production operations are characterised.

8.3.1 Failure modes of the different components

There are different ways in which the individual components can fail, thus leading to a disruption of the overall copper production process. Assuming that there is no redundancy scheme in the process flow, i.e. no backup machinery for individual components, it can be seen that the failure of one component would hamper the whole down-stream process system. For the simulation model presented here, the various sub-systems of the copper production operations; along with the possible failure modes considered, are outlined next:

i. Crushing, scalping and screening process

Movable mechanical components are involved in this process; hence mechanical breakdown is more likely. Therefore, two different failure modes are considered:

- a. Mechanical breakdown of the primary crusher,
- b. Failure of scalping and screening equipment.

ii. SAG and ball milling

Rotating drums are involved, so mechanical breakdowns are more likely. One failure mode is considered:

- a. Mechanical breakdown of the ball/SAG mills.

Note that electrical energy is required to rotate the ball mills. It is assumed that the electricity is external to the system and is constantly available (by some captive technology, local backup or otherwise). Therefore in this context, dysfunctional mills due to lack of electricity cannot have maintenance and breakdown (MB) costs, although they would give rise to business interruption.

iii. Flotation process

The main components are the static flotation cells. Hence, leakage of one or more cells is more likely to disrupt operations than any other form of mechanical breakdown. Thus, only one failure mode is considered:

- a. Leakage of cell.

iv. Recycling, water treatment and storage in tailings dam

This composite process can be affected by multiple failure modes:

- a. Failure of the water pumping station,
- b. Temporary stoppage of tailings deposition owing to heavy rainfall,
- c. Damage to tailings dam owing to earthquakes, excessive rainfall or other natural disaster,
- d. Damage to tailings dam owing to design fault or overloading.

Note that failure of the water pumping station would also affect other upstream processes like grinding and flotation, which use the recycled water as input. This is automatically taken care of by the model structure and the corresponding processes go into the NoInput state of the finite state machine model. In the simulation example presented here only the first kind of failure is considered, as this case is based on real plant maintenance data. The latter three failures are related to environmental catastrophes and are not considered in the present example.

v. Copper concentrate thickening, filtration and drying

There are a few mechanical rotating components in thickening and compressed air drying units, hence only one failure mode is considered:

- a. Failure of the filtration/drying unit.

vi. Pyro metallurgical smelting

This unit consists of various sub units and the main failure modes can be identified as follows:

- a. Failure of smelting system and slag treatment system,
- b. Failure of sulphur recovery system.

If the smelting or the slag recovery system fails then the process would stop. However, the copper production process can go on even if the sulphur recovery system fails with the side effect that there would be significant increase in SO_x emissions into the atmosphere. However, with stricter environmental regulations,

this might mean that the management is forced to shut down the whole plant to keep the environmental emissions within acceptable levels.

vii. Electro-refining

All the different smaller failure modes like leakage of electrolyte etc. are grouped under one term:

- a. Failure of the electro-refining unit.

These failure modes represent a broad class of failures lumped together and grouped under a specific category. In cases where the company knows that a certain critical component is more likely to fail, it would enforce a proper maintenance strategy, so that the outages can be addressed in a short time or preventive action may be taken to decrease the possibility of an outage.

The failure data represents a wide class of smaller failures which disrupt operation of the particular equipment. Therefore, the maintenance team would have detailed data of all the faults that occurred, the time taken to clear the fault and the actual maintenance costs associated with it. For example, the failure mode of the mechanical breakdown of a crusher would include the breakdowns associated with the electric motors which drive the unit, the wear and tear of the mechanical equipment involved in the crushing operation, failure of the control system etc. All these failures eventually affect the output of the particular equipment and disrupt the operation. These are combined in one term representing the generic failure mode i.e. failure of the crushing operation. Table A.1 (in Appendix) shows the various failure modes with the corresponding fitted data of the TBF and OD probability curves. The TBF is fitted with an exponential probability distribution curve and the OD is fitted with a lognormal probability distribution. The data is shifted and scaled to maintain anonymity. Scaling is also implemented to discern the failures easily in the final output curves.

8.3.2 Systems framework based risk model development

A Matlab tool for operational risk modelling of the copper processing unit is developed using the modelling methodology developed in this research. The various subsystems are modelled as risk items with multiple failure modes. The distributions of the TBF, OD and MBCost are given in Table A.1. The state transition functions for the failure modes are taken as in Equations (4.1)-(4.3). The functions for $Cost_{MB}$

calculations are shown in Equations (4.7)-(4.8) and the function (Ψ) in the Health-Loading logic is as in Equation (4.9). The loading (Γ) follows Equation (4.11) and the function Φ in Equation (4.10) is:

$$\Phi_i = LF_c \times Outp_{i_{Rated}} \quad \forall i \in \{1, 2, \dots, m\} \quad (8.5)$$

where LF_c is the composite load factor calculated as

$$LF_c = \prod_{j=1}^n \left(\frac{Inp_j}{Inp_{j_{Rated}}} \right) \quad (8.6)$$

The rationale for this is that, depending on the current value of the input variable at any specific time, a load factor is calculated as a fraction of the rated input values. The output is suitably multiplied by this load factor. Hence, if there is a lower value of the input than the rated value, then the output is proportionally lowered with respect to the input. The inputs and outputs of each risk item with their corresponding rated values as used in the model are given in Table A.2.

8.3.3 System characterisation and simulation results

The system layout shown Figure 8.1 and the data presented in Table A.1 are used to simulate the risks associated with the operation for a period of one year. Figure 8.3 shows one realisation of the refined copper output. Since the original data is scaled up, the failures are magnified in the simulation and the plant remains operational for intermittent periods. Figure 8.4 shows the time evolution of the cumulative maintenance and breakdown costs of the process operation. The first subplot in Figure 8.4 shows the total cumulative cost over time and the second sub plot shows the time evolution of the cumulative costs of each failure mode. A comparison between Figure 8.3 and Figure 8.4 shows that during the periods where there is output of copper from the system, the maintenance and breakdown curves are flat. At times when there is no output, some components of the MB costs associated with some failure mode increases. The other failure modes are not active during those time instants. The curve of the total cumulative costs also increases every time there is a failure (owing to the failure modes that are active during the corresponding period of time). This agrees with the intuitive understanding of the problem. It is also shown that no MB costs are associated with periods when the system is running smoothly, whereas MB costs are incurred during breakdowns. Another important

consideration in this model is that there are no parallel processes with redundancy schemes or storage elements in the system. The consideration of parallel processes would make the system modular and its output could take discrete values in between the maximum and the minimum range. This is because even if one component of the parallel path fails, others can process a part of the input and produce a lower level of output, instead of the plant coming to a complete standstill. In the next section, redundancy schemes and storage elements are introduced and the system is optimised for performance.

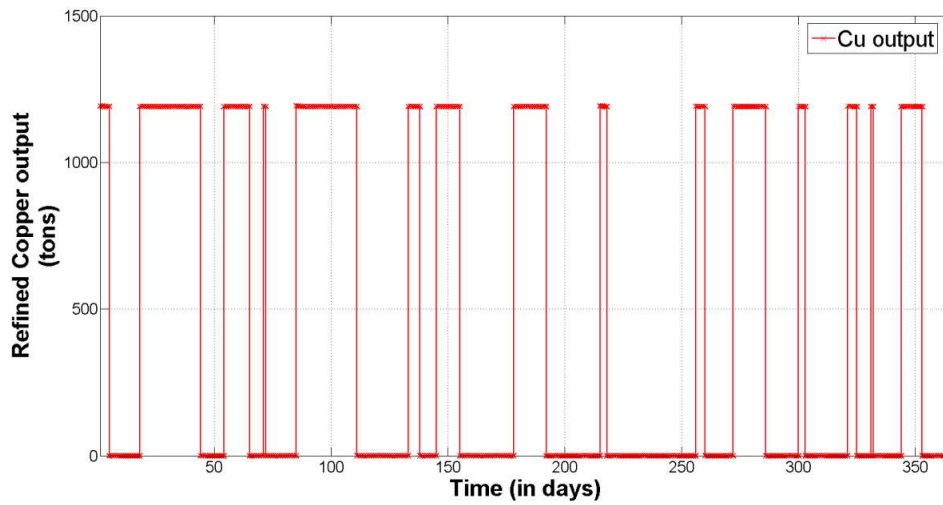


Figure 8.3: Refined copper produced over a period of 1 year. The data is scaled up to represent higher system losses.

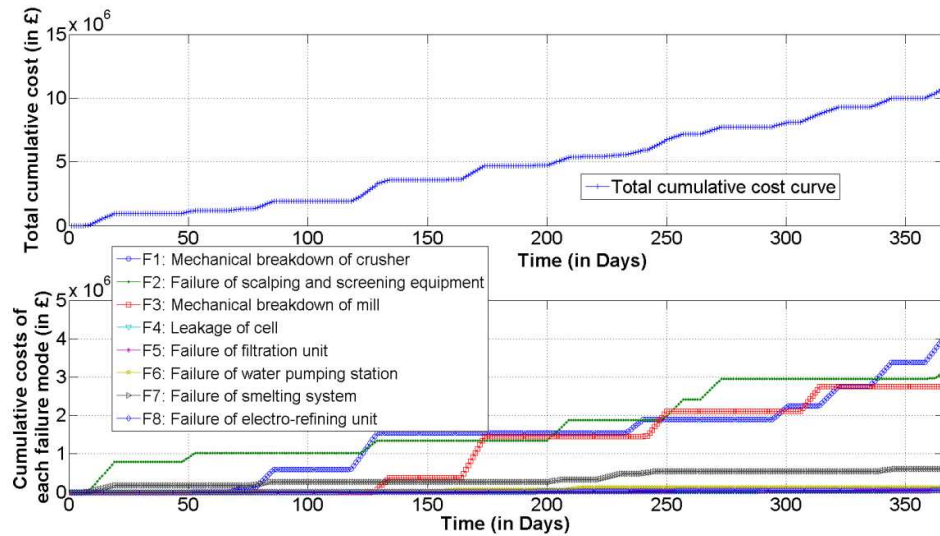


Figure 8.4: Total cumulative costs owing to maintenance and breakdown and its individual components over time

Another consideration in the model involves the event start and end points. A breakdown event can occur which starts near the end of the year and carries over to the next year. It can even be that the system is in a state of partial breakdown at the start of the year and recovers fully during the year. For the former, the model calculates the costs associated with the event till the end of the year, i.e. the period of interest. For the latter, the model assumes that the system is in a fully healthy state and none of the components have failed at the start of the year. However, these can easily be changed to include other scenarios depending on the specific problem at hand.

The system is run in a Monte Carlo (MC) fashion for a hundred times and the total cumulative cost owing to maintenance and breakdown is plotted over time. Since each of the individual failure modes fail at different times, the state trajectories of these curves are different. Also, as is evident from Figure 8.5, there is a large variance of the total cumulative cost at the end of one year.

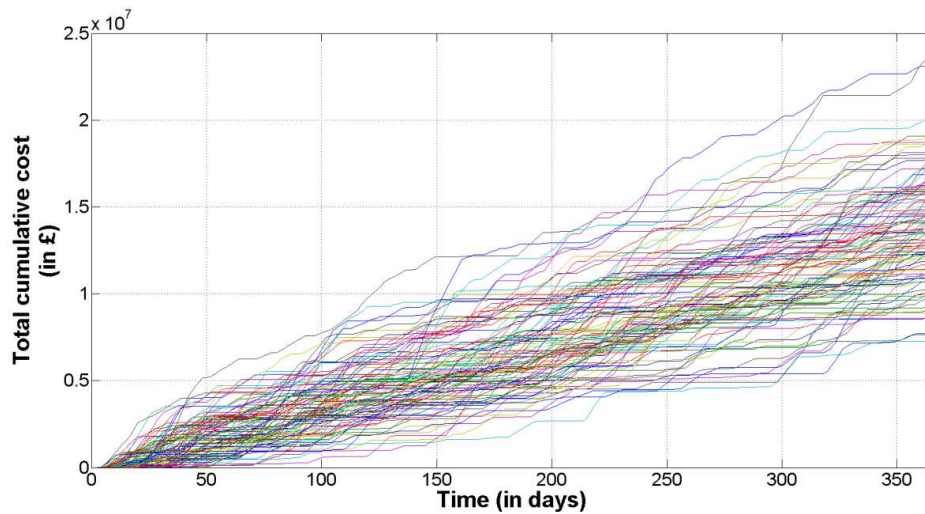


Figure 8.5: One hundred Monte Carlo runs of the evolution of total MB costs over time

To characterise the system effectively, one thousand MC simulations are carried out and histograms of the total copper output and the total MB costs are plotted in Figure 8.6. The mean (μ) and variance (σ) of the data from the two histograms are plotted in Table 8.1. The histograms show that the curves have a slightly lingering tail towards the right. This indicates that there are a few rare incidents where the MB costs incurred are very large. However, in most cases the costs lie within a certain specified band near the centre of the histogram.

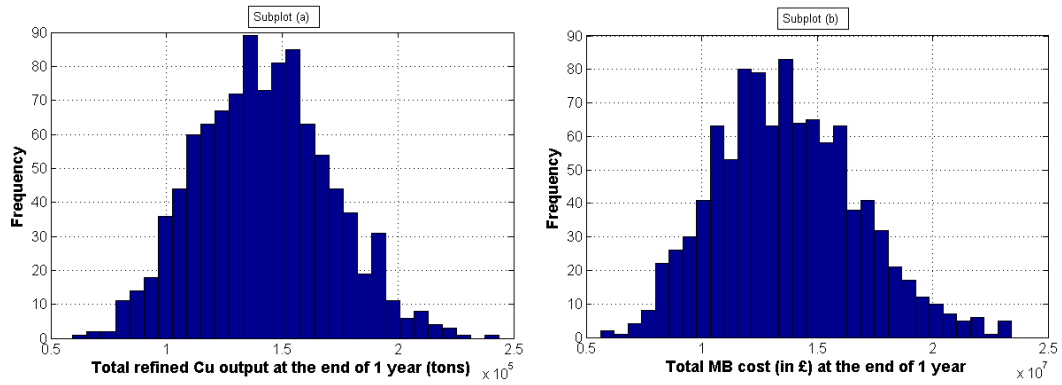


Figure 8.6: Histogram of the a) total refined Cu output and b) the total MB cost at the end of 1 year for 1,000 Monte Carlo simulations.

Table 8.1: Mean and variance of the important outputs from the model

Item	μ	σ
Copper Output (tonnes)	141,856	28,881
Total MB cost (£)	13,714,990	3,086,199

The information from these simulations can help the management to allocate a certain amount of fund each year to cover these losses. The concepts of Value at Risk and Tail Value at Risk, as discussed previously, can be used here to identify possible risk mitigation mechanisms.

8.4 System optimisation

The advantage of abstracting the process knowledge in the framework of such a tool is that it can be used not only to characterise the system, but also to optimise the system and decide the appropriate policies to employ for risk mitigation. The optimisation problem might be posed in many different ways. One way might be to look at an optimal maintenance schedule for all the equipment, so that there is least amount of downtime of the system. Another approach would be to frame the problem in a multi-objective fashion with two or more conflicting objectives. Frequent maintenance schedule might make implementation financially prohibitive or labour intensive. So, a trade-off can be made between the associated costs and the benefits achieved through such implementation. This problem can be analysed through multi-objective optimisation. Yet another way to frame the optimisation problem is to look for the effect of redundancy schemes on the system throughput. It

is obvious that keeping backup equipment would ensure less system downtime. But the key question is to identify the number and type of equipment and more importantly to quantify the effect of reduced downtime versus the initial capital costs for installation. The system can also be simulated to look at how long it takes to recover the initial capital investments owing to the improvements it offers and the resulting steady supply of production volume. A third way to frame the optimisation problem is to have some form of stockpiles at each intermediate stage, so that even if some operation fails in the process chain, the downstream production can continue unaffected for a short time until the upstream processes resume again. There is of course a trade-off between the capital investments required in such a case versus the improvement in reliability achieved by putting this equipment in place. A more complicated optimisation problem might be framed by looking at the effect of all three options coupled together.

Intelligent evolutionary or swarm based algorithms are useful for optimising these kinds of models. It is difficult to incorporate the standard methods of convex optimisation or MINLP (Mixed Integer Non Linear Programming) owing to the inclusion of finite state machines, stochastic time dependent phenomena and other constraints which might be non-convex. It is, therefore, expedient to use intelligent techniques like genetic algorithms, differential evolution, particle swarm optimisation etc. in such cases. In the following sections, the risk model of the copper processing and metal production system is optimised using Multi-Objective Genetic Algorithms (MOGA). The next section briefly outlines the salient features of the evolutionary GA and MOGA optimisation techniques. This is followed by a formulation of the objective functions and a brief description of the system and its performance objectives. The section ends with the results of the optimisation process and an interpretation of the same.

8.4.1 System description for optimisation and performance objectives

The basic process flow of the copper processing and metal production system is considered and a few modifications are made to it for process optimisation. It is noted from the simulations and also from the system data in Table A.1, that the failures associated with the crushing, scalping and screening processes are much higher (since the TBF multiplication factors are much lower). In order to mitigate the effect of the failures to some extent, a backup system with the same configuration

and failure properties is introduced in the system. This acts as a standby and allows the system to be online in the case of failure of one of the units. Along with the redundancy scheme, two stockpiles are introduced in the process flow. One of them is introduced before the SAG and ball milling process and the other one is introduced before the pyro-metallurgical smelting unit. The choices are made in line with site specific process knowledge. The rationale behind stockpiling before the SAG grinding operation is to keep the processing plant running at a specified feed rate in case there is variability in the production of the ores in the upstream mining process. This also helps to safeguard against the failures associated with the grinding process. In most cases, the output of the thickening and filtration unit in a concentrator is shipped to a distant pyro-metallurgical smelter. Therefore, the rationale behind putting another stockpile before the pyro-metallurgical unit is that it might represent material brought from other locations as well. Also, it is desirable to keep the smelting unit in constant operation throughout and a storage unit of the feed material helps in the smooth operation of the system.

In the simulation study presented, the idea is to use an optimisation algorithm and find out the values of the capacities of the stockpiles at each point in the process flow chain. The choice of the number of redundant units and the location of the stockpiles could also be taken as decision variables in the optimisation algorithm and optimised together with the stockpiling capacities. But in cases where process knowledge exists, it is wiser to design some parts of the system by taking the practical design constraints into consideration. This is because there are a lot of factors in terms of ease of use and practicality which can be easily deduced by intuitive judgement but is hard to express in terms of mathematical equations.

It is evident that a higher capacity stockpile would give the system high reliability since in case of upstream failures, the stockpile would supply the required materials to keep the operation running. However, the cost associated with the land requirement of a stockpile and the maintenance of the same increases in proportion to its capacity. Hence, a judicious trade-off is required to find out the optimal values. This is achieved by framing the problem in a multi-objective optimisation and solving it using MOGA.

Also, in the present case the upstream production from the mine is considered to be constant for a period of 1 year, which is also the period under study. Therefore, there

is no provision of replenishing the stockpiles in the case they get exhausted. This in turn implies that the initial capacity of the stockpile at the beginning of the year must be sufficient enough to make up for the losses in production during that year. In a more practical scenario, the process systems manager can ask the mine to ramp up the production to a certain level for a few weeks to reach the capacities of the stockpiles, once they are depleted following an earlier breakdown. This would result in a lower capacity requirement for the stockpile and, thus, lower allocation of capital for the same. This, however, requires heuristic inputs from the management and is also constrained by mine planning, the operating life of the mine, labour availability etc. and hence is not included in the present simulation scenario.

