

https://dspace.org/

A methodology for evaluating utilisation of mine planning software and consequent decision-making strategies in South Africa

Gen, Bekir

2015-05-05
http://hdl.handle.net/10539/17626

Downloaded from a DSpace repository. https://dspace.org/

A METHODOLOGY FOR EVALUATING UTILISATION OF MINE PLANNING SOFTWARE AND CONSEQUENT DECISION-MAKING STRATEGIES IN SOUTH AFRICA

Bekir Genç

A thesis submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2015

DECLARATION

I declare that this thesis is my own, unaided work. Where use has been made of the work of others, it has been duly acknowledged. It is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before in any form for any degree or examination in any other University.

Signed:

_____

Bekir Genc

This _____ day of _____ 2015

ABSTRACT

Mine planning software has and continues to contribute to the development of the South African mining industry. As mine planning software usage continues to be more widespread, it is imperative that a methodology to evaluate mine planning software utilisation for enhanced decision-making strategies in South Africa is established. An existing online database available on the website link http://db.mining.wits.ac.za was developed prior to this study in September 2012 (initial data collection date). However, the database only acted as a snapshot of mine planning software data repository and lacked a framework to evaluate utilisation of mine planning software in the South African mining industry.

In this thesis, a methodology was developed to measure the utilisation of mine planning software to enhance decision-making strategies in the South African mining industry. The methodology for the evaluation of utilisation of mine planning software in various commodity sectors was developed on the basis of three variables, namely, commodity, functionality, and time factor, as a key evaluation criteria. Even though the calculations can be done on any commodity in a similar manner, in this research, calculations were only performed on four different commodities, namely coal, diamond, gold and platinum group metals which are the most significant minerals in South Africa. Six functionalities namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation were applied on the four different commodities using two different time-stamps (September 2012 and April 2014). The following software providers availed information that was used to populate the database: Geovia, MineRP Solutions, Sable, RungePincockMinarco, Maptek, Cyest Technology and CAE Mining. Note that the CAE Mining data was only made available in April 2014 (second data collection date). However, the results indicated that the market leaders in terms of mine planning software utilisation in South Africa differs, depending on the commodity that is being mined as well as the functionality that is being used.

In addition, this thesis also proposed a framework to estimate the future use of mine planning software on an evolving dataset by considering the fact that the database will be continually updated in the future. By using Artificial Neural Networks (ANN), specifically supervised learning, time-series analyses were performed. Results from

the time-series analyses were used to establish the framework for estimating the future use of mine planning software utilisation in the South African mining industry. By using this newly developed framework, utilisation of the various mine planning software was measured leading to the formulation of different decision-making strategies for the various mine planning software stakeholders.

By using this newly developed framework to estimate and measure mine planning software utilisation, and proposing a framework for time-series analyses on an evolving dataset, this thesis serves a number of beneficiaries; firstly, the South African mining industry to position themselves better by acquiring optimal combination of mine planning software that is being used in South Africa so that they can improve their production levels, secondly, tertiary education institutions and mining consulting firms which make use of mine planning software, and lastly, the aforementioned software providers by strategically positioning themselves in a limited mine planning software market. However, this newly developed framework could be used by involved parties for corporate strategic decision-making.

LIST OF PUBLICATIONS

The publications listed below have emanated from this research work so far:

1. Katakwa, TP, Musingwini, C and **Genc, B (2013)** Online database of mine planning and peripheral software used in the South African mining industry. *Journal of The Southern African Institute of Mining and Metallurgy*, Volume 113, No. 6, June 2013, pp. 497-504. An electronic copy of the paper is available on SAIMM website at INTERNET: http://www.saimm.co.za/Journal/v113n06p497.pdf
2. **Genc, B**, Musingwini, C and Katakwa, TP **(2014)** Estimating mine planning software utilisation for decision-making strategies in the South African platinum group metals mining sector. In *Proceedings of the 6th International Platinum Conference* (20-22 October), Sun City, South Africa, pp. 1-10.
3. **Genc, B**, Musingwini, C and Celik, T **(2015)** Estimating mine planning software utilisation for decision-making strategies in the South African gold mining sector. Accepted by the *Journal of The Southern African Institute of Mining and Metallurgy*, in July, 2014.
4. **Genc, B**, Musingwini, C and Celik, T **(2015)** Measuring "optimisation" software utilisation in the South African mining industry. Accepted by the *International Journal of Mining, Reclamation and Environment*, in October, 2014. DOI: 10.1080/17480930.2014.993833

The paper abstracts are included in Appendix 9.1.

ACKNOWLEDGEMENTS

I would like to acknowledge the following individuals for their guidance and encouragement, and the following companies for their contributions in making this research work successful:

- My supervisor, Prof C. Musingwini (Head, School of Mining Engineering, University of Witwatersrand), for his invaluable academic guidance throughout the research and reading through draft chapters and draft theses.
- Prof T. Celik (University of Witwatersrand, School of Computer Science) for providing insights into various scripting issues.
- Prof F.T. Cawood (Former Head, School of Mining Engineering, University of Witwatersrand), for providing School's support needed to successfully execute the research programme.
- Mr. T.P. Katakwa (Cyest Corporation) for assistance during database updates.
- My wife Farhana and sons Irfaan and Erhan for selflessly giving their love and patience while I undertook this research work.
- Software companies that participated by providing the much-needed data to populate the database.

Lastly, the opinions expressed in this thesis are those of the author and may not necessarily represent the policies of the companies and institutions mentioned in this thesis.

DEDICATION

To my late mother, father and brother who have been my pillar of support in life.

CONTENTS **Page**

DECLARATION ..... **II**

ABSTRACT ..... **III**

LIST OF PUBLICATIONS ..... **V**

ACKNOWLEDGEMENTS ..... **VI**

DEDICATION ..... **VII**

LIST OF FIGURES ..... **XI**

LIST OF TABLES ..... **XV**

LIST OF ABBREVIATIONS ..... **XVII**

1 INTRODUCTION **1**

1.1 CHAPTER OVERVIEW 1
1.2 PROBLEM STATEMENT 1
1.3 RELEVANCE OF THE RESEARCH AND RESEARCH QUESTION 2
1.4 OBJECTIVES OF THE RESEARCH 3
1.5 RESEARCH METHODOLOGY 5
1.6 OVERVIEW OF HOW STAKEHOLDERS CAN USE THE METHODOLOGY DEVELOPED 6
1.6.1 Mining companies 6
1.6.2 Educational institutions 6
1.6.3 Mining consulting companies 7
1.6.4 Software providers 7
1.7 STRUCTURE OF THESIS 7
1.8 SUMMARY 8

2 LITERATURE REVIEW **9**

2.1 INTRODUCTION 9
2.2 CATEGORIES OF THE MINE PLANNING SOFTWARE 12
2.2.1 Mapping software 14
2.2.2 Geological data management software 15
2.2.3 Geological modelling and resource estimation software 16
2.2.4 Design and layout software 18
2.2.5 Planning and scheduling software 20
2.2.6 Financial valuation software 21
2.2.7 Mine optimisation software 22
2.2.8 Other software (i.e. ERP, monitoring, reporting, etc.) 24
2.2.8.1 ERP software 24
2.2.8.2 Monitoring and reporting software 24
2.2.8.3 Simulation software including simulators 26
2.2.8.4 Virtual Reality software 27
2.3 SUMMARY 28

3 ONLINE SOFTWARE DATABASE 29

3.1 INTRODUCTION 29
3.2 RATIONALE FOR COLLATING MINE PLANNING SOFTWARE 29
3.3 MINE PLANNING AND PERIPHERAL SOFTWARE DATABASE 30
3.4 DATABASE STRUCTURE 32
3.5 DATABASE FEATURES 37
3.6 DISCUSSION OF THE RESULTS 42
3.7 SUMMARY 43

4 DEFINING AND DERIVING MINE PLANNING SOFTWARE UTILISATION .. 45

4.1 INTRODUCTION 45
4.2 OVERALL EQUIPMENT EFFECTIVENESS 45
4.3 SOFTWARE UTILISATION 48
4.4 RESULTS FOR COAL 54
4.4.1 Geological Data Management software results for coal 55
4.4.2 Geological Modelling and Resource Estimation software results for coal 57
4.4.3 Design and Layout software results for coal 58
4.4.4 Scheduling software results for coal 61
4.4.5 Financial Valuation software results for coal 63
4.4.6 Optimisation software results for coal 64
4.5 RESULTS FOR PLATINUM GROUP METALS 66
4.5.1 Geological Data Management software results for PGM 67
4.5.2 Geological Modelling and Resource Estimation software results for PGM 68
4.5.3 Design and Layout software results for PGM 70
4.5.4 Scheduling software results for PGM 72
4.5.5 Financial Valuation software results for PGM 76
4.5.6 Optimisation software results for PGM 77
4.6 DISCUSSION OF THE RESULTS AND INDICATIVE DECISION-MAKING STRATEGIES . 79
4.7 SUMMARY 80