For system optimisation the following two cost functions are considered:

$$J_1 = \int_{t=0}^{365} (P_r - P_a) dt \quad (8.7)$$

$$J_2 = \int_{t=0}^{365} (MB_{tot}) dt + \sum_{i=1}^N H_{\text{tank}_i} \quad (8.8)$$

where, P_r is the rated production per day, P_a is the actual production per day, MB_{tot} is the total maintenance and breakdown cost associated with the whole operation per day. The first objective function J_1 in Equation (8.7), represents the loss of production owing to the various failures in the system. The second objective function J_2 in Equation (8.8) represents the total cumulative cost owing to maintenance and breakdown along with the holding and capital costs associated with the stockpiles. N represents the maximum number of stockpiles. This is proportional to the size of the tanks. The values of these are calculated as in Equations (8.9) and (8.10). It is obvious that owing to the underlying stochastic dynamics of the different failure modes, both J_1 and J_2 would evaluate to different values each time and, hence, a boot strapping method is required to calculate the expected values of these integrals.

$$H_{\text{tank}_1} = 10000 + 50 * Cap_{\text{tank}_1} \quad (8.9)$$

$$H_{\text{tank}_2} = 15000 + 80 * Cap_{\text{tank}_2} \quad (8.10)$$

where Cap_{tank_1} and Cap_{tank_2} are the capacities of tanks (stockpiles) 1 and 2 in tonnes respectively.

Since J_1 represents the lost production, it can be multiplied by a cost factor (£/ton) to convert it into cash. Then both J_1, J_2 might be assigned weighting factors and can be cast into a single objective and optimised. However, in the context of insurance, J_1 essentially gives the business interruption costs, which is treated differently than the MB costs given by J_2 . Therefore, a multi-objective optimisation framework is chosen, which can quantify the trade-off between these two.

8.4.2 Optimisation results and discussions

The system is optimised through the NSGA II algorithm which has been discussed previously. The number of individuals in each population is taken as 30 and the NSGA II algorithm is run for 50 generations. The expected value of Equations (8.7) and (8.8) is calculated by evaluating the objective functions for 10 times at each point and averaging the result. A penalty function method is adopted for solutions which fall in the infeasible regions of the search space. A tournament selection is adopted with a tournament size of 2. An intermediate crossover function is chosen which creates children by random weighted average of the parent genes. The crossover fraction is chosen as 0.8. A Gaussian mutation function is chosen for the mutation operation. The Pareto front population fraction is chosen as 0.7 of the total population. The range of optimisation variables (tank capacities) are taken as $[40000, 400000]$ and $[5000, 70000]$ for Tank₁ and Tank₂ respectively. Figure 8.7 shows the solutions in the final generation of the algorithm. The blue points represent the dominated solutions and the red ones indicate the non-dominated or the Pareto solutions. The Pareto solutions indicate the best trade-off that has been found by the NSGA II algorithm. A reduction in one of the objective functions would invariably result in an increase in the other objective function. From the Pareto front, three representative solutions are chosen and their function values along with the optimisation variables (capacities of the stockpile) are reported in Table 8.2.

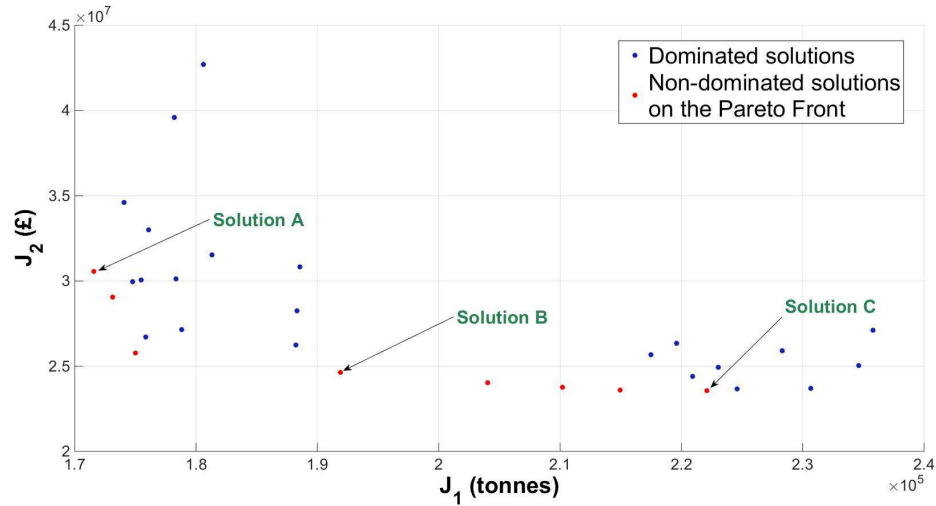


Figure 8.7: Solutions in the final generation of MOGA showing the Pareto front

Table 8.2: Representative solutions on the Pareto front

Solution	J1 (tonnes)	J2 (£)	Stockpile 2 capacity (tonnes)	Stockpile 1 capacity (tonnes)
A	171,574	30,548,128	54,532	123,774
B	191,906	24,637,750	31,785	54,981
C	222,106	23,579,387	10,000	40,002

The system is simulated with these values of stockpile capacities and the cumulative losses in Cu production and the cumulative MB and other costs are plotted in Figure 8.8 and Figure 8.9 respectively. The values of J_1 and J_2 are the final values of these curves at the end of one year. It is to be noted that since the system is stochastic in nature, the same value of J_1 and J_2 is not obtained in this independent simulation run. However, the general trend and the nature of the solutions are preserved. For example, in Figure 8.8, the total cumulative losses for copper production is the smallest for Solution A, while it is the largest for Solution C. Solution B lies in between these two extremes. However this comes at a cost. In Figure 8.9 this is

exemplified where it can be seen that the Solution A has the highest cumulative costs for a year while Solution C has the lowest. Solution B again lies in between these two extremes. Therefore, this clearly illustrates the trade-off between reliability and escalating costs of the system. Another noticeable effect is that the cost curves start from different initial values in Figure 8.9 for the different representative solutions. This is explained by the fact that the tank capacities for Solution A are higher than that of solutions B and C as can be seen from Table 8.2. The initial values in Figure 8.9 represent the total capital costs and running costs, which increase in proportion with the size of the storage unit. Another observation from Figure 8.8 is that from the start of the simulation time, the solution A does not incur any cumulative losses in production for a large period of time. This trend is followed by B and C for shorter periods of time. This is owing to the fact that the capacities of the stockpiles in Solution A are higher than those of Solutions B and C. Therefore, any shortfall in production is supplied by the stockpiles as long as there is material available in it. In this simulation no provision is made for replenishing the stockpiles during the period of operation of the whole year. Hence, the stockpiles become ineffective after they run out of material to supply.

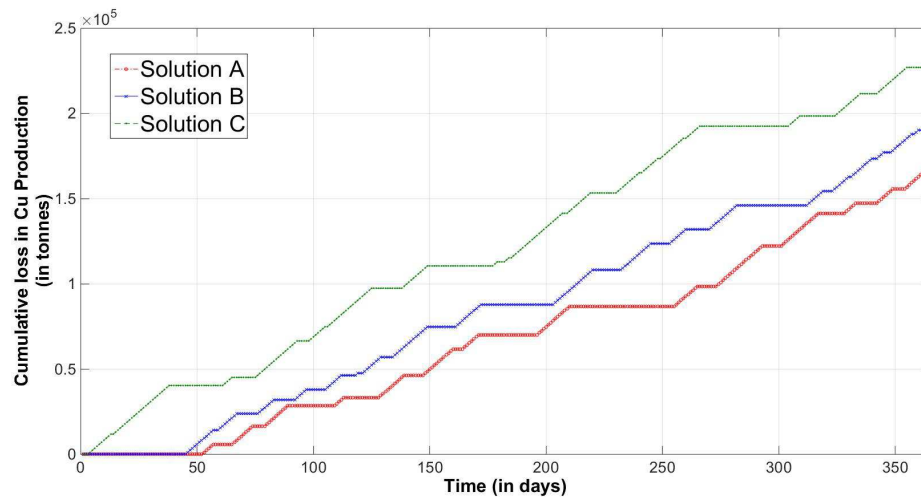


Figure 8.8: Cumulative losses in copper metal production for the three representative solutions on the Pareto front

The two figures, Figure 8.10 and Figure 8.11 illustrate this point more effectively. Figure 8.10 shows the production of copper metal over time for the three cases. Figure 8.11 shows how the level of materials in one of the stockpiles fall with time, as it is used to keep up the rated supply when there is a failure.

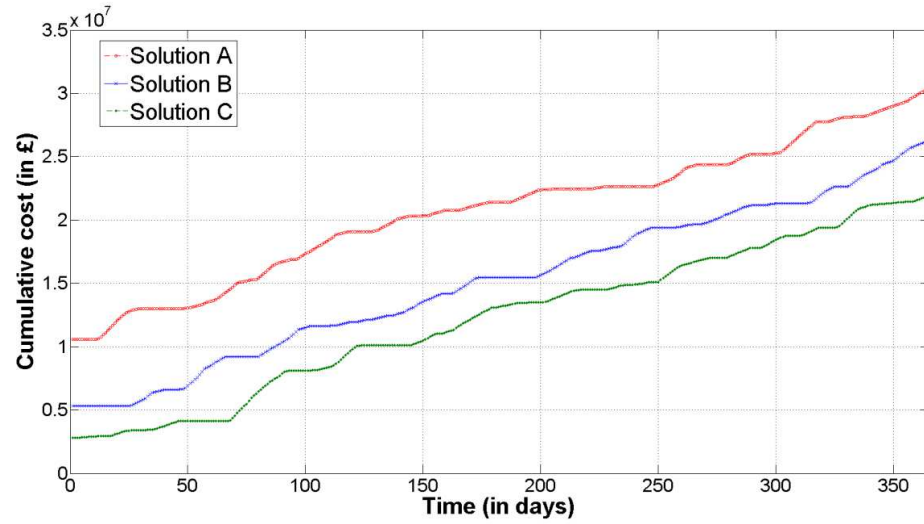


Figure 8.9: Cumulative costs for the three representative solutions on the Pareto front

In Figure 8.11, the initial levels are different, with Solution A having the highest initial value owing to largest capacity, while Solution C having the lowest.

In Figure 8.10 another noticeable effect is that there is a steady output for a little more than 50 days for Solution A and a little less than 50 days for Solution B. This implies that the stockpiles supply the rated output during that period.

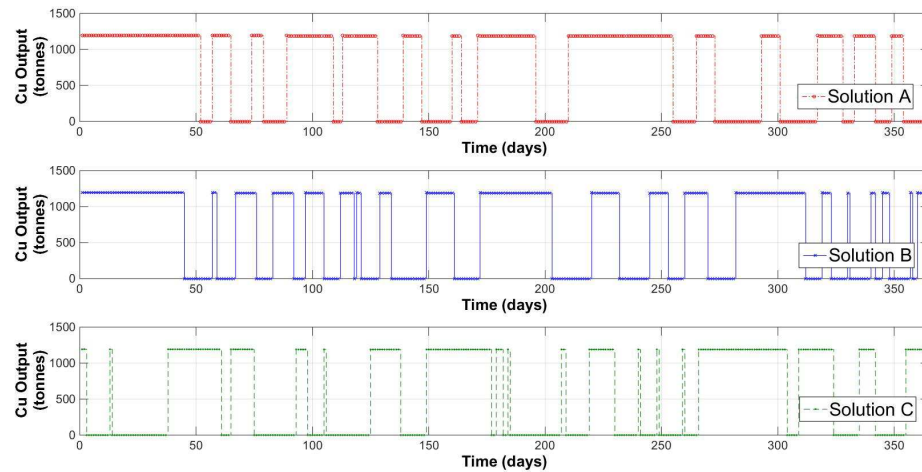


Figure 8.10: Production of copper metal over time for the three representative Pareto solutions

This is also clear from Figure 8.11 where the level of the tanks fall during the initial period. For Solution C, there seems to be an anomaly, that there are periods of failure in the initial days but the level of the stockpiles does not fall during that period to provide for the failure. This is not an anomaly, as the stockpiles are situated in-

between the process chain and cater to the upstream failures only and not the downstream ones. This implies that the failures have occurred in the downstream components of the stockpile, i.e. in the pyro-metallurgical smelter or the electro-refining equipment.

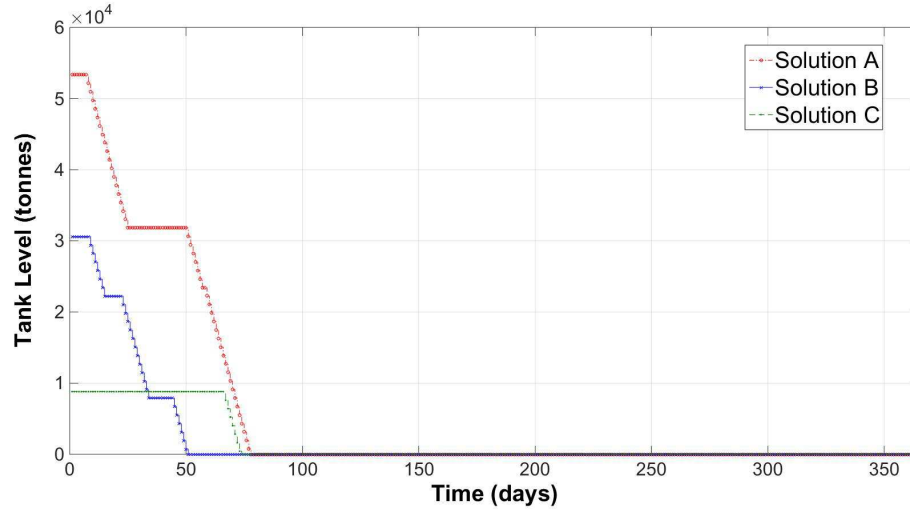


Figure 8.11: Amount of materials over time in the stockpile before the smelter

It can be intuitively understood that if a stockpile is kept at the end of the process chain, after the electro refining unit, then it would have the maximum effect on maintaining a rated product output. However, the material at this stage is already marketable and it does not make much financial sense to keep reserves of already finished goods in anticipation of a loss in production. Therefore, this is not considered in the present study. Any in-between storage units would not be able to exert any control over downstream failures, as also exemplified in this case.

8.5 Fuzzy logic based modelling

As discussed in the preceding chapters, this systems based approach is useful for a base case situation, where it is possible to get reliability data for each individual risk item. However, this does not solve all the issues from an insurer's point of view, since reliability data would not be available for all similar process operations. The fuzzy logic based mechanism is useful in such circumstances where there is 'soft' information about the process operation whose risk profile needs to be characterised. The subsequent portions of this section applies this fuzzy inference system to the the case of the copper processing and metal production operation.

The copper processing and metal production system as simulated in the previous sections is considered as the base case and the corresponding expert inputs about the Company rating, location and age of the plant, as considered for the fuzzy methodology, is shown in Table 8.3. Next, a couple of more cases are considered and the corresponding expert ratings about these processes in relation to the base case is outlined in Table 8.3. The $OD_{coeff-fuzzy}$ and $TBF_{coeff-fuzzy}$ as obtained from the fuzzy inference mechanism is also reported in Table 8.3.

Table 8.3: Cases with expert inputs about the plant and their outputs from the fuzzy system

Case	Company Rating	Location	Age of Plant	$OD_{coeff-fuzzy}$	$TBF_{coeff-fuzzy}$
Base	0.5	0.5	0.5	0.5	0.5
Worse	0.25	0.75	0.9	0.673	0.327
Better	0.8	0.1	0.2	0.142	0.859

For these three cases, the probability distributions of the base case are modified by the adjustment factors as discussed before. The corresponding systems are simulated for 1,000 Monte Carlo evaluations and the total amount of copper output and the total MB cost are reported in Figure 8.12. It can be observed from the figures that for the “Worse” case, the refined copper output is less and the MB cost is greater than the “Base” case. Also, the whiskers are much higher for the worse case, indicating that there are more extreme events and the interquartile range might not be an appropriate representation of the overall statistics. For the “Better” case, the refined copper output is higher and the MB cost is lower than the “Base” case. Also the variation outside the interquartile range is smaller. These results fit with the intuitive understanding of the different cases and also match with the qualitative information from the underwriters.

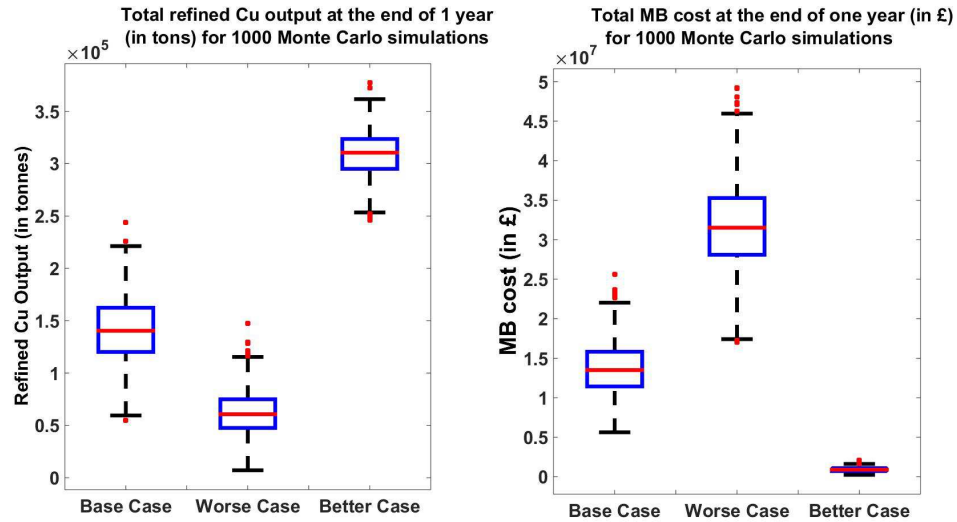


Figure 8.12: Box-and-whisker plots for 1,000 Monte Carlo runs of the three different cases with total refined Cu output and total MB cost respectively

As also done in the previous sections, the basic process flow operation is modified with a redundancy scheme and couple of stockpiles. The multi-objective optimisation is run on all the three cases.

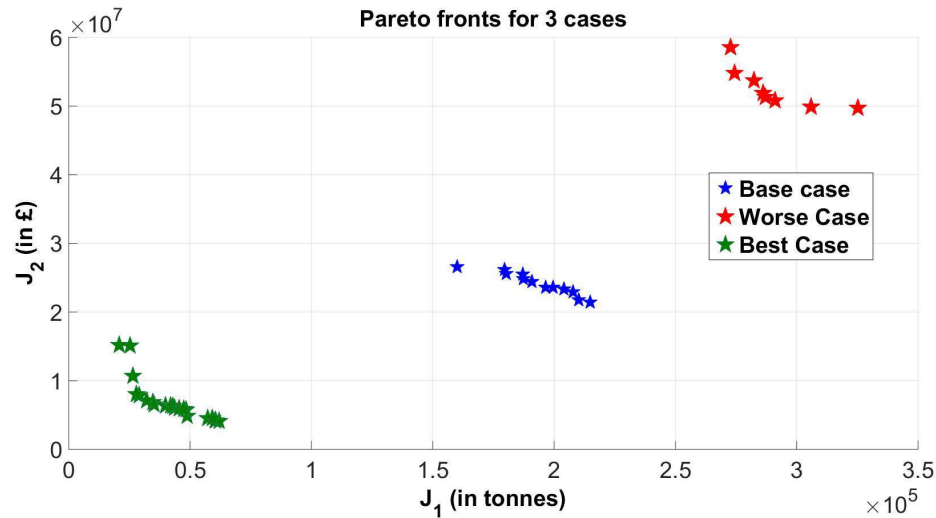


Figure 8.13: Multi-objective optimisation results for the three cases

The Pareto fronts obtained from the multi-objective optimisation are shown in Figure 8.13. The results are in accordance with intuitive understanding and the Pareto front for the worse case is dominated by the base case which in turn is dominated by the best case. The system designer needs to choose one of the trade-off points for the final design. This kind of analysis might be used by the insurer to put a rough estimate on the costs for improving the reliability of the process system. This might

be offered as a suggestion to the company and, if followed, the insurance premium might be appropriately adjusted to reflect the lower risk of the operation.