5 PREDICTING SOFTWARE UTILISATION USING ARTIFICIAL NEURAL NETWORKS 82

5.1 INTRODUCTION 82
5.2 WHAT IS ANN? 82
5.2.1 Pattern classification 84
5.2.2 Function approximation 84
5.2.3 Prediction/forecasting 85
5.3 ANNS OVERVIEW 86
5.3.1 Neurons 86
5.3.2 Architecture/Topology of an ANN 88
5.3.3 Supervised learning 89
5.3.4 Error-correction rule 90
5.4 FEEDFORWARD NETWORKS 92
5.4.1 Single-layer Feedforward Networks 93
5.4.2 Multilayer Feedforward Networks 93
5.5 TRAINING/LEARNING OF ANNS 97
5.6 SUMMARY 98

6 TIME-SERIES ANALYSIS WITH ARTIFICIAL NEURAL NETWORKS 99

6.1 INTRODUCTION 99
6.2 WHAT IS A TIME-SERIES? 99

6.3 TIME-SERIES ANALYSIS USING BACKPROPAGATION FEEDFORWARD NEURAL NETWORKS 100
6.3.1 Data simulation 102
6.3.2 Data validation 104
6.3.3 Performance evaluation 104
6.4 BACKPROPAGATING FEEDFORWARD NEURAL NETWORK 105
6.5 PREDICTING SOFTWARE UTILISATION LAGS 109
6.6 SIMULATION RESULTS 111
6.7 DISCUSSION OF THE RESULTS 120
6.8 SUMMARY 120

7 CONCLUSIONS AND RECOMMENDATIONS 122

7.1 INTRODUCTION 122
7.2 KEY FINDINGS 122
7.3 RESEARCH CONTRIBUTIONS: METHODOLOGY AND DECISION-MAKING STRATEGIES 124
7.4 RESEARCH LIMITATIONS 125
7.5 RECOMMENDATIONS FOR FUTURE RESEARCH WORK 125

8 REFERENCES 127

9 APPENDICES 134

9.1 SUMMARY OF PAPER ABSTRACTS ON THE RESEARCH 134
9.1.1 Paper 1 134
9.1.2 Paper 2 135
9.1.3 Paper 3 136
9.1.4 Paper 4 137
9.2 THE SCRIPT DETAILS USED TO CALCULATE MINE PLANNING SOFTWARE UTILISATION 138
9.2.1 Utilisation 138
9.2.2 Validation 141
9.3 THE SCRIPT DETAILS USED TO PREDICT TIME-SERIES ANALYSES 151
9.3.1 Backpropagation 151
9.3.2 Feedforward ANN 151
9.3.3 Forecasting 151
9.3.4 Simulations 154
9.3.5 Testing 155
9.3.6 Training 156

LIST OF FIGURES

Figure **Page**

Figure 2.1: GEOVIA Surpac displaying drillholes (GEOVIA Surpac, 2013) 13

Figure 2.2: ArcGIS geological cross sections and map subsurface features in 2D and 3D (Esri, 2014) 15

Figure 2.3: SABLE Data Warehouse software logging drillhole data (Sable, 2013) 16

Figure 2.4: CAE's Studio software, showing access to the orebody (adopted from CAE Studio3, 2013) 17

Figure 2.5: Maptek Vulcan software pit design (Vulcan, 2013) 19

Figure 2.6: GEOVIA MineSched creating activities (GEOVIA MineSched, 2013) ... 21

Figure 2.7: RungePincockMinarco's XERAS logo (RungePincockMinarco's XERAS, 2013) 22