8.6 Conclusion

The risk profile of a copper ore processing and metal production operation was characterised using the systems modelling framework developed in this research. The fuzzy inference system was able to quantify the risk profiles for similar process operations based on expert information. Multi-objective optimisation was carried out to obtain a set of trade-off solutions of risk vs. reliability. The coupling of fuzzy inference and multi-objective optimisation shows the flexibility and the versatility of the systems risk modelling tool. The flexible paradigm allows for customising the tool using bespoke modules to address context specific challenges to address a practical problem of interest.

The next chapter deals with risk characterisation of an open pit mine. The process flow operation for an open pit mine can be naturally represented as a discrete event system. The systems based risk framework, as used in this chapter, needs to be appropriately modified using queuing theory constructs to model an open pit mining operation. This is presented in the next chapter. In addition, a questionnaire based approach is used for risk quantification for underground mining operations.

Chapter 9 Systems risk modelling case study: Mining operations

9.1 *Introduction*

This chapter applies the developed risk quantification methodologies to the case of mining operations. The risk profiles for an open pit mine, underground soft rock and underground hard rock mine are characterised. The Bayesian network based formalism is used in conjunction with the queuing theory based systems level risk models for quantifying the catastrophic risks in the case of open pit mining operation. For the case of underground mining operations, an expert questionnaire based approach is used to identify the different modes of failure leading to catastrophic events. An integrated model incorporating the outputs from the expert questionnaire and the systems based risk model is used to characterise the risk profiles for the underground operations.

9.2 *Application to open pit mining operation*

9.2.1 Brief overview and system model

Figure 9.1 illustrates the schematic of the process flow of an entire copper ore production operation which is used in the present work. The operation consists of two areas A and B producing a total (ore+waste) of 3.3 and 2.46 million tonnes/year respectively. The stripping ratios in areas A and B are 4:1 and 3:1 respectively. This represents a total of 1.63 million tonnes/year (.82 mt from area A and .81 mt from

area B) of copper ore being sent by the primary crusher to the mill. Alternatively it can be said that the mine produces 11,000 tonnes/day of aggregated ore and waste from area A. It also produces 8,200 tonnes/day of aggregated ore and waste from area B.

The cut-off grades in the areas A and B are 0.8% and 0.5% respectively. The ore and waste dump trucks for the two areas are all 100 tonnes in capacity with different loading times. The number of ore and waste dump trucks for area A is 3 and 12 respectively and for area B it is 3 and 9 respectively. The other relevant model parameters are given in the appendix in Table A 4.

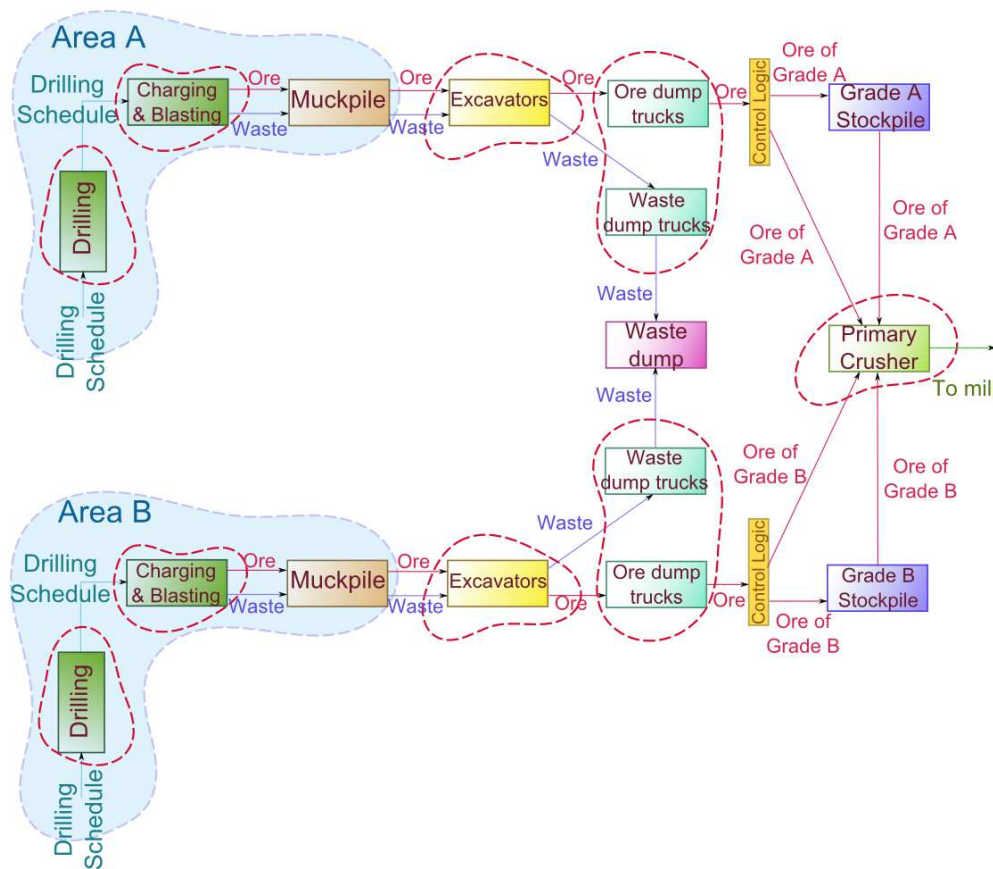


Figure 9.1: Systems representation of the copper ore production process system operation

Each subsystem as shown in Figure 9.1 has a TBF and OD probability distribution associated with it and remains in the states of failure or no-failure depending on a random number drawn from the probability distribution when entering each state. The costs are dependent on the OD and are calculated when the system transits to the failure mode. The actual model parameters of the system are given in Table A 3. For

all the simulations in this chapter, the TBF distributions are centered on the multiplication factor.

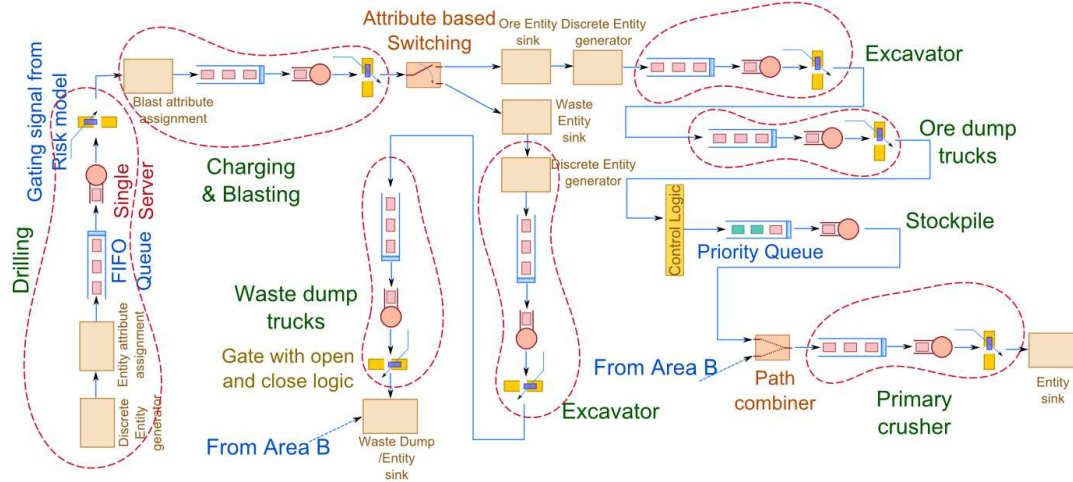


Figure 9.2: Queuing theory based representation of the process operation coupled with the systems level risk models

Figure 9.2 shows the representation of area A of the process system as a network of servers and queues. The systems level risk models control the gating mechanisms at the output of each of the servers. The stockpile containing the ore of Grade A is implemented as a priority queue with control logic. This mechanism ensures that the primary crusher is fed with a constant rate of ore A. In situations where some of the upstream processes fail and the ore dump trucks temporarily stop transferring ore to the primary crusher, the mechanism ensures that the primary crusher still receives a constant supply of ore A from the stockpile until it is exhausted. In a case where the primary crusher itself breaks down, the priority queue based control mechanism ensures that the trucks dump all the ores to the stockpile and the upstream processes are not hampered.

The charging and blasting block in the queuing model has a blast attribute assignment variable which assigns the quantity of the blasted ore or waste with some variability. Also each of the entity is assigned a tag denoting whether it represents an ore block or a waste block. This assignment is implemented randomly over time, but the aggregate number of waste entities to those of the ore is in proportion to the stripping ratio. At the end of the charging and blasting block there is a switch which routes the entities to either the ore stream or the waste stream depending on their assigned attribute. The cumulative sum of the quantities of ore or waste is calculated

at the entity sink, which represents the ore or the waste muckpile respectively. At this point new entities are generated in the excavator block, which represents the quantised amounts of ore or waste that would be loaded and carried by the trucks. These are taken off from the muckpile in discrete quantities and therefore the level of ore in the muckpile varies over time. The priority queue based control logic mechanism as discussed earlier, is used to supply the ore to the primary crusher. The path combiner at the entry of the primary crusher block takes the entities from both the areas in a specific ratio and combines them into one entity, to maintain a uniform grade of ore going into the crusher. These composite entities from the output of the primary crusher represent the final product leaving the system boundary to the mineral processing plant.

9.2.2 Simulation results

The system is simulated with the process parameters as in Table A 4 and the failure mode parameters as given in Table A 3. Figure 9.3 shows the arrival of the entities over time at different points in the process flow for the simulation time of one year.

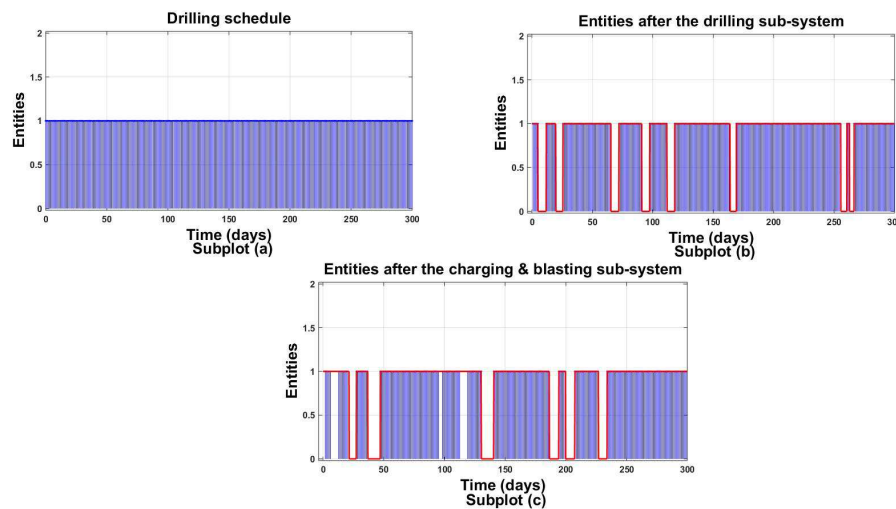


Figure 9.3: Entity arrivals at different points in the process flow of area A

The drilling block generates a drilling schedule operation everyday as shown in Figure 9.3 (a). Figure 9.3 (b) shows the entities coming out of the drilling subsystem. This is superimposed with the health of the drilling sub-system in red colour. It can be seen that only when the system is healthy, the entities come out of the sub-system. During the times in which the health is zero, there are no entities at the output. Figure 9.3 (c) shows the entities coming out of the charging and blasting block, with

the health of the block superimposed in red colour. It can be seen that in some cases no entities come out even though the health of the charging and blasting block is one. This is owing to the failure modes in the previous drilling block.

Figure 9.4 (a) and (b) show the arrivals of the ore and the waste entities respectively, at the muckpile. It can be observed that the number of waste entities is much higher than the ore entities. This is owing to the stripping ratio of 1:4 of ore to waste. These entities fill up the level of ore or waste in the muckpile. Since the excavators are also simultaneously loading to the trucks, the amount of ore or waste in the muckpile changes dynamically. This can be seen in Figure 9.4 (c) and (d) for the case of the ore and the waste respectively.

Figure 9.5 (a) shows the temporal evolution of the cumulative costs owing to the different failure modes in the system. Figure 9.5 (b) shows the cumulative costs owing to all the failure modes taken together. Figure 9.5 (c) shows the arrival of the different entities at the output of the primary crusher. Each of these discrete entities brings a certain quantity of ore at the output and the cumulative sum of the ore produced over the period of one year is shown in Figure 9.5 (d).

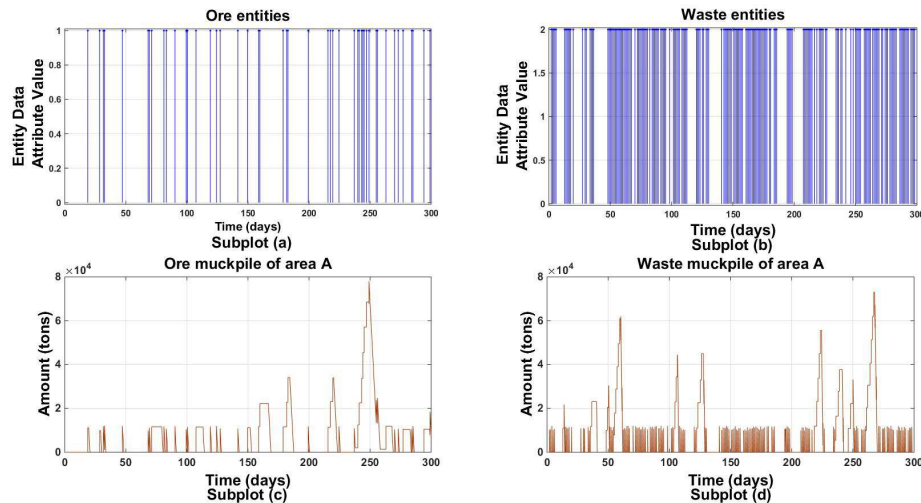


Figure 9.4: Entity arrivals and state of muckpile of area A

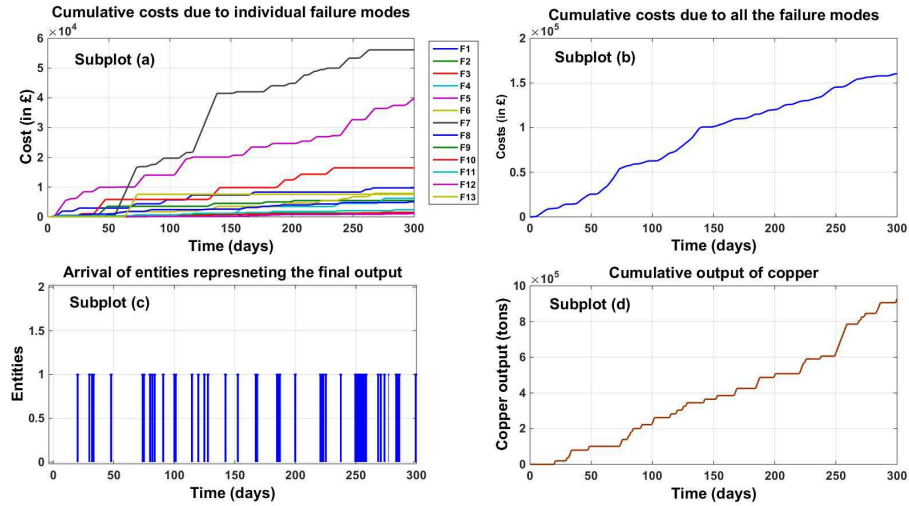


Figure 9.5: Temporal evolution of the final output of copper and total costs

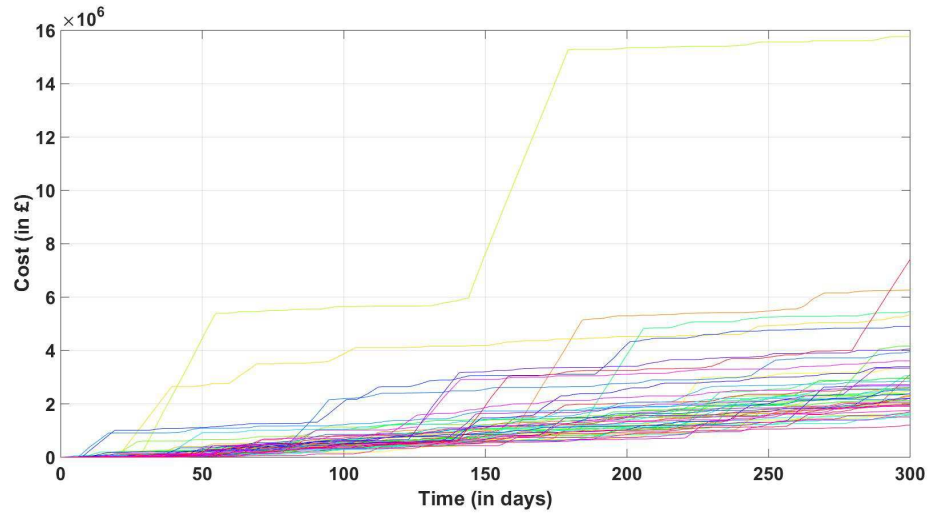


Figure 9.6: 50 Monte Carlo simulations showing the evolution of the costs over time owing to the various failure modes considered

Figure 9.6 and Figure 9.7 show the trajectories of the cost and copper ore production for 50 Monte Carlo simulation runs. Figure 9.6 illustrates that for one of the cases the time evolution of the cost is much higher. This is owing to the long tail lognormal distributions employed in the simulation model, which allows for small number of events with high variability.

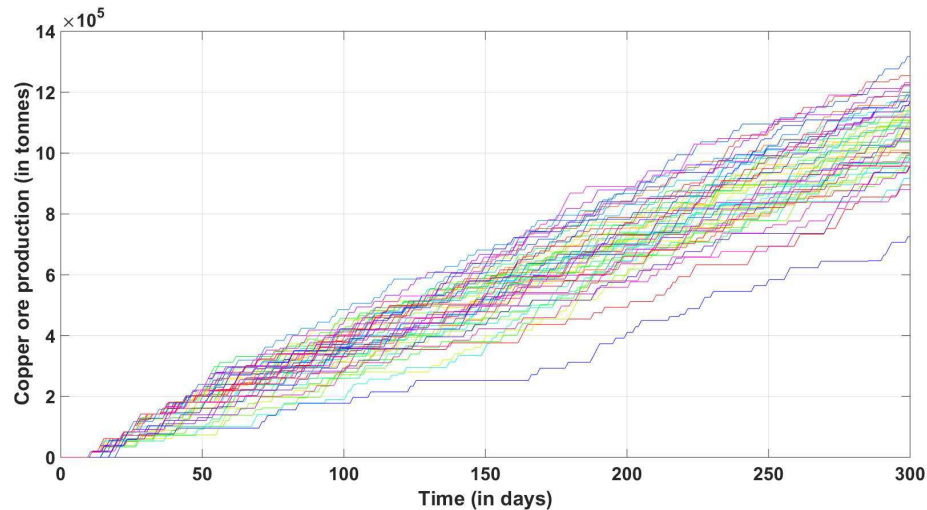


Figure 9.7: 50 Monte Carlo simulations showing the evolution of the copper production over time

Next, the catastrophic risks owing to slope collapse and flood are integrated in the risk model. Based on different partially available evidence scenarios, different cases are constructed and the corresponding risk profiles are characterised.

For the case of slope collapse, two different cases are constructed, with partially available information indicating a higher level of impact for case P and a lower level of impact for case Q. The corresponding Bayesian network based inference for both the cases are depicted in Figure 9.8 and Figure 9.9 respectively. Table 9.1 summarises the probabilities of different impacts for the two cases, as obtained from the Bayesian networks.

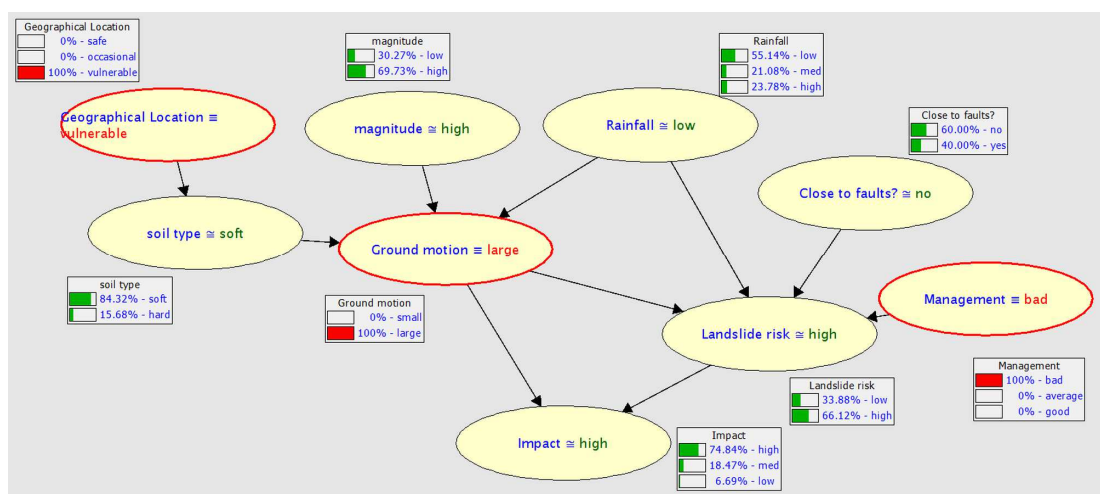


Figure 9.8: Case P: Bayesian network with evidence about high impact slope collapse risks

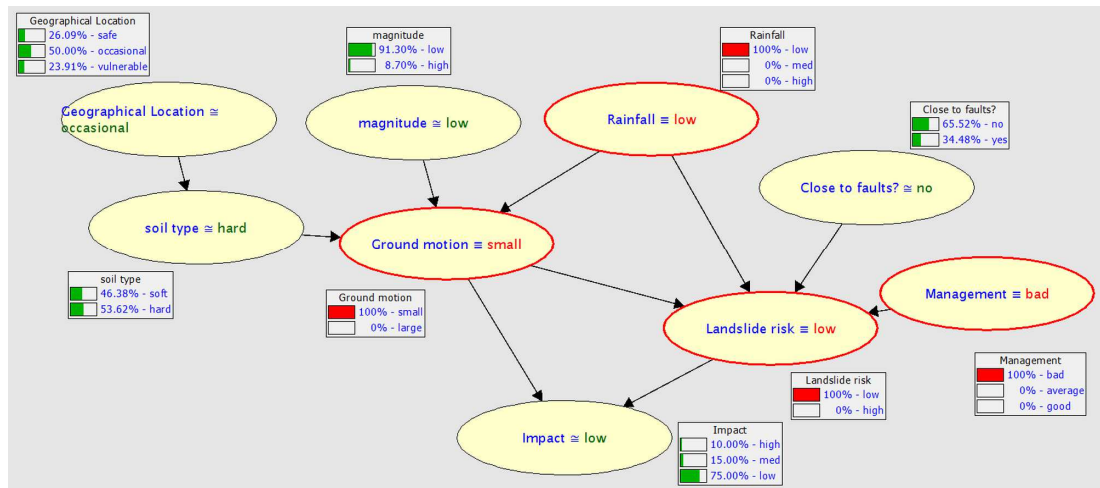


Figure 9.9: Case Q: Bayesian network with evidence about low impact slope collapse risks

Table 9.1: Probabilities of different impacts of slope collapse as obtained from the Bayesian network

Cases	Impact of Slope Collapse		
	Low	Medium	High
Case P	6.69%	18.47%	74.84%
Case Q	75%	15%	10%

Two different cases are constructed for the flood risk as well. The evidence in case R indicates a high risk of flood and a consequent higher impact. The evidence in case S implies a lower level of flood risk. Figure 9.10 and Figure 9.11 show the Bayesian networks for cases R and S respectively. The outputs of the Bayesian inference system is summarised in Table 9.2.

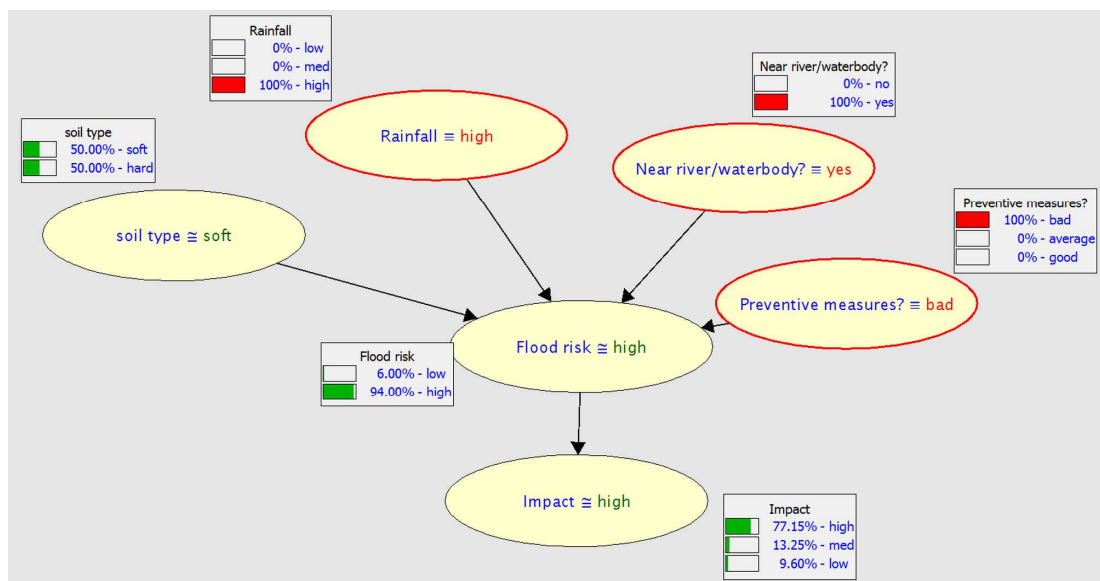


Figure 9.10: Case R: Bayesian network with evidence of high impact flood risks

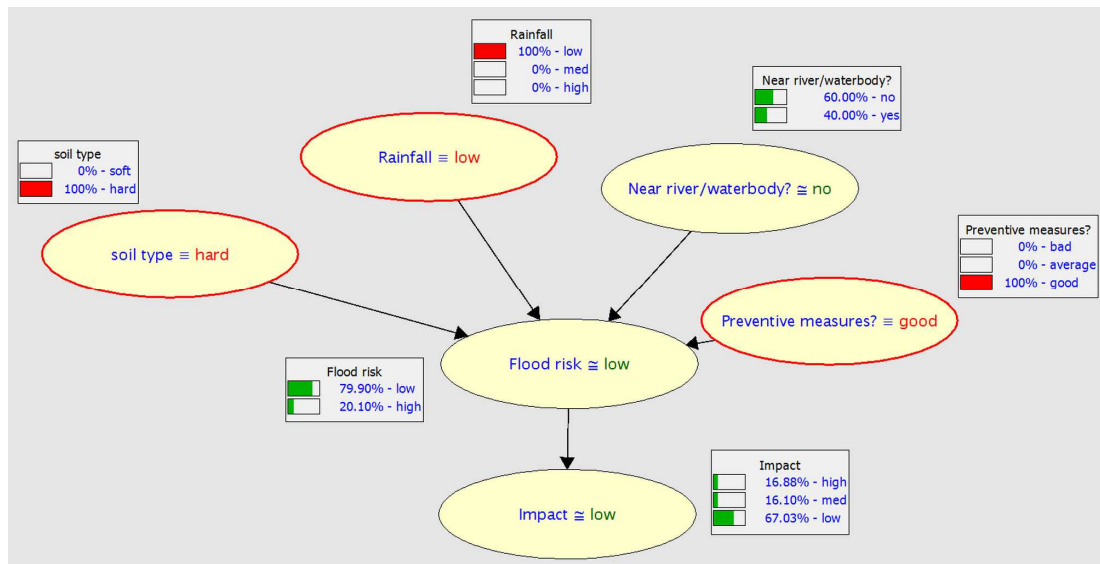


Figure 9.11: Case S: Bayesian network with evidence of low impact flood risks

Table 9.2: Probabilities of different impacts of flood risk as obtained from the Bayesian network

Cases	Impact of Flood risk		
	Low	Medium	High
Case R	9.60%	13.25%	77.15%
Case S	67.03%	16.10%	16.88%

Figure 9.12 shows the box plots of the ore outputs for all the 10,000 Monte Carlo runs. The ‘No Bayes’ case is the output from the systems based risk model which involves only the maintenance and breakdown losses without any catastrophic risks like flood or slope collapse. As can be observed from Figure 9.12, all the four cases (P, Q, R and S) have significant outliers (with low values of ore output), owing to the incorporation of the catastrophic risks. Figure 9.13 shows the box plots of the losses incurred in all five cases. It can be seen that for all the catastrophic risk cases, there are significant outliers (high values of losses). Even the ‘No Bayes’ case has quite a few large losses. This can be attributed to breakdowns and maintenance of costly equipment which does not occur very frequently.

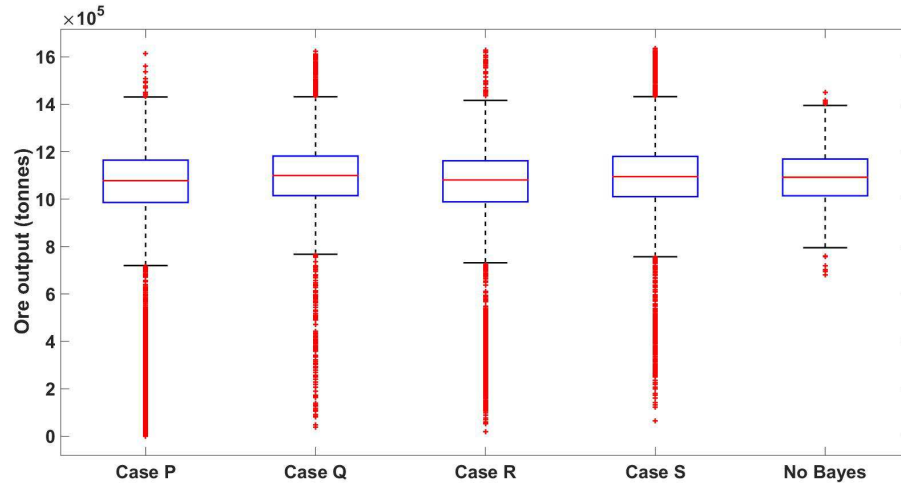


Figure 9.12: Box plots of annual ore output for all the different cases in 10,000 Monte Carlo evaluations

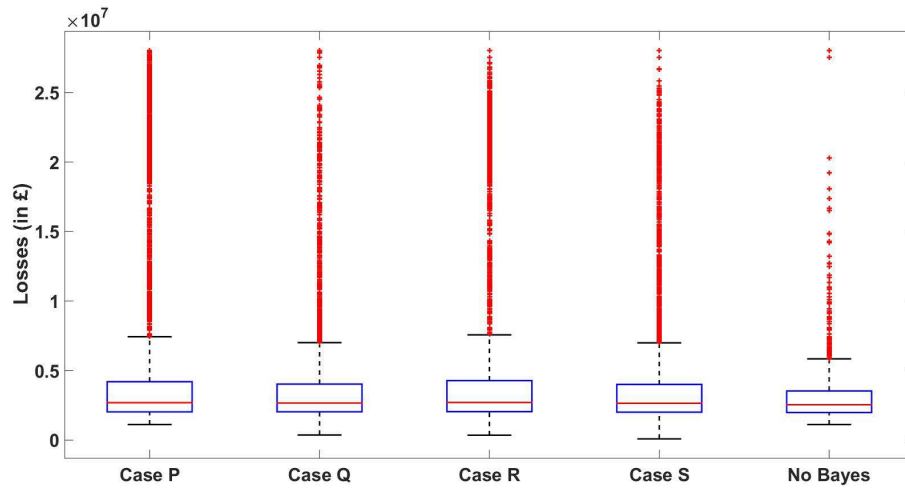


Figure 9.13: Box plots of losses for all the different cases in 10,000 Monte Carlo evaluations

In Figure 9.12 and Figure 9.13 most of the events in the interquartile range are the effect of the systems based risk model which characterises the maintenance and breakdown events. The outliers, on the other hand are mostly from the catastrophic risk events. In order to compare the effect of the different cases which incorporates the Bayesian networks, the histograms of the tail events are plotted in Figure 9.14 and Figure 9.15.

A comparison of Cases P and Q (for the slope collapse risk) in Figure 9.14, shows that Case P is more severe and has less annual ore output than Case Q. As can be observed in Figure 9.15, the corresponding financial losses are also higher in Case P than in Case Q. This result is in agreement with the inference from the Bayesian

network in Table 9.1 which shows that the severity of the risks in Case P is much higher than in Case Q.

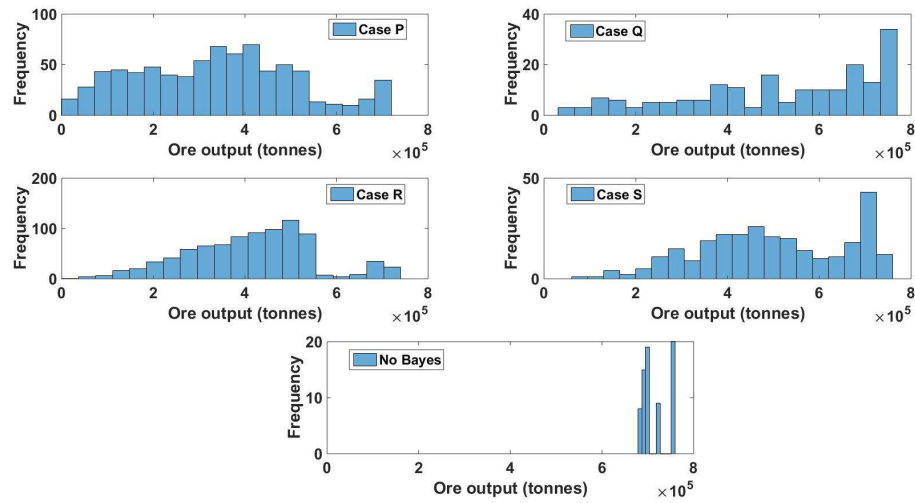


Figure 9.14: Histogram showing the ore output of the extreme events (outliers) for each case

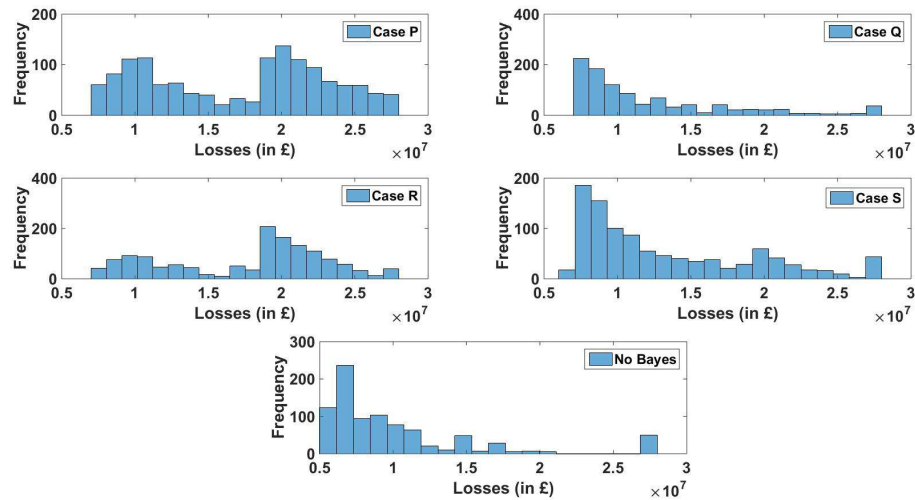


Figure 9.15: Histogram showing the financial losses of the extreme events (outliers) in each case

A comparison of cases R and S in Figure 9.14 shows that Case R is more severe and has a much higher number of events resulting in low annual output of the ore. This is also reflected in the histogram showing the financial losses in Figure 9.15 where Case S has more number of low magnitude losses than Case R. The results are in agreement with the level of severity as obtained from the Bayesian network for flood risks in Table 9.2.

Finally the ‘No Bayes’ case, without incorporating the catastrophic risks are also shown in Figure 9.14 and Figure 9.15. There are fewer of outliers for this case,

showing that the maintenance and breakdown risks do not affect the operation drastically. Also, most of the financial losses are smaller in magnitude as compared to the other four cases.

9.3 *Process models of different underground mining operations*

The overground open pit mining operations are relatively simpler and have similar configurations which can be modelled in a common template as shown in Figure 9.1. However, the underground mining process is more complicated and depends on the type of mining method, rock characteristics, local geology and management practices. Therefore, the operations were broadly categorised into underground soft rock and hard rock mines. They were further sub-classified and common templates of process operation were developed based on the different mining methods. Figure 9.16 shows the process operation for underground soft rock extraction (coal) using the longwall mining method. In Figure 9.16 there are three different process sub-systems (indicated by blue dotted rectangles) which might be sequential or concurrent with each other.

The first block represents the main development phase where the access to the coal deposit in the sub-surface needs to be built using shafts, drifts or roadways. Generally this process would precede the other two and would be in action only when the mine is starting its underground operations. The next process operation is the production development phase where gate roads are built in or around the coal resource to be extracted. The gate roads can be constructed in the coal itself, in which case the coal produced would be sent to the washery. In other cases this results in the production of waste rock which is sent to the waste dump. The final stage is the production phase at the face of the longwall operation. Generally, when there is face production at one part of the mine, there is also a concurrent production development at another end of the mine. So, both these two operations run simultaneously.

Each of these three processes has two different mechanisms of taking the coal or waste out of the underground mine. One is through a bunker using a shaft/skip. The other is through a drift conveyor. Depending on the geology of the mine one of these routes would be feasible and it can be suitably activated in the model.