Figure 2.8: CAE NPV Scheduler – Mine flow optimiser (CAE NPV Scheduler, 2013) 23

Figure 2.9: GEOVIA Whittle - Core module application (GEOVIA Whittle, 2013) 23

Figure 2.10: Modular Mining Systems' DISPATCH software monitoring equipment movement (Modular's DISPATCH, 2013) 25

Figure 2.11: Immersive Technologies CM360-B simulator for coal mining (Immersive Technologies, 2013) 26

Figure 2.12: ViMINE VR software virtually showing 3D orebody (ViMINE, 2013).... 28

Figure 3.1: How the server-side scripting produces dynamic web pages (PHP and MySQL) (Bogotobogo, 2014) 32

Figure 3.2: Summary of data representation within the mine planning software database (Katakwa, et al., 2013) 35

Figure 3.3: Website structure for the web-based interface of the database (Katakwa, et al., 2013) 36

Figure 3.4: Homepage of the database (Katakwa et al., 2013) 36

Figure 3.5: Main features of navigation and interface design for the web-based interface of the database (Katakwa, et al., 2013) 38

Figure 3.6: Available software in the database with their functionality (Katakwa, et al., 2013) 39

Figure 3.7: Distribution of commodities at mining sites 40

Figure 3.8: Adoption of mine planning software 40

Figure 3.9: Market impact analysis (using number of mining sites) 41

Figure 3.10: Relative market impact for software 41

Figure 3.11: Distribution of software client sites by company profile 42

Figure 4.1: Time factors effecting equipment utilisation (adapted from Shirose, 2013) 47

Figure 4.2: Updated list of mine planning software including CAE Mining software and their functionalities along the mining value chain 53

Figure 4.3: Geological Data Management functionality software utilisation for coal 56

Figure 4.4: Geological Modelling and Resource Estimation functionality software results for coal 58

Figure 4.5: Design and Layout functionality software utilisation for coal for both time-stamp 60

Figure 4.6: Scheduling functionality software utilisation for coal for both time-stamp 63

Figure 4.7: Financial Valuation functionality software utilisation for coal 64

Figure 4.8: Optimisation functionality software utilisation for Coal for both time-stamp 66

Figure 4.9: Geological Data Management functionality software utilisation for PGM 68

Figure 4.10: Geological Modelling and Resource Estimation functionality software utilisation for PGM 70

Figure 4.11: Design and Layout functionality software utilisation for PGM 72

Figure 4.12: Scheduling functionality software utilisation for PGM as of September 2012 75

Figure 4.13: Scheduling functionality software utilisation for PGM as of April 2014 75

Figure 4.14: Financial Valuation functionality software utilisation for PGM 77

Figure 4.15: Optimisation functionality software utilisation for PGM 79

Figure 5.1: Schematic drawing of a neuron (Jacobson, 2013) 83

Figure 5.2: Commonly used neural networks tools: a) Pattern classification; b) Function approximation and c) Prediction/forecasting; (adopted from Jain et al., 1996) 86

Figure 5.3: Neuron model (McCulloch and Pitts, 1943) 87

Figure 5.4: Different types of activation functions (McCulloch and Pitts, 1943) 88

Figure 5.5: Categories of network architectures (Jain et al., 1996) 88

Figure 5.6: Feedforward Network architecture with three-layers (Jain et al., 1996). 93

Figure 5.7: The role of hidden units (Jain et al., 1996) 96

Figure 6.1: An error function with four stationary points (Bishop, 1996) 102

Figure 6.2: Impact of sampling rate to the data simulation process ($M = 1000,\ K = 4$ with variable ts) 103

Figure 6.3: Impact of sampling rate to the data simulation process ($M = 100,\ K = 4$ with variable ts) 103

Figure 6.4: Additional sampling rate examples ($M = 100,\ K = 4$ with variable ts) 104

Figure 6.5: Architecture of Feedforward Neural Network considered in utilisation prediction 107

Figure 6.6: Training of Feedforward Neural Network 108

Figure 6.7: Obtaining software utilisation prediction of Feedforward Neural Network 109

Figure 6.8: Training FNN to predict $uk(t+1)$ software utilisation 110

Figure 6.9: Predicting $uk(t+2)$ software utilisation 110

Figure 6.10: Predicting $uk(t+3)$ software utilisation 111

Figure 6.11: How RMSE values change with T_f 112

Figure 6.12: Actual value comparison with different prediction stages 112

Figure 6.13: Initial software utilisation before prediction 114

Figure 6.14: Software utilisation prediction for k=1 (from *t*+1 to *t*+6) 114

Figure 6.15: Software utilisation prediction for k=1 (from *t*+7 to *t*+10) with RMSE values 115