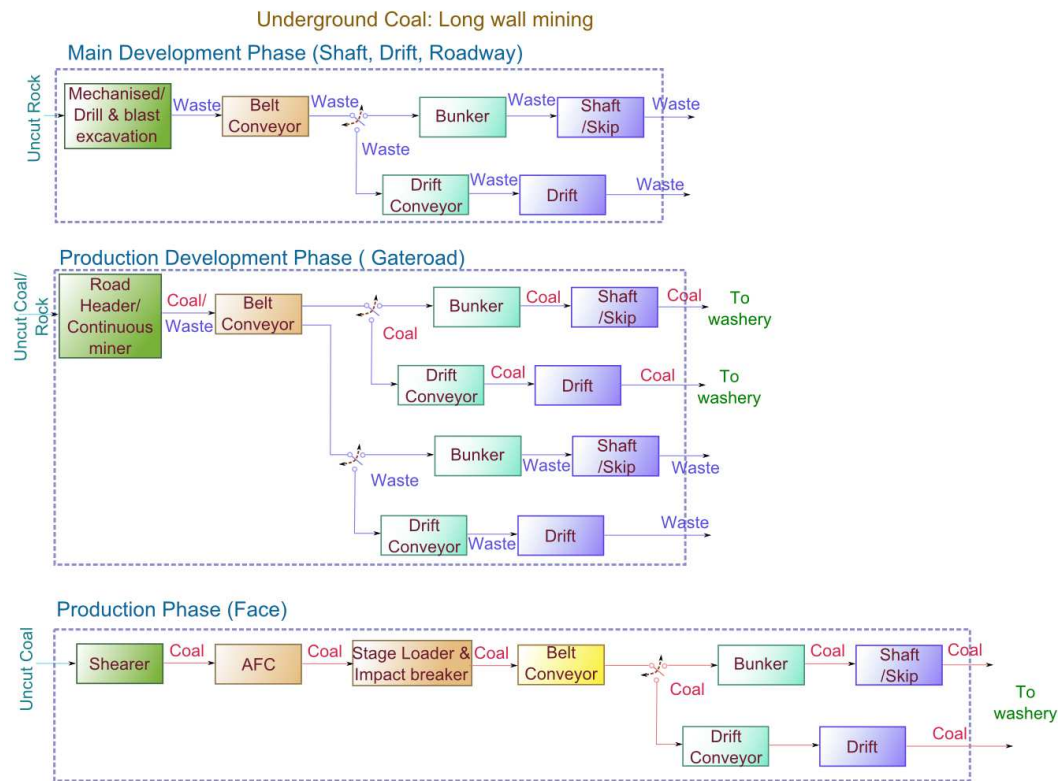


Figure 9.16: Process model for an underground longwall coal mining operation

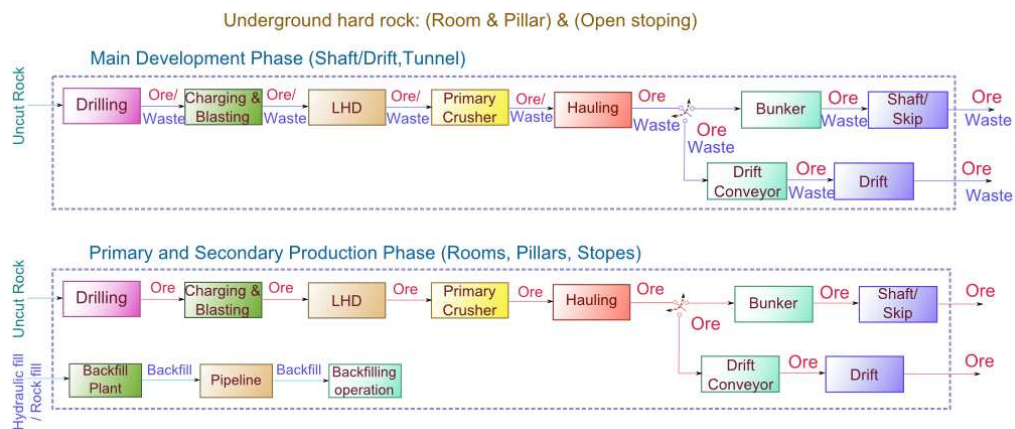


Figure 9.17: Process model for underground hard rock room & pillar and open stoping operation

Figure 9.17 shows the process model for an underground hard rock room & pillar and open stoping operation. Owing to the nature of the mining method, there is an additional backfilling operation in the process chain. Only the longwall coal mining and the room and pillar hard rock mining operations are considered as case studies in this chapter and are presented here. Several other mining operations based on

different mining methods are additionally presented in the Appendix in Figures A.1 to Figure A.4.

9.4 Risk model application for an underground coal mining operation

9.4.1 Brief overview of the risk model

Most of the underground mining operations are heavily dependent on human workforce and therefore the effectiveness of the mine management plays a crucial role in the risks associated with such operations. Traditionally the risk assessment for such mines has involved a heuristic risk rating approach which is carried out by experts who visit the site. In the present study this risk rating approach (for the mine management) is used to modify the probability distributions in the systems based risk model. For catastrophic risks like spontaneous combustion, rock outbursts, mine floods etc. a separate risk index is developed which modifies the loss distribution models for catastrophic risks. Table A 6 shows a sample questionnaire developed, which considers relative weights and user inputs used to calculate normalised risk rating indices for different catastrophic risks (RI_C) and management risks (RI_M). The questions marked with a red font are used in the calculation of the management risk index. The risk rating indices are on a normalised scale of $[0,1]$ where a value of one indicates very high risk and a value of zero indicates very low risk. The RI_M value as obtained from the questionnaire modifies the OD and the TBF distributions in the process model by multiplying the distributions with two values β_1, β_2 respectively which are given by:

$$\beta_1 = k_1 \times \left(\frac{RI_M}{RI_{M_{base}}} \right) \quad (9.1)$$

$$\beta_2 = k_2 \times \left(\frac{1 - RI_M}{1 - RI_{M_{base}}} \right) \quad (9.2)$$

where k_1, k_2 are the constants of proportionality which can be used for calibrating the model and $RI_{M_{base}}$ is the risk index for the base case. The rationale behind such a choice is that this simple set of update equations in (9.1) and (9.2) captures the gross variation of the OD and TBF with respect to the risk index RI_M . For example, if the

management risk index is a high value (close to one), then intuitively the outage duration should increase and more frequent failures would occur resulting in a lower value of TBF. Therefore in Eqn. (9.1), the factor RI_M / RI_{Mbase} would be greater than one and consequently β_1 would be greater than one (assuming k_1 is greater than one). Since the OD distribution is multiplied by β_1 , it would result in the distribution having a larger spread than the base case. The converse would happen if the management risk index is a low value. This is accounted for in Equations (9.1) and (9.2) by the two adjustment factors β_1, β_2 .

For the case of catastrophic risks like methane explosions, rock outbursts etc, a loss probability distributions based approach is adopted. The catastrophic risk indices for each of the failure modes (as obtained from the questionnaire), decides the usage of the corresponding probability distribution curve and the relative number of samplings. The schematic of the whole modelling methodology is shown in Figure 9.18.

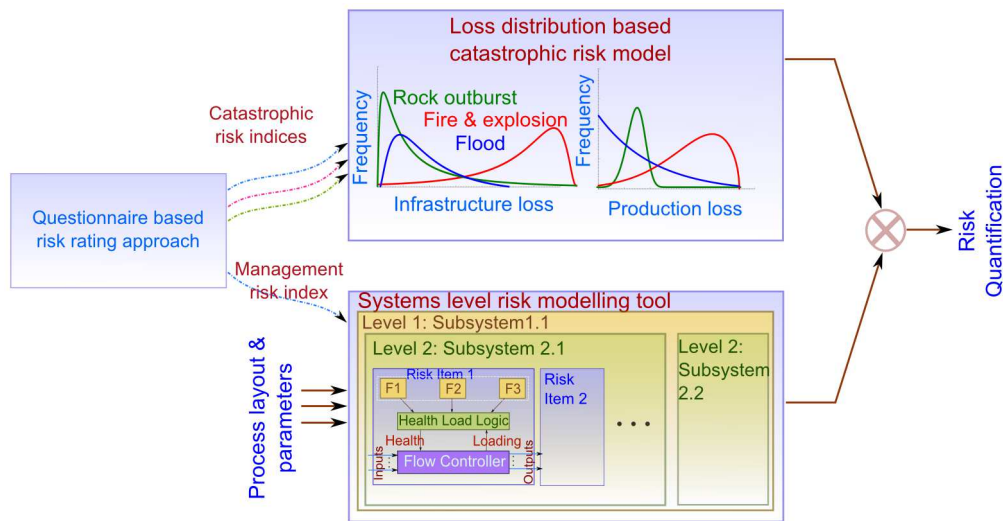


Figure 9.18: Risk quantification modelling for underground mining operations leveraging on questionnaire based risk rating approach

The catastrophic risks are categorised separately into different risk indices which represent the different modes by which a catastrophic failure may occur. These are not combined together as such an averaging would under estimate the risks in certain cases. For example, consider that there are two different catastrophic risks (say flood and fire risk). The feedback from the questionnaire indicates that the risk due to fire is much higher than that due to flood. An averaging of the individual risk indices

would overlook the high risk owing to the fire hazard and would result in a moderate risk profile. Treating each individual failure mode separately eliminates this issue.

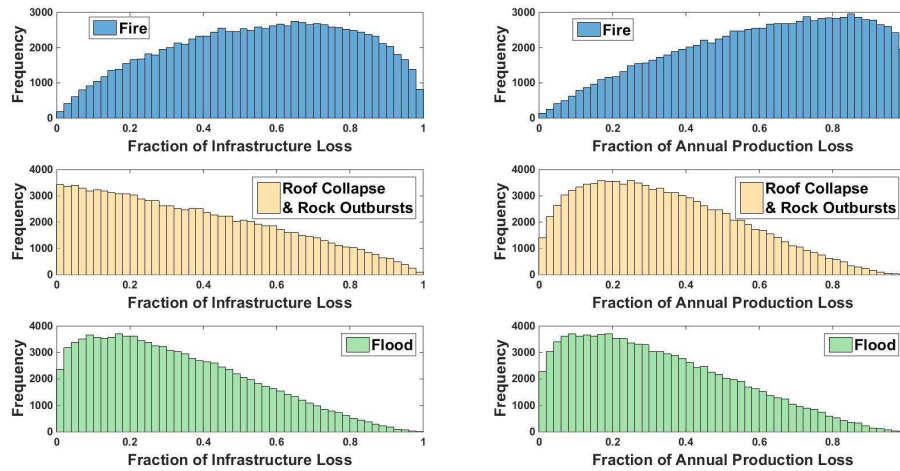


Figure 9.19: Priors for the loss distribution based catastrophic risk model

Figure 9.19 shows the priors for the different catastrophic risks considered in the model. Corresponding to each failure mode (fire, flood etc.) there are two different distributions – one which relates the fraction of mine infrastructure loss to the frequency of occurrence and the other which relates the fraction of annual production loss to the frequency. The abscissas are normalised in a range of $[0,1]$ and need to be multiplied by the capital cost of the mine infrastructure and the annual production rates respectively to obtain the actual values for a specific mine. These are obtained from expert opinion and are then fitted with probability distributions. The histograms in Figure 9.19 are constructed by drawing 100,000 samples from the fitted probability distributions.

As can be observed from Figure 9.19, the losses owing to the various failure modes are significantly different. A fire in an underground mine is much more hazardous than say a rock outburst or a flood. In other words, a fire results in a significant mine infrastructure loss and annual production loss in most occasions. This is because fire (as in the case of dust explosions) can travel to other parts of the mine which are quite far from its point of origin. Also, since it happens very quickly, there is less time for the management to react and put preventive controls if these have not already been in place. This mode of failure could also generally result in loss of lives of the underground workers and, therefore, the production might be held up for

several months in lieu of safety issues in the mine. Therefore, even though the mine might recover from the event in a few weeks, it might not be permitted by the authority to run its operations until there is an overhaul of the management team and proper safety regulations are in place. The flood risk on the other hand might affect a part of the mine and since the management has time to react by putting remediation measures like pumping out water or sealing off the damaged portion of the mine, it might be possible to prevent a large loss to the mine infrastructure. Also if the water can be effectively contained in a part of the mine, the production operations can resume in other parts of the mine, which would reduce the shortfall in the annual production targets. The case of rock outbursts is also quite localised, in the sense that it would generally affect a small part of the mine and it would be possible for the mine to recover from such an event without very significant losses to the mine infrastructure and the production targets. In case of roof-falls, there can be extreme cases where there is a cascading effect and the failure of one part of the mine results in increased stresses in other pillar supports which yield under such pressures. This cascading failure can take out the whole mine in some cases.

The algorithm to simulate the integrated risk model in Figure 9.18 is as follows. For a specific mine, the questionnaire as outlined in Table A 6 gives six separate risk indices for catastrophic risks and one risk index for management risk. The risk indices for the different catastrophic risks include –

- Spontaneous combustion risk index
- Methane explosion risk index
- Dust explosion risk index
- Rock outburst risk index
- Roof fall risk index and
- Flood risk index.

All these seven risk indices are in the range of [0,1] with a lower value indicating low risk and a higher value indicating high risk. The catastrophic risk indices are mapped into a number (ψ_p) which determines the relative ratio of simulating the loss distribution based risk models versus the systems level risk model. Each time the model is simulated, a random number is drawn from a uniform distribution and if

it is greater than ψ_p the catastrophic risk models are simulated and if it is less than ψ_p the process model is simulated. This mechanism ensures that for a given number of runs the loss distribution based risk models would be simulated more number of times if the catastrophic risk indices are higher and vice-versa. Therefore, in the final aggregated output, there would be more elements obtained from the loss distribution based risk models quantifying the higher overall level of risk.

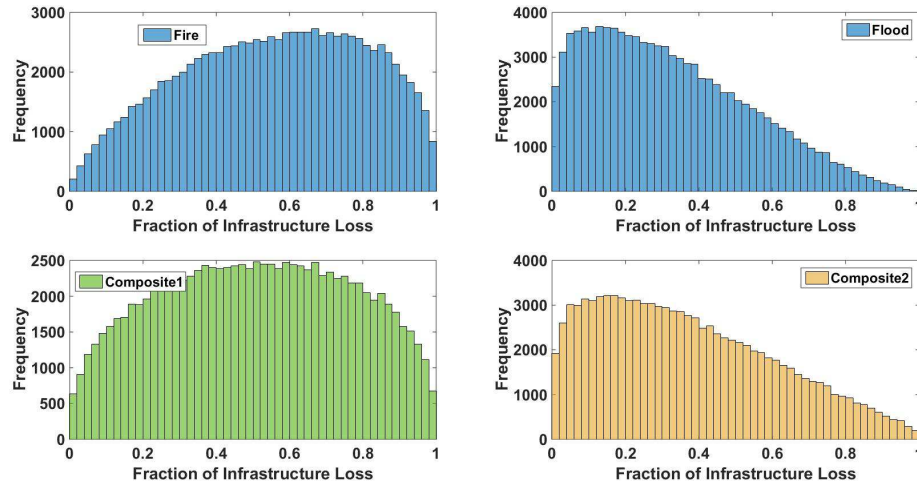


Figure 9.20: Histograms of the individual and the composite risk modes as can be obtained from the risk model

Also, once the loss distribution based risk model is activated, the relative differences between the catastrophic risk indices are used to determine the distribution to use for drawing a loss sample. Since the losses associated with different failure modes are different, this mechanism ensures that the final catastrophic loss distribution is a composite function of the various individual underlying losses. An illustrative example of this is shown in Figure 9.20. The first two subplots in Figure 9.20 show the histograms of the fraction of infrastructure loss for the failure modes owing to fire and flood respectively. The next two histograms (Composite1 and Composite2), show the output owing to the sampling strategy which draws samples from the fire and flood distributions in the ratio of 0.8:0.2 and 0.2:0.8 respectively. As is expected, the histogram labelled Composite1, has a higher level of infrastructure loss owing to the predominance of the fire events, but it also has smaller levels of infrastructure losses owing to the presence of flood events. Similarly in the case of the histogram labelled as Composite2, it is more skewed towards low infrastructure losses owing to

the presence of a larger component of the flood risk. However, it has a fatter tail than the distribution of the flood risk owing to the presence of the high impact fire risks.

9.4.2 Simulation results

The longwall coal mining operation is chosen as a sample case for demonstrating the risk modelling methodology. The adopted process system model for this operation is the one shown in Figure 9.16. The coal mine has a shaft access with a production rate of 8,400 tonnes per day. The total number of days of operation in a year is considered as 300 days. The capital cost and quantity of individual equipments are enumerated in Table A 8. The probability distributions of the TBF and the OD for the different failure modes are shown in Table A 7. Two different cases are considered as shown in Table 9.3. The values of the management and catastrophic risk indices are obtained from the questionnaire in Table A 6. The corresponding values of β_1, β_2 are calculated by considering $RI_{M_{base}} = 0.5$ and $k_1 = k_2 = 1$. As can be observed from Table 9.3, Case A has a higher risk rating than Case B, as obtained from the expert questionnaire.

Table 9.3: Different cases of catastrophic and management risk indices simulated for the underground coal mining risk case

Risk Indices from questionnaire	Case A	Case B
Catastrophic Risk Indices		
Spontaneous combustion Risk Index	0.443	0.060
Methane explosion Risk Index	0.853	0.282
Dust explosion Risk Index	0.667	0.271
Rock outburst Risk Index	0.578	0.250
Roof fall Risk Index	0.952	0.190
Flood Risk Index	0.375	0.125
Management Risk Index	0.949	0.246

The integrated risk model as shown in Figure 9.18 is simulated for 10,000 Monte Carlo runs for each of the cases A and B. The corresponding annual production of coal for both the cases is represented in the form of histograms in Figure 9.21. It can be observed that Case A has a much higher number of events where the annual production is significantly low. Case B on the other hand has a consistent output about a mean value. This matches with our intuitive understanding of the risk ratings

as given in Table 9.3. The catastrophic events in Case A recur more often causing significant shortfall in annual production in those cases. Figure 9.22 shows the box-plots of the total losses for the Monte Carlo simulation runs for both cases A and B. It can be seen that Case A has significant high impact tail losses (outliers) as compared to Case B. This is owing to the costs incurred by the events corresponding to the catastrophic risks. Also the inter-quartile range for Case A is bigger than that of Case B. This can be attributed to the inefficient management and a consequent higher management risk index of Case A over Case B (as shown in Table 9.3).

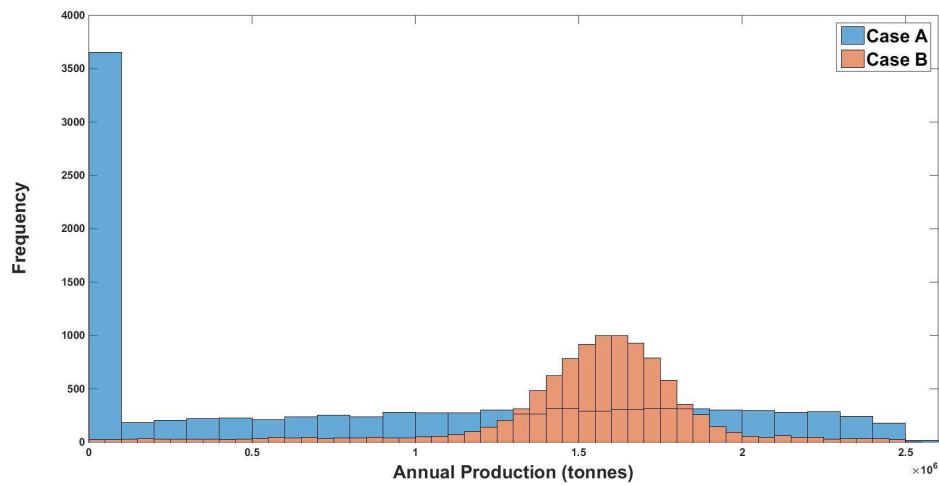


Figure 9.21: Histograms of annual production of coal from 10,000 Monte Carlo simulations for both cases A and B

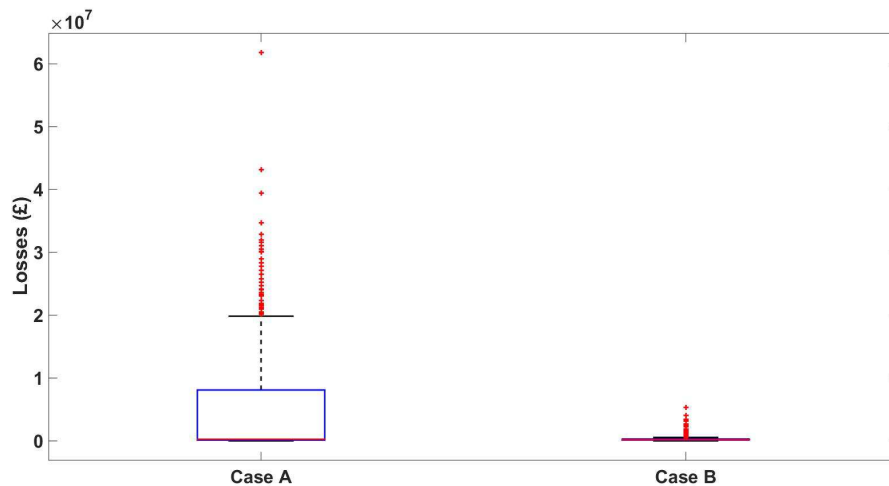


Figure 9.22: Box plots of total costs (losses) obtained from 10,000 Monte Carlo simulations for both cases A and B

9.5 Risk model application for an underground hard rock mining operation

9.5.1 Brief overview and system model

Owing to the geological properties, soft-rock mines (especially coal mines) are more hazardous and consequently many different modes of failures must be taken into account. The hard rock mines (e.g. metal mines) are comparatively less hazardous. Therefore, even though the overall risk modelling methodology for both cases would be the same, the modes of failures, associated catastrophic risks and impacts would be different. For this reason, a separate case study of an underground room and pillar Zn-Pb metal mine is considered for risk assessment in the hardrock case. The daily production is 2,000 metric tonnes per day, with 300 working days in a year. The risk model parameters for the underground metal mine are tabulated in Table A 10. The quantity and cost of the different equipments are enumerated in Table A 11. The process system model considered for simulation is the one illustrated in Figure 9.21. A similar integrated risk modelling methodology, as outlined in Figure 9.18, is applied to this underground hard rock mining operation as well. The catastrophic failure modes are, however, different than the underground soft-rock coal model. Therefore, a separate questionnaire is constructed for the underground hard rock mines as outlined in Table A 9. The answers to these questionnaires along with their corresponding weights are used to determine the normalised management and catastrophic risk indices.

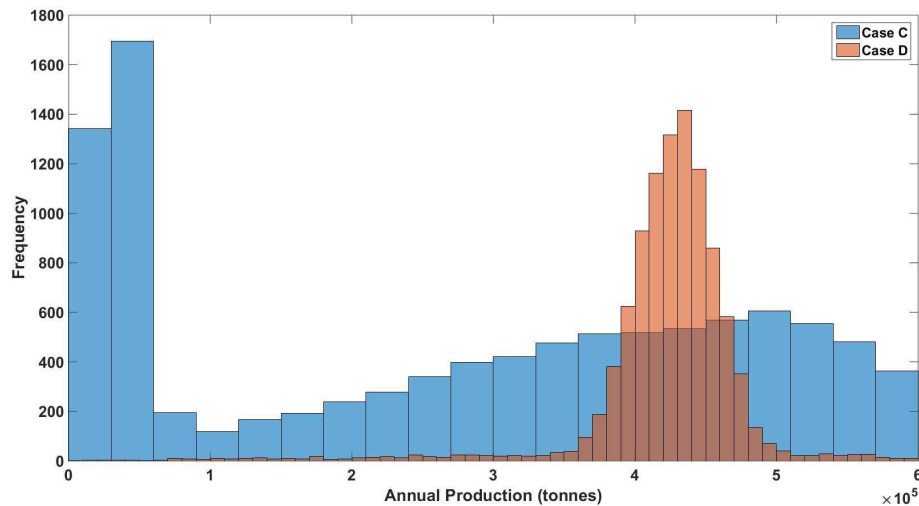
9.5.2 Simulation results

The risk profile is characterised for two different cases – a high risk scenario (Case C) and a low risk scenario (Case D). These different cases along with their corresponding risk indices as obtained from the questionnaire are tabulated in Table 9.4. The values of $RI_{M_{base}} = 0.5$ and $k_1 = k_2 = 1$ as in the previous case.

Table 9.4: Different cases of catastrophic and management risk indices simulated for the underground hard rock mining model

Risk Indices from questionnaire	Case C	Case D
Catastrophic Risk Indices		
Roof fall risk index	0.816	0.122
Stope failure risk index	0.327	0.102
Backfill failure risk index	0.909	0.045
Flood risk index	0.750	0.125
Management Risk Index	0.816	0.132

The integrated risk model is run for 10,000 times in a Monte Carlo fashion, for both the cases C and D. The corresponding histograms of the annual production of ore for both these cases are shown in Figure 9.23. Similar to the previous case of the underground soft-rock coal mine, it can be observed that the system with high catastrophic risk indices (Case C) has significant number of events where the output production tonnage is very low. This is in contrast to the low risk scenario (Case D) which has a higher mean annual production and with a smaller standard deviation. Figure 9.24 shows the costs associated with these events in the form of box plots. Case C which has higher value of risk indices has much larger number of tail events with high financial losses.

**Figure 9.23:** Histograms of annual production of ore (Zn-Pb) from 10,000 Monte Carlo simulations for both cases C and D

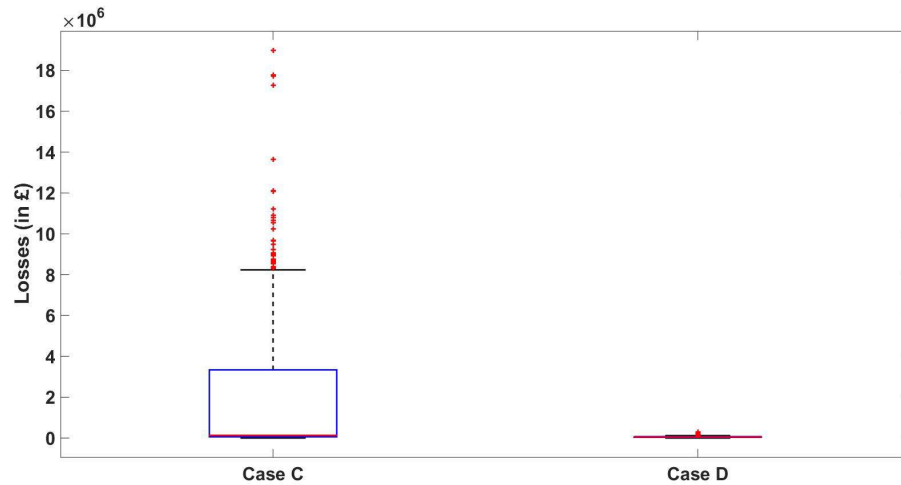


Figure 9.24: Box plots of total costs (losses) obtained from 10,000 Monte Carlo simulations for both cases C and D

9.6 Conclusions

This chapter applied the developed risk modelling tools to the case of different mining operations. For the open pit mining case, the queuing theory based models were used to represent the underlying process operation and the risk profile for the maintenance and breakdown failures were characterised. This process model was then integrated with Bayesian networks for predicting the catastrophic risks owing to flood and slope collapse. Subsequently, the risk profiles of an underground soft rock coal mine and a hard rock metal mine were characterised. An expert questionnaire based approach was used to identify the different underground risks like rock outbursts, roof-falls, fire and explosion etc. This was integrated with the systems level risk model to obtain the risk profiles including both the maintenance and breakdowns risks along with the catastrophic risks.

The open architecture of the systems risk modelling framework was useful for modelling the underlying process operation using servers and queues. This can therefore be easily extended to other modelling scenarios employing queuing theory. For example, a risk quantification can be done for cyber attacks by modelling the underlying transmission of communication packets within a given computer network using queuing theory and coupling it with the systems risk model. Other applications might include risk modelling due to disruption of transportation networks (airlines, trucks etc.), automated manufacturing process lines (e.g. soft drink manufacturing and packaging plants, silicon chip design plants for computer processors) etc.

Chapter 10 Conclusions and recommendations for further research

10.1 *Achievements and conclusions from the research*

This PhD research has developed a new tool for quantitative risk modelling of engineering process operations. The tool is generic in nature and is applicable to any engineering process system. The risk model is built on the principles of systems thinking and is of an open architecture in the sense that it is easily able to integrate diverse modelling strategies and leverage on the additional insights provided by these modules. In addition to the afore-mentioned advantages, the systems thinking paradigm which forms the core of the model, can handle missing or vague information in the process chain. This is very valuable from the application point of view, since detailed process systems data might not be available to external risk analysts or insurers. This approach helps to overcome the limitations of the mass balance based models which need the values of the process variables at each point in the process chain. It is also a trade-off between these complicated mass balance based process simulators and the oversimplified historical data based approaches which miss out on the underlying root causes for the failures.

The model has mainly been developed keeping in mind the challenges faced by the external insurers and risk analysts who often have to deal with missing data or rely on expert judgements to arrive at a risk profile. Nevertheless, the model is useful to

the management of the engineering process operation as well. It can be used by the management team to identify the key bottlenecks in their process operation and gauge the level of risks involved in the whole system. This might help in their decision to set aside a fixed amount of capital for such contingencies or implement adequate maintenance strategies or redundancy schemes as appropriate. Where there is a higher chance of losses, they might even consider transferring the risks to a third party like an insurance company. Depending on the circumstances, the process operator can also partially self-insure for the smaller losses and transfer the catastrophic risks to the insurance company.

The systems risk model is implemented using finite state machines with a switching logic between the states and is parameterised by three probability distributions which can be obtained from reliability data of the risk items. The methodology offers the flexibility of incorporating storage tanks, maintenance schedules, redundancy schemes etc. which are important in modelling any realistic scenario. These additional improvisations have been implemented in addition to the basic model functionality and validation studies have been conducted to show the effective operation of these modules.

Since many process systems can be represented as a network of servers and queues, it is imperative to integrate the systems based risk model with this paradigm, to make the modelling methodology truly generic. This has been illustrated successfully in this thesis and several validation studies have been conducted to verify that the model is working as expected.

Another issue with any kind of risk model is the uncertainty in the model parameters. Since prior historical data is very scarce and might not be available in all cases, the model parameters cannot be learned from this data-set. To circumvent this problem, expert opinion is encoded in the form of a rule base and a fuzzy inference system is developed which modifies the risk model parameters based on some expert input rating indices for each process operation. The fuzzy inference system is essentially a pre-processor to the systems based risk model and does not need to call it during run time. This helps in speeding up the simulations which involve multiple Monte Carlo runs. The fuzzy inference system is also validated with case studies to show the effective working of the methodology.

Another critical issue in any risk modelling framework is the incorporation of catastrophic environmental risks which have huge financial implications. Generally, to arrive at detailed site specific risk indices, various Geographical Information System (GIS) datasets might be used. However this is an elaborate process and needs to be done on a case by case basis. Therefore this is infeasible, especially considering the short amount of time in which underwriting decisions must be made. Also, in many cases, some of the data might not be available at the time of making a decision. To circumvent all these issues, a Bayesian network (BN) is constructed with the key parameters which drive each of the catastrophic risk modes. These are represented as the nodes of the network with the inter-dependencies among them being represented by the edges. The advantage of such a framework is that the BN can improve the predictions on the basis of partially available evidence. It does not necessarily require evidence for all the parameters to make an inference. This ensures that the risk model makes best use of the available data. The BN based methodology for predicting catastrophic risks is coupled with the systems based approach which models the maintenance and breakdown risks. Therefore, this results in a composite risk profile incorporating both the low and high severity risks.

Finally, the thesis applies the developed risk modelling techniques to the case of a mineral processing operation and various open pit and underground mining activities. This validation shows that the model works well for real life cases. For the case of mineral processing, the finite state machine based system model is used to simulate the risks for the operation. The fuzzy inference system is used to incorporate the expert opinion and deal with uncertainties in the risk model parameters. Additionally a multi-objective optimisation is carried out to propose various alternative configurations which might reduce the risks in the operation. A Pareto optimal profile is obtained which indicates that the alleviation of risk comes at the cost of additional capital expenses. This kind of optimisation is useful for suggesting various alternative redundancy schemes or maintenance schedules facilitating effective risk mitigation and control.

For the case of the open pit mining operations, the queuing theory based models are used to construct the underlying process operation. This is then coupled with the systems based risk model, as described previously, to arrive at a risk profile. The Bayesian network based catastrophic risk modelling technique is also integrated with

this model to incorporate the effects of slope collapse and flood risk. For the underground mining operations, two different scenarios are simulated – the soft rock (coal) case and the hard rock (metal) case. An expert questionnaire based approach was adopted to arrive at a management risk index and various catastrophic risk indices. These indices were used to modulate the parameters of the systems based risk model and a loss distribution based catastrophic risk model. Different case studies were constructed considering various levels of management and catastrophic risks and the corresponding risk profiles were characterised.

10.2 Limitations and unanswered questions

The case studies conducted in this thesis demonstrate the viability of the proposed methodology. However, more improvisations can be made to fine tune the model. For example, different failure modes can be cascaded together and a logic can be incorporated which triggers one failure mode if the impact owing to the other failure mode reaches a certain threshold. A sample scenario for such a case would be that there is a rock outburst, which damages some of the underground diesel equipment, which goes on to catch a fire and the fire spreads to other areas of the mine causing significant damage. This involves a non-linear dependence among the various failure modes which cannot be characterised with the present model.

In this thesis, the fuzzy logic mechanism is used based on three input (company rating, location specific risk and age of plant) and two output parameters ($OD_{\text{coeff-fuzzy}}$ and $TBF_{\text{coeff-fuzzy}}$). However, the methodology outlined is very general and can be easily extended to other key process parameters of interest and can be used to update the risk model accordingly. In cases where the process flow of the base case system is markedly different from the process system whose parameters are updated by the fuzzy inference system, the method would not yield meaningful results. However, it can be used if the base case is detailed enough and the new system is a subset of the base case, e.g. comprising fewer components. Then the fuzzy inference model should update only those parameters which are relevant for the new case and the failure modes of the other components can be deactivated. In such case, the presence of additional subsystems would not affect the risk profile, as they are considered non-failing and only transform one set of input variables to the other.

Also it might be that there are backup services for some of the critical components in the process chain. In such cases the system reliability would be higher than the base case situation without any redundancies. To include this possibility in the fuzzy framework it is necessary to define an extra input parameter in the fuzzy inference scheme which reflects the reliability of the system based on the design of backup services. The expert rule base needs to be appropriately updated to reflect how this parameter, in conjunction with the others, would affect the risk profile of the overall system.

The present design and coupling of the fuzzy inference system with the risk simulator, modifies all the probability distributions of each risk model with some constant factor as inferred by the fuzzy system. This is useful in the case where there is very limited information regarding the plant operation. However, in other cases there might be equipment specific data which can be incorporated to improve the risk model. The fuzzy methodology as outlined in this thesis is also applicable in such circumstances. Fuzzy rule bases need to be created for each separate process equipment and the corresponding probability distributions in the risk model can be updated based on the particular rule base.

Another issue is that of calibration of the model. The base model can be calibrated with sufficient accuracy as there would be enough historical information available for the same. However, this would not be the case for the models updated with the fuzzy rules, in general. It might be possible to obtain historical failure data of a few more process operations and test the validity of the fuzzy inference system with those. If enough of these historical sets of data are available, an optimisation routine can be called, which updates the fuzzy membership functions so that the error (between the actual and the model predicted values) in the final distribution of the risk profile is minimised. In cases where none of the above is feasible, the fuzzy method is the best possible use of the available data and expert knowledge. This can be iteratively improved in a case by case manner as new projects are accepted based on this model and with the knowledge of whether they actually incurred sufficient loss or not.

For the underground mining case, much of the risks are owing to the management and this feature is included in the model through a risk index which is based on a questionnaire. The current questionnaire is not exhaustive and it is possible to

improve the list of questions and the mapping of these questions to the risk index. This would help in capturing the risk profile better by effectively combining all the available information sources.

In many cases, sufficient information may not be available to construct a meaningful risk profile. The model must be able to identify such situations and output some measure of confidence which indicates that the obtained risk profile is not reliable or is based on insufficient data. This would help in making better underwriting decisions and judgements. At present this is not incorporated in the model.

The present model can distinguish the relative levels of risk between different cases. However it is difficult to relate this with the industry practices, since the current risk level assignment practices are mostly heuristic and based on different risk rating indices. This limitation would be eliminated over time as more and more cases are analysed using the risk tool forming a database of the previously analysed cases. Any new case can be compared with the earlier ones to judge the level of risk considering the historical cases. This would help in the creation of a new and more quantitative risk rating index, which could be used in its own right.

Also at the moment, all the risk model coding is implemented using Matlab, and there is no provision for parallelisation. Obtaining the probability distribution of the financial losses entails multiple calls to the risk model which can be easily done in parallel by deploying it on clusters with large number of processors. Making the code parallel would speed up the total simulation run time.

10.3 Recommendations for future research

The main focus of the PhD was the development of a novel risk modelling tool which would be of practical value to external insurers and risk analysts and apply this tool to the case of the mining and mineral processing operations. This is valuable from an industry perspective as it can be used to identify and rank different operations based on their risk exposure. The obtained risk profiles can be further used for calculating insurance premiums of such operations. Future work could look at applying the modelling tool to characterise the risk profile of other process systems like chemical process plants, supply chains, energy networks etc. The model might need to be adapted to each specific case to incorporate expert knowledge into

the modelling methodology (as it was done by using a questionnaire based approach for the underground mining case).

There is further scope for improvement of the model by trying to eliminate the limitations outlined in the previous section. This entails improvement of the questionnaire, evaluating reliability of the results, code parallelisation etc. among others.

Apart from the fuzzy and Bayesian methods, various add-on modules can be developed to incorporate the effects of different kinds of information in the final risk profile. One of the key challenges in such situations is the reliability of information as obtained from the different sources. Since many of the risk indices are calculated heuristically by experts, there might be personal biases among them. Therefore, situations might arise where one expert or surveyor thinks that there is a potential risk from the circumstantial evidence, while another expert might infer that the risk is negligible. Clearly the risk ratings as obtained from both the experts would be very different and if each of them is separately used as a model input, then the outputs would also be markedly different. Some form of data fusion techniques must be used to infer the most probable risk ratings by fusing multiple information sources. This would help in streamlining the collective inputs from multiple sources and help in arriving at a more realistic risk estimate which is relatively free from personal biases.