Figure 6.16: Software utilisation prediction for k=2 (from *t*+1 to *t*+6) 115

Figure 6.17: Software utilisation prediction for k=2 (from *t*+7 to *t*+10) with RMSE values 116

Figure 6.18: Software utilisation prediction for k=3 (from *t*+1 to *t*+6) 116

Figure 6.19: Software utilisation prediction for k=3 (from *t*+7 to *t*+10) with RMSE values 117

Figure 6.20: Software utilisation prediction for k=4 (from *t*+1 to *t*+6) 117

Figure 6.21: Software utilisation prediction for k=4 (from *t*+7 to *t*+10) with RMSE values..... 118

Figure 6.22: Software utilisation prediction for *t*+1 118

Figure 6.23: Software utilisation prediction for *t*+3 119

Figure 6.24: Software utilisation prediction for *t*+5 119

Figure 6.25: Software utilisation prediction for *t*+10 120

LIST OF TABLES

Table **Page**

Table 2.1: Most commonly used mine planning software in South Africa 10

Table 3.1: Categorisation of data (Katakwa et al., 2013) 34

Table 4.1: Number of software solutions by company 49

Table 4.2: Software providers and their headquarter locations 55

Table 4.3: Geological Data Management functionality software utilisation for coal . 56

Table 4.4: Geological Modelling and Resource Estimation functionality software results for coal 57

Table 4.5: Design and Layout functionality software utilisation for coal as of September 2012 59

Table 4.6: Design and Layout functionality software utilisation for coal as of April 2014 60

Table 4.7: Scheduling functionality software utilisation for coal as of September 2012 61

Table 4.8: Scheduling functionality software utilisation for coal as of April 2014 62

Table 4.9: Financial Valuation functionality software utilisation for coal 63

Table 4.10: Optimisation functionality software utilisation for coal as of September 2012 65

Table 4.11: Optimisation functionality software utilisation for coal as of April 2014 . 66

Table 4.12: Geological Data Management functionality software utilisation for PGM as of September 2012 67

Table 4.13: Geological Data Management functionality software utilisation for PGM as of April 2014 68

Table 4.14: Geological Modelling and Resource Estimation functionality software utilisation for PGM as of September 2012 69

Table 4.15: Geological Modelling and Resource Estimation functionality software utilisation for PGM as of April 2014 69

Table 4.16: Design and Layout functionality software utilisation for PGM as of September 2012 71

Table 4.17: Design and Layout functionality software utilisation for PGM as of April 2014 72

Table 4.18: Scheduling functionality software utilisation for PGM as of September 2012 73

Table 4.19: Scheduling functionality software utilisation for PGM as of April 2014 . 74

Table 4.20: Financial Valuation functionality software utilisation for PGM..... 76

Table 4.21: Optimisation functionality software utilisation for PGM as of September 2012 78

Table 4.22: Optimisation functionality software utilisation for PGM as of April 2014 78

Table 5.1: Well-known learning algorithms (adopted from Jain et al., 1996) 89

Table 6.1: Data showing Actual and Predicted values 113

Table 7.1: Commodity/Functionality table showing most utilised mine planning software in 2014 124

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
APMOT	Anglo Platinum Mine Optimisation Tool
DMR	Department of Mineral Resources
ERP	Enterprise Resource Planning
FNN	Feedforward Neural Network
GDP	Gross Domestic Product
JSE	Johannesburg Stock Exchange
KPI	Key Performance Indicator
ML	Machine Learning
MLP	Multilayer perceptron
MRO	Mineable Reserves Optimizer
NPV	Net Present Value
OEE	Overall Equipment Effectiveness
PGM	Platinum Group Metals
RMSE	Root-Mean-Square Error
SAIMM	Southern African Institute of Mining and Metallurgy
SIMRAC	Safety in Mines Research Advisory Committee
TOP	Trimble Open Pit Design
TPM	Total Productive Maintenance
VR	Virtual Reality