The catastrophic risk modelling modules can be drastically improved by incorporating geo-spatial data into the methodology. This would however take significant time and resources. A methodology can be developed where the risks associated with the geospatial information is pre-computed and stored in a database. Various automated data-mining and machine learning techniques can be used for this purpose. The risk model can take these pre-computed values as site specific inputs and incorporate this information for predicting the risk profile. This would dispense off with expensive computations at run time and also help in speeding up underwriting decisions.

There are many other possible avenues to explore and build upon the basic risk modelling framework as developed in this thesis. One of the key elements which would drive towards a better risk estimate, is the interplay between encoding heuristic expert information and intelligent processing of scarce data. Future work

can be directed keeping this general objective in mind, so that all the available information can be effectively leveraged to arrive at a more accurate risk profile.

Appendix

The cost data presented in some of the following tables have been interpreted from the Infomine database made available by Sciemus Ltd.

Table A.1: Shifted, scaled and anonymised data of a mineral processing operation, used for the simulation.

Components	Failure modes						
	Name	TBF (exponential) ($Dist_1$)		OD (lognormal) ($Dist_3$)			$MB_{\text{costfactor}_{OD}}$
		Multiplication factor (days)	μ	Multiplication factor (days)	μ	σ	(£ / day)
Crushing, Scalping & Screening process	Mechanical breakdown of crusher (F1)	43	0.75	5	0.13	0.24	8,179
	Failure of scalping and screening equipment (F2)	32	1.21	8	0.18	0.29	6,728
Sag & ball milling	Mechanical breakdown of mill (F3)	81	0.8	6	0.36	0.31	10,872
Flotation	Leakage of cell (F4)	64	1.3	4	0.11	0.16	1,190
Thickening and filtration unit	Failure of filtration unit (F5)	87	1.16	3	0.14	0.22	912
Recycling, Storage & water treatment	Failure of water pumping station (F6)	96	1.45	5	0.15	0.43	807
Pyro metallurgical smelting	Failure of smelting system (F7)	73	0.7	4	0.59	0.3	1,537
Electro refining	Failure of electro- refining unit (F8)	110	0.53	3	0.05	0.2	534

Table A.2: Inputs and outputs of each risk item with their corresponding rated values.

Components	Inputs	Rated Inputs	Outputs	Rated Outputs
Crushing, Scalping & Screening process	Energy (kWh)	17,200	8" Cu Ore (tonnes/day)	40,000
	Raw Ore (tons/day)	40,000		
Sag & ball milling	Energy (kWh)	865,200	180 µm Cu Ore (tonnes/day)	40,000
	8" Cu Ore (tonnes/day)	40,000		
	Recycled water (tonnes/day)	16,800		
Floatation	Energy (kWh)	64,400	Cu Concentrate (tons/day)	1,196
	180 µm Cu Cu Ore (tonnes/day)	40,000	Process water (tonnes/day)	113
	Recycled water (tonnes/day)	57,200	Tailings (tons/day)	38,760
			Waste water (tonnes/day)	69,600
Thickening and filtration unit	Cu Concentrate (tonnes/day)	1,196	Low Solid Content Cu (tonnes/day)	1,196
	Process water (tonnes/day)	113	Waste water (tonnes/day)	67
Recycling, Storage & water treatment	Tailings (tonnes/day)	38,760	Recycled water (tonnes/day)	74,000
	Waste water (tonnes/day)	69,667		
	Makeup water (tonnes/day)	30,120		
Pyro metallurgical smelting	Low Solid Content Cu (tonnes/day)	1,196	High grade Cu cathode (99.6%)(tonnes/day)	1,196
Electro refining	High grade Cu cathode (99.6%) (tonnes/day)	1,196	Very High grade Cu (99.99%)(tonnes/day)	1,196

Table A.3: Distributions of the failure modes of the different system level components of the open pit copper ore production operation.

Components	Name	Failure modes					
		TBF (exponential)		OD (lognormal)			MBcost
		mult fact	mu	mult fact	mu	sigma	£/day
Drilling	Drilling risk model (F1) Area A	62.00	0.13	4.00	0.49	0.33	375
	Drilling risk model (F2) Area B	62.00	0.13	4.00	0.49	0.33	375
Charging & Blasting	C&B risk model (F3) Area A	86.00	0.52	1.20	0.68	0.21	2,163
	C&B risk model (F4) Area B	86.00	0.52	1.20	0.68	0.21	2,163
Excavator	Ore excavator (F5) Area A	77.50	0.81	2.50	0.57	0.24	3,235
	Waste excavator (F6) Area A	80.00	0.76	2.00	0.58	0.21	1,892
	Ore excavator (F7) Area B	77.50	0.81	2.50	0.57	0.24	2,817
	Waste excavator (F8) Area B	80.00	0.76	2.00	0.58	0.21	1,034
Dump Truck	Ore dump truck (F9) Area A	150.00	0.34	2.00	0.26	0.63	3,245
	Waste dump truck (F10) Area A	150.00	0.34	2.00	0.26	0.63	5,192
	Ore dump truck (F11) Area B	150.00	0.34	2.00	0.26	0.63	3,245
	Waste dump truck (F12) Area B	150.00	0.34	2.00	0.26	0.63	3,894
Primary Crusher	Mechanical breakdown of primary crusher (F13)	186.00	0.83	4.00	0.41	0.75	7,355

Table A.4: Model parameters for the open pit copper ore production operation

Risk model block	Name	Value		Units
		Area A	Area B	
DRILLING BLOCK	Drilling schedule from mine planning at fixed intervals	1	1	days
CHARGING & BLASTING BLOCK	Waste to ore ratio in each blast	4:1	3:1	
	Gross quantity of waste/ore in each blast	11000	8200	tons/day
	Variability in each blast	10%	10%	of gross qty.
EXCAVATOR BLOCK	Interval of loading ore by excavator	$1/(3*36)$	$1/(3*27)$	in days
	Interval of loading waste by excavator	$1/(12*36)$	$1/(9*27)$	in days
	Gross ore quantity of each excavator bucket	100	100	in tons
	Gross waste quantity of each excavator bucket	100	100	in tons
	Variability of ore quantity in each excavator bucket	+5	+5	in tons
	Variability of waste quantity in each excavator bucket	+5	+5	in tons
STOCKPILE/CONTROL LOGIC	Qty. of packets for sending to stockpile/crusher	100	80	in tons
	Intergeneration time for sending to stockpile	15	15	in mins

Table A.5: Equipment quantity and cost for the open pit copper ore production operation

Equipment Prices	Numbers	dollars/unit	Total equipment costs
Primary crusher	1	4,387,000	4,387,000
Area A			
Hydraulic shovel	1	1,214,000	1,214,000
Wheel loader	1	1,474,000	1,474,000
Rear dump trucks	15	1,108,500	16,627,500
Rotary drill (crawler)	1	620,740	620,740
Crawler tractor (dozer)	4	320,900	1,283,600
Road grader	1	465,000	465,000
Water truck	1	40,175	40,175
Service truck	4	61,800	247,200
Shot loader truck	1	85,500	85,500
Explosives truck	1	92,600	92,600
Area B			
Hydraulic shovel	1	962,000	962,000
Wheel loader	1	912,500	912,500
Rear dump trucks	12	1,108,500	13,302,000
Rotary drill (crawler)	1	620,740	620,740
Crawler tractor (dozer)	4	320,900	1,283,600
Road grader	1	465,000	465,000
Water truck	1	40,175	40,175
Service truck	3	61,800	185,400
Shot loader truck	1	85,500	85,500
Explosives truck	1	92,600	92,600
Total Equipment costs			44,486,830 \$
			28,026,703 £

Table A.6: Sample questionnaire with user inputs and relative weights for calculating the catastrophic risk index and the management risk index for an underground coal mine. The hatched user inputs changes the relative importance of the questions that follow whose weights are marked in green colour. The questions in red font indicate those associated with management risks.

	Risk indicators	Monitoring/Measures	Weights/ Relative importance	User Input	Weighted Index
1	Spontaneous combustion	1. Considered at design stage?	10	Yes	0
		2. Intrinsic Risk Factors: Rank of coal (Low/high)	7	High	14
		3. Extrinsic Risk Factors: Is the Mining method a risk factor?	7	No	14
		4. CO levels monitored? Mine wide monitoring?	10	No	100
		6. Gas sampling & analysis?	8	No	80
		7. Is there proper action plan?	5	Yes	0
2	Methane related risks	1. Is the mine classified as gassy?	10	Yes	100
		2. Can the mine keep the methane concentration at the face and the return gate <1% at all times?	10	Yes	0
		3. Mine wide monitoring?	0	Yes	0
		4. Face monitoring?	10	No	100
		5. Gas sampling & analysis	8	No	80
		6. Methane drainage monitoring?	0	No	0
		7. Suitability of methane drainage with mining method?	0	No	0
		8. Poor management?			
		8.1 Flame proofing of electrical equipment?	4	No	40
		8.2 Intrinsic safety measures (gassy mines must have certification)	5	No	50

		8.3 Avoiding frictional ignitions?	6	No	60
		8.4 Sprays on machines?	5	No	50
		8.5 Continuous maintenance and calibration of monitoring equipment	10	No	100
3	Coal dust explosions	1. Considered at design stage?	10	No	100
		2. Is the mine gassy?	10	No	0
		3. Dust levels monitored? (Hand held electrostatic equipment)	3	Yes	0
		4. Regulating ventilation?	4	No	40
		5. Dust control measures?			
		5.1 Cleaning dust from belt conveyors, roadways	4	No	40
		5.2 Have stone dust or water barriers?	6	Don't Know	30
		5.3 Maintained properly or not?	2	No	20
		6. Poor management?			
		6.1 Avoiding frictional ignitions?	6	No	60
		6.2 Sprays on machines?	3	No	30
4	Outburst	1. Considered at design stage?	10	Yes	0
		2. Are the coal seams outburst prone?	10	Yes	100
		3. Is mining method leading to high stress concentrations?			
		3.1 Gas pressure and coal permeability as outburst indicators?	7	No	70
		3.2 Mining in fault zones? Tectonically disturbed zones?	5	Yes	50

		3.3 Are there high pressure free gas pockets in coal?	7	No	0
		4. Probing and injecting water by fracturing coal?	7	No	70
		5. Fault monitoring?	8	No	80
		6. Gas drainage?	10	Yes	0
5	Roof fall	1. Geological area risk prone? Earthquakes?	2	No	0
		2. Induced seismicity due to mining?	10	Yes	100
		2.1 Pillar robbing?	10	Yes	100
		2.2 Are gate roads and deformation monitored (extensometers?)	5	No	50
		2.3 Are the supports suited to the mining method?	8	No	80
		2.4 Are the supports maintained well?	7	No	70
6	Flood	1. Considered at design stage?	10	Yes	0
		2. Water body nearby?	5	Yes	50
		3. Old workings filled with water nearby?	5	Don't Know	25

Table A.7: Risk model parameters for the underground soft rock (coal) model.

Components	Failure modes					
	TBF (exponential)		OD (lognormal)			Mbcost factor
	Mult. Fact.	mu	Mult. Fact.	mu	sigma	£/day
Continuous miner	80	0.85	4	0.23	0.31	514
Belt conveyor	142	0.76	4	0.15	0.35	1,756
Bunker	189	0.72	2	0.49	0.28	4,366
Shaft/hoist	285	0.91	3	0.12	0.29	4,767
Shearer	173	0.54	4	0.84	0.71	10,051
AFC	235	0.56	2	0.12	0.57	3,997
Stage loader & Impact breaker	161	0.53	3	0.33	0.79	298

Table A.8: Equipment quantity and costs for the underground coal mining case

Equipment Prices	Numbers	dollars/unit	Total equipment costs
Longwall Shearer	2	3,196,000	6,392,000
Face Shields	260	174,700	45,422,000
Armored Face Conveyor	2	3,045,000	6,090,000
Electrical Power Pack	2	24,000	48,000
Hydraulic Power Pack	2	24,000	48,000
Mobile Roof Support	4	579,000	2,316,000
Continuous Miner /Roadheader	3	2,144,500	6,433,500
Feeder/Breakers	2	390,600	781,200
Gateroad Conveyor	2	2,011,000	4,022,000
Primary Haul Conveyor	6	975,400	5,852,400
Production Hoist	1	6,740,500	6,740,500
Service Hoist	1	3,930,200	3,930,200
Bin Loader Conveyor	2	59,000	118,000
Load-out Storage Bins	4	115,300	461,200
Load-out Bin Dust Collector	4	6,600	26,400
Percussion Rockbolter Drill	2	1,480,000	2,960,000
Power Supply	1	200,000	200,000
Primary Ventilation Fan	2	177,900	355,800
Auxiliary Ventilation Fan	4	16,000	64,000
Mine Drainage Pumps	4	35,000	140,000
Development Drill Jumbo	1	1,125,300	1,125,300
Supply LHD/Scoop or Rail Transport system	4	1,354,100	5,416,400
Personnel Carriers/Rail or other	2	236,700	473,400
Total Equipment costs			99,416,300 \$
			62,632,269 £

Table A.9: Sample questionnaire with user inputs and relative weights for calculating the catastrophic risk index and the management risk index for an underground metal mine. The questions in red font indicate those associated with management risks.

	Risk indicators	Monitoring/Measures	Weights/ Relative importance	User Input	Weighted Index
1	Roof fall	1. Geological area risk prone? Earthquakes?	2	Don't Know	10
		2. Induced seismicity due to mining?	10	No	0
		3. Pillar robbing?	10	No	0
		4. Are gate roads and deformation monitored (extensometers?)	5	No	50
		5. Are the supports suited to the mining method?	8	Yes	0
		6. Is the extraction area extensively roof bolted?	7	Yes	0
		7. Are the supports maintained well?	7	Yes	0
2	Stope Failure	1. Considered at design stage?	10	Yes	0
		2. Is rockmass strength and deformation behaviour characterised?	5	No	50
		3. Is mining method leading to high stress concentrations?	8	No	0
		4. Monitoring stope instability?	8	Yes	0
		5. Are supports inadequate or are there adverse loading conditions?	10	No	0
		6. Are supports well maintained and monitored?	8	Yes	0
3	Backfill failure	1. Geological area risk prone? Earthquakes?	2	Don't Know	10
		2. Induced seismicity due to mining?	10	No	0
		3. Has poor quality backfill been used?	10	No	0
4	Flood	1. Considered at design stage?	10	Yes	0

		2. Water body nearby?	5	No	0
		3. Old workings filled with water nearby?	5	Don't Know	25

Table A.10: Risk model parameters for the underground metal (Zn-Pb) mine model

Components	Failure modes					
	TBF (exponential)		OD (lognormal)			Mbcost factor
	Mult. Fact.	mu	Mult. Fact.	mu	sigma	£/day
Drilling	58.33	0.36	2.67	0.21	0.32	153
Charging & Blasting	37.50	0.43	2.25	0.33	0.45	1828
LHD	56.00	0.81	2.20	0.27	0.67	11,139
Primary Crusher	60.00	0.92	3.00	0.48	0.54	17,208
Hauling	96.67	0.67	1.67	0.51	0.33	986
Bunker (Service operations)	52.50	0.55	1.00	0.18	0.15	177
Shaft/ skip	150.00	0.83	3.00	0.26	0.29	566
Backfill plant	100.00	0.69	1.00	0.36	0.11	5,924

Table A.11 : Equipment quantity and costs for underground metal mining case

Equipment Prices	Numbers	dollars/unit	Total equipment costs
Bulk explosive chargers	1	653,100	653,100
Low profile compressors	2	48,600	97,200
Continuous haulage belt conveyor systems	1	274,600	274,600
Crawler-Mounted Belt Tailpiece	1	217,600	217,600
Complete Eight Unit Continuous Haulage System with Feeder Breaker	1	2,503,000	2,503,000
Roadheader	1	2,144,500	2,144,500
Drill Jumbo type 1	1	955,800	955,800
Drill Jumbo type 2	1	646,200	646,200
Complete Blasthole drilling system	1	407,350	407,350
LHD (81 hp)	2	502,700	1,005,400
LHD (185 hp)	2	864,000	1,728,000
Mine hoist (drum type)	1	2,795,000	2,795,000
Roof bolters	2	978,000	1,956,000
Scaler	1	575,000	575,000
Shotcreting systems	1	71,000	71,000
Rear dump trucks	2	869,000	1,738,000
Utility vehicles			
> Service truck	2	319,000	638,000
> Lube truck	2	319,000	638,000
> ANFO loader truck	2	419,000	838,000
> Personnel carrier	2	236,700	473,400
> Multi purpose vehicle	2	98,000	196,000
Axial Ventilation Fan	2	173,000	346,000
Axial fan accessories	2	48,200	96,400
Auxilliary Ventilation Fan	1	16,000	16,000
Gyratory crusher	1	10,264,000	10,264,000
Slurry pump and backfill plant	2	50,000	100,000
Mine drainage pumps	2	19,700	39,400
Total Equipment costs			31,412,950 \$
			19,790,159 £

Underground mining process models

Figure A.1 shows the process operation for an underground room and pillar coal mining operation. The main development phase and the production development phase are almost similar to the longwall mining case. The primary production also includes pillar extraction in this case and employs a continuous miner instead of a

shearer as in the longwall case. Due to the distributed nature of the room and pillar mining operations, shuttle cars are used instead of armoured face conveyors (AFCs).

Figure A.2-Figure - A.4 shows the process operation schematics for different underground hard rock mining methods. All of them have the same basic operations of drilling, charging and blasting, loading and hauling, crushing and then transporting it overground through a shaft or drift. However there are slight differences in operation depending on the mining method. For example in the sub-level caving (in Figure A.2), the drifting and support operations are performed first, followed by fan drilling. For the block caving mining method in Figure A.3, there is a secondary drawing which is sometimes performed when the ores get stuck in the chutes. The cut and fill method (in Figure A.4) has additional backfilling operations in the production phases.

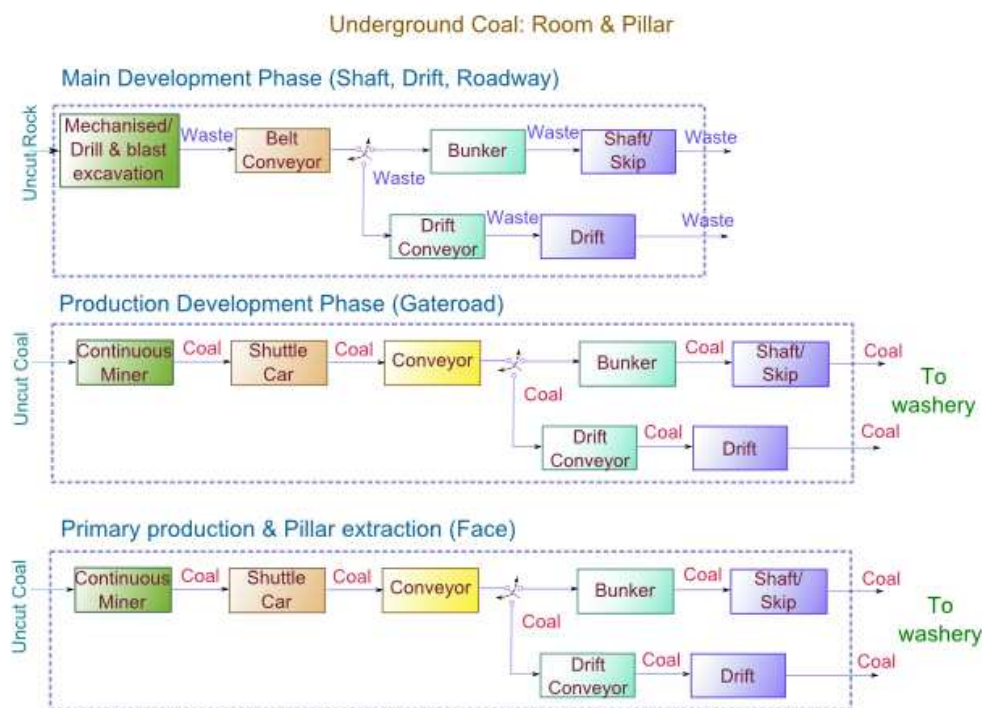


Figure A.1: Process model for an underground room and pillar coal mining operation

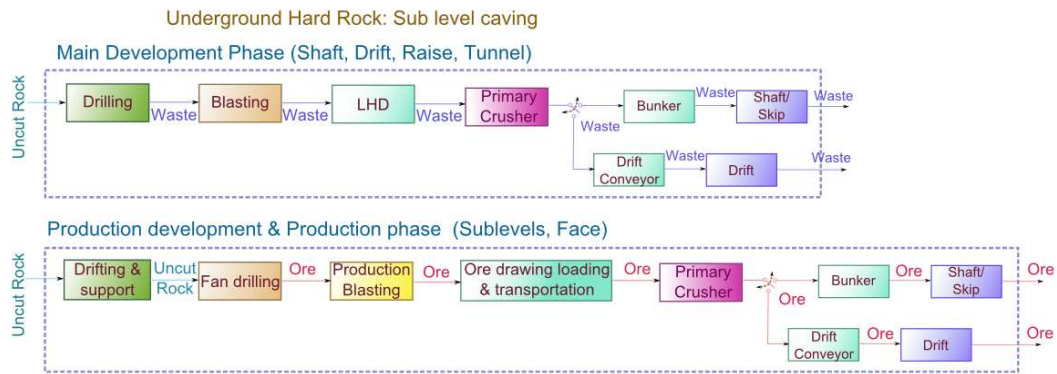


Figure A.2: Process model for an underground hard rock sublevel caving operation

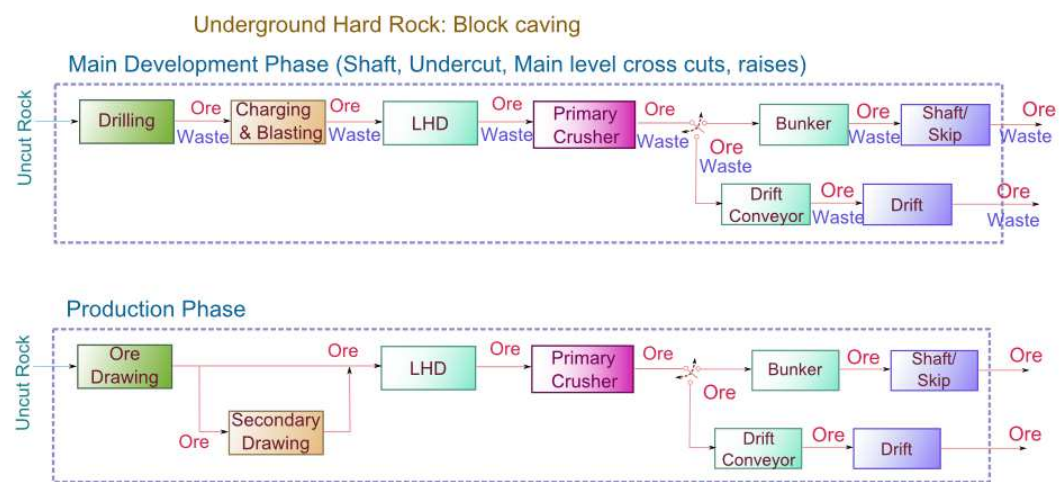


Figure A.3: Process model for an underground hard rock block caving operation

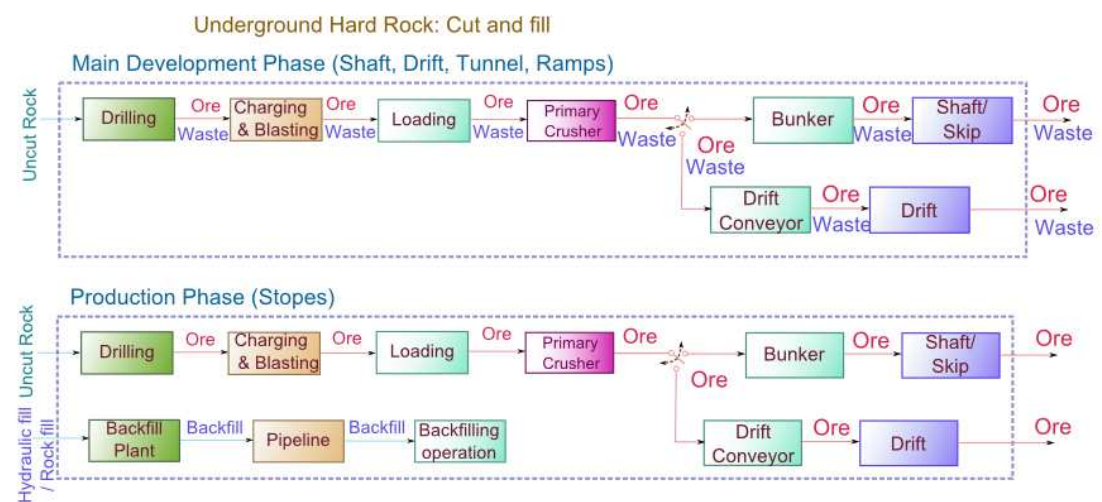


Figure A.4: Process model for an underground cut and fill operation

References

- Abbaspour, K.C., 1994. Bayesian risk methodology for crop insurance decisions. *Agricultural and forest meteorology* 71, 297–314.
- Ades, A., Lu, G., 2003. Correlations between parameters in risk models: estimation and propagation of uncertainty by Markov Chain Monte Carlo. *Risk Analysis* 23, 1165–1172.
- Ahmad, R., Kamaruddin, S., Azid, I.A., Almanar, I.P., 2012. Failure analysis of machinery component by considering external factors and multiple failure modes- A case study in the processing industry. *Engineering Failure Analysis* 25, 182-192.
- Bakheet, M.T., 1995. Contractors risk assessment system (surety, insurance, bonds, construction, underwriting). Ph. D. Dissertation, Georgia Institute of Technology.
- Bauwens, L., Laurent, S., Rombouts, J.V.K., 2006. Multivariate GARCH models: a survey. *Journal of applied econometrics* 21, 79–109.
- Beyer, H.-G., 2000. Evolutionary algorithms in noisy environments: theoretical issues and guidelines for practice. *Computer methods in applied mechanics and engineering* 186, 239–267.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics* 31, 307–327.
- Brockett, P., Cooper, W., Golden, L., Xia, X., 1997. A case study in applying neural networks to predicting insolvency for property and casualty insurers. *Journal of the Operational Research Society* 1153–1162.
- Brockett, P.L., Cooper, W.W., Golden, L.L., Pitaktong, U., 1994. A neural network method for obtaining an early warning of insurer insolvency. *Journal of Risk and Insurance* 402–424.

- Brockett, P.L., Xia, X., Derrig, R.A., 1998. Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud. *Journal of Risk and Insurance* 245–274.
- Brooks, C., Clare, A., Dalle Molle, J.W., Persaud, G., 2005. A comparison of extreme value theory approaches for determining value at risk. *Journal of Empirical Finance* 12, 339–352.
- Byström, H.N.E., 2005. Extreme value theory and extremely large electricity price changes. *International Review of Economics & Finance* 14, 41–55.
- Cassandras, C.G., Lafontaine, S., 2008. *Introduction to discrete event systems*. Springer.
- Castillo, E., 1988. Extreme value theory in engineering (Statistical modeling and decision science). *Recherche* 67, 02.
- Checkland, P., 1999. *Systems thinking, systems practice: includes a 30-year retrospective*, Wiley.
- Chib, S., Nardari, F., Shephard, N., 2002. Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* 108, 281–316.
- Chinbat, U., 2011. *Risk Analysis in the Mining Industry*. InTechOpen.
- Chinbat, U., Takakuwa, S., 2009. Using simulation analysis for mining project risk management, in: *Winter Simulation Conference (WSC), Proceedings of the 2009*, 2612–2623.
- Coello, C.A., 2000. An updated survey of GA-based multiobjective optimization techniques. *ACM Computing Surveys (CSUR)* 32, 109–143.
- Dalla Valle, L., 2009. Bayesian copulae distributions, with application to operational risk management. *Methodology and Computing in Applied Probability* 11, 95–115.
- Danielsson, J., 2010. *Financial risk forecasting*. Wiley Finance.
- Darwiche, A., Casico, K., Allen, D., Chan, H., Chavira, M., Park, J., Zalzenny, D., Zalzenny, M., 2001, SamIam: Sensitivity analysis, modeling, inference, and more. *Software available from <http://reasoning.cs.ucla.edu/samiam>*.
- Davenport, W.G., Biswas, A.K., 2002. *Extractive metallurgy of copper*. Pergamon Pr.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on* 6, 182–197.

- Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock market returns and a new model. *Journal of empirical finance* 1, 83–106.
- Dougherty, E., 1989. Human Reliability Analysis. *IEEE Transactions On Reliability* 38, 509.
- Embrechts, P., Resnick, S.I., Samorodnitsky, G., 1999. Extreme value theory as a risk management tool. *North American Actuarial Journal* 3, 30–41.
- Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* 987–1007.
- Fonseca, C.M., Fleming, P.J., 1995. An overview of evolutionary algorithms in multiobjective optimization. *Evolutionary computation* 3, 1–16.
- Genccay, R., Selccuk, F., Ulugülyağci, A., 2003. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics* 33, 337–356.
- Gilli, M., Këllezi, E., 2006. An application of extreme value theory for measuring financial risk. *Computational Economics* 27, 207–228.
- Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of finance* 1779–1801.
- Goebel, R., Sanfelice, R.G., Teel, A., 2009. Hybrid dynamical systems. *Control Systems, IEEE* 29, 28–93.
- Goonan, T.G., 2005. Flows of selected materials associated with world copper smelting. US Geological Survey.
- Hardy, M.R., 2002. Bayesian risk management for equity-linked insurance. *Scandinavian Actuarial Journal* 2002, 185–211.
- Hasofer, A., Lind, N.C., Division, U. of W.S.M., 1973. An exact and invariant first-order reliability format. University of Waterloo, Solid Mechanics Division.
- Heuberger, R., 2005. Risk analysis in the mining industry. *Journal of the South African Institute of Mining and Metallurgy* 105, 75–80.
- Hohenbichler, M., Rackwitz, R., 1981. Non-normal dependent vectors in structural safety. *Journal of the Engineering Mechanics Division* 107, 1227–1238.
- Holt, C.C., 2004. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting* 20, 5–10.
- Horn, J., Nafpliotis, N., Goldberg, D.E., 1994. A niched Pareto genetic algorithm for multiobjective optimization, in: *Evolutionary Computation, 1994. IEEE World*

- Congress on Computational Intelligence., Proceedings of the First IEEE Conference on., 82–87.
- Iman, R.L., Helton, J.C., 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis* 8, 71–90.
- Jang, J. S. R., Sun, C. T. 1996. *Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence*. Prentice-Hall, Inc.
- Jarrow, R.A., Lando, D., Turnbull, S.M., 1997. A Markov model for the term structure of credit risk spreads. *Review of Financial studies* 10, 481–523.
- Jin, Y., Branke, J., 2005. Evolutionary optimization in uncertain environments-a survey. *Evolutionary Computation, IEEE Transactions on* 9, 303–317.
- Johansson, K.H., Egerstedt, M., Lygeros, J., Sastry, S., 1999. On the regularization of Zeno hybrid automata. *Systems & Control Letters* 38, 141–150.
- Kapur, K.C., Lamberson, L.R., 1977. *Reliability in engineering design*. New York, John Wiley and Sons, Inc.
- Katz, R., 1999. Extreme value theory for precipitation: Sensitivity analysis for climate change. *Advances in Water Resources* 23, 133–139.
- Kijima, M., 1998. Monotonicities in a Markov chain model for valuing corporate bonds subject to credit risk. *Mathematical Finance* 8, 229–247.
- Kijima, M., Komoribayashi, K., 1998. A Markov chain model for valuing credit risk derivatives. *The Journal of Derivatives* 6, 97–108.
- Knowles, J.D., Corne, D.W., 2000. Approximating the nondominated front using the pareto archived evolution strategy. *Evolutionary computation* 8, 149–172.
- Koller, D., Friedman, N., 2009. *Probabilistic graphical models: principles and techniques*. MIT press.
- Lauridsen, K., Christou, M., Amendola, A., Markert, F., Kozine, I., Fiori, M., 2001. Assessing the uncertainties in the process of risk analysis of chemical establishments: Part 1 and 2, in: *Proceedings of the International Conference on ESREL*.
- Lavaja, J.H., Bagajewicz, M.J., 2004. Managing financial risk in the planning of heat exchanger cleaning. *Computer Aided Chemical Engineering* 18, 235–240.
- Li, L., Wang, J., Leung, H., Jiang, C., 2010. Assessment of catastrophic risk using Bayesian network constructed from domain knowledge and spatial data. *Risk analysis* 30, 1157–1175.
- Li, L. F., Wang, J. F., & Leung, H., 2010a. Using spatial analysis and Bayesian network to model the vulnerability and make insurance pricing of catastrophic

- risk. *International Journal of Geographical Information Science*, 24-12, 1759-1784.
- Li, L., Wang, J., Leung, H., Zhao, S., 2012. A Bayesian method to mine spatial data sets to evaluate the vulnerability of human beings to catastrophic risk. *Risk analysis*, 32-6, 1072-1092.
- LICYMIN, 2003, Life cycle assessment of mining projects for waste minimisation and long term control of rehabilitated sites, EU Project Report.
- Lygeros, J., Johansson, K.H., Sastry, S., Egerstedt, M., 1999. On the existence of executions of hybrid automata, in: *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*. pp. 2249–2254.
- Marhavilas, P., Koulouriotis, D., 2011. A combined usage of stochastic and quantitative risk assessment methods in the worksites: Application on an electric power provider. *Reliability Engineering & System Safety* 97.1, 36-46.
- Matlab Inc., 2010. Matlab Stateflow Userguide.
- McKee, T.E., Lensberg, T., 2002. Genetic programming and rough sets: A hybrid approach to bankruptcy classification. *European Journal of Operational Research* 138, 436–451.
- Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica: Journal of the Econometric Society* 347–370.
- Olivieri, A., Pitacco, E., 2010. *Introduction to insurance mathematics: technical and financial features of risk transfers*. Springer Verlag.
- Pan, I., Das, S., 2013. Frequency Domain Design of Fractional Order PID Controller for AVR System Using Chaotic Multi-objective Optimization. *International Journal of Electrical Power & Energy Systems*.
- Pan, I., Korre, A., Das, S., Durucan, S., 2012. Chaos suppression in a fractional order financial system using intelligent regrouping PSO based fractional fuzzy control policy in the presence of fractional Gaussian noise. *Nonlinear Dynamics* 70, 2445–2461.
- Park, J., 1993. *Bankruptcy prediction of banks and insurance companies: an approach using inductive methods*, UT Austin.
- Podofillini, L., Dang, V., 2012. Conventional and Dynamic Safety Analysis: Comparison on a Chemical Batch Reactor. *Reliability Engineering & System Safety* 106, 146-159.
- Podofillini, L., Zio, E., Mercurio, D., Dang, V., 2010. Dynamic safety assessment: Scenario identification via a possibilistic clustering approach. *Reliability Engineering & System Safety* 95, 534–549.

- Rackwitz, R., Flessler, B., 1978. Structural reliability under combined random load sequences. *Computers & Structures* 9, 489–494.
- Raphael, W., Faddoul, R., Feghaly, R., Chateauneuf, A., 2011. Analysis of Roissy Airport Terminal 2E collapse using deterministic and reliability assessments. *Engineering Failure Analysis* 20, 1–8.
- Robinson, D.G., 1998. A survey of probabilistic methods used in reliability, risk and uncertainty analysis: Analytical techniques i. Sandia National Lab, Report SAND98-1189.
- Rosenblatt, M., 1952. Remarks on a multivariate transformation. *The Annals of Mathematical Statistics* 23, 470–472.
- Ross, T.J., 2009. *Fuzzy logic with engineering applications*. John Wiley & Sons.
- Rubinson, T., Yager, R.R., 1996. Fuzzy logic and genetic algorithms for financial risk management, in: *Computational Intelligence for Financial Engineering, Proceedings of the IEEE/IAFE 1996 Conference on.*, 90–95.
- Shapiro, A.F., 2002. The merging of neural networks, fuzzy logic, and genetic algorithms. *Insurance: Mathematics and Economics* 31, 115–131.
- Shapiro, A.F., 2004. Fuzzy logic in insurance. *Insurance: Mathematics and Economics* 35, 399–424.
- Shin, K.S., Lee, Y.J., 2002. A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications* 23, 321–328.
- Taylor, S.J., 1986. *Modelling financial time series*. John Wiley & Sons Ltd.
- Tomlin, C.J., Mitchell, I., Bayen, A.M., Oishi, M., 2003. Computational techniques for the verification of hybrid systems. *Proceedings of the IEEE* 91, 986–1001.
- Tsutsui, S., Ghosh, A., 1997. Genetic algorithms with a robust solution searching scheme. *Evolutionary Computation, IEEE Transactions on* 1, 201–208.
- Varetto, F., 1998. Genetic algorithms applications in the analysis of insolvency risk. *Journal of Banking & Finance* 22, 1421–1439.
- Vaughn, M.L., Ong, E., Cavill, S., 1997. Interpretation and knowledge discovery from a multilayer perceptron network that performs whole life assurance risk assessment. *Neural Computing & Applications* 6, 201–213.
- Venkatasubramanian, V., Zhao, J., Viswanathan, S., 2000. Intelligent systems for HAZOP analysis of complex process plants. *Computers & Chemical Engineering* 24, 2291–2302.
- Wu, C.H., Tzeng, G.H., Goo, Y.J., Fang, W.C., 2007. A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting

- bankruptcy. *Expert Systems with Applications* 32, 397–408.
- Wu, W., Cheng, G., Hu, H., Zhang, Y., 2012. A knowledge-based reasoning model using causal table for identifying corrosion failure mechanisms in refining and petrochemical plants. *Engineering Failure Analysis*.
- Yang, J., Huang, H.Z., He, L.P., Zhu, S.P., Wen, D., 2011. Risk evaluation in failure mode and effects analysis of aircraft turbine rotor blades using Dempster-Shafer evidence theory under uncertainty. *Engineering Failure Analysis* 18, 2084–2092.
- Zhang, J., Johansson, K.H., Lygeros, J., Sastry, S., 2001. Zeno hybrid systems. *International Journal of Robust and Nonlinear Control* 11, 435–451.
- Zhang, Q., Li, H., 2007. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *Evolutionary Computation, IEEE Transactions on* 11, 712–731.
- Zimmerman, D., Wahl, K., Gutjahr, A., Davis, P., 1990. A review of techniques for propagating data and parameter uncertainties in high-level radioactive waste repository performance assessment models. Sandia National Labs., Albuquerque, NM (USA).
- Zitzler, E., Thiele, L., Zitzler, E., Zitzler, E., Thiele, L., Thiele, L., 1998. An evolutionary algorithm for multiobjective optimization: The strength Pareto approach., Swiss Federal Institute of Technology Zürich (ETH), 43.