1 INTRODUCTION

1.1 Chapter overview

South Africa is one of the countries in the world with significant mineral resources. It is the largest producer of platinum, and one of the leading producers of gold, diamonds, base metals and coal. The South African mining industry's contribution towards the country's gross domestic product (GDP) varies between 5% and 10% and employs around 500,000 people (Department of Mineral Resources, 2012). The Chamber of Mines of South Africa (2014) reported that the mining sector alone helped to create 1,365,892 jobs in the South African economy in 2012, which includes both direct and indirect jobs. Furthermore, the South African mining sector is a very important part of the Johannesburg Stock Exchange (JSE), and it accounts for almost 25% of the exchange (Chamber of Mines, 2014). According to the 2011 Citigroup Bank report, South Africa was ranked as the world's richest country in terms of its mineral reserves, worth an estimated USD 2.5 trillion followed by Russia and Australia (Prinsloo & Marais, 2014). To capitalise on these valuable mineral resources, it is vital for the country to have a mining industry that is capable of utilising sector specific mine planning software to improve efficiencies and remain significant to the South African economy.

1.2 Problem statement

Kapageridis (2005) indicated that almost every aspect of the mining industry is today using some form of mine planning software. From exploration to rehabilitation, the use of software is becoming more widespread (Kapageridis, 2005).

South Africa is no exception to observations made by Kapageridis (2005). Despite mine planning software in the South African mining industry continuing to be more widespread in use, there is no established methodology to evaluate mine planning software utilisation to enhance decision-making strategies in South Africa. The first attempt to collect data on the extent of mine planning software utilisation was the development of an online database by Katakwa et al. (2013). Although this database could be seen as an attempt to establish a baseline to display the extent of mine

planning software usage in South Africa, it lacked a methodology to evaluate mine planning software utilisation for decision-making strategies in the South African mining industry which is crucial to improve efficiencies within the mining sector. To improve efficiencies within the mining sector, a methodology to evaluate mine planning software utilisation for enhancing decision-making strategies in South Africa needed to be developed.

After successful development of this methodology to measure the level of mine planning software utilisation, it is also important to predict the future use of software utilisation within reasonable confidence levels by conducting time-series analyses. Through these analyses, the future market influence of different mine planning software can be predicted in this limited market where mergers and acquisitions often occur. Furthermore, the developed methodology allows software to be compared but interoperability of the software will still be a long lasting issue as long as there is no agreed platform to enhance compatibility issues within mine planning software providers. The Open Group (2010) has been attempting to address the issue of interoperability but with limited success to date.

1.3 Relevance of the research and research question

This research provides the South African mining industry with a methodology to increase their efficiency through evaluating level of utilisation of the various mine planning software. Choosing the correct mine planning software solution is very important as they add value to the mining operations.

The South African mining industry is always seeking ways to improve the value of mining operations through advanced software solutions. For instance, Whittle & Burks (2010) explained that the net present value (NPV) of a mine can be increased by 5-35% or more, often even before the expenditure of significant amounts of capital by using their new 10-step simultaneous mining optimisation software. This can be achieved by using advanced business models focusing on factors such as pit shell optimisation, underground mine design, scheduling, cut-off grades and stockpiling. Using a number of case studies, it was established that the Anglo Platinum Mine Optimisation Tool (AMPOT) can improve mine optimisation when used in conjunction

with mine planning software such as CADSMine and Mine2-4D (Smit & Lane, 2010). These examples demonstrate the fact that correct software solutions improved production efficiencies.

From the foregoing sections it can be noted that this thesis seeks to provide answers to the question, *"as the mine planning software usage in South African mines is becoming more widespread, can a methodology be developed in such a way that it leads to the evaluation of the utilisation of the mine planning software by each specific mining sector in South Africa so that the evaluation can enhance decision-making strategies on software utilisation in the South African mining industry?"*.

The research is relevant to the South African mining industry, mine planning software providers, institutions that offer mining related qualifications as well as mining consultancy companies in several ways. The South African mining industry will benefit as this research provides a methodology to evaluate mine planning software utilisation to enhance decision-making strategies. Through these enhanced decision-making strategies, the South African mining industry will be in a position to choose optimal software combinations, which will lead to improved production efficiencies. Educational institutions in South Africa such as the University of the Witwatersrand (WITS) and University of Pretoria will also benefit out of this research by exposing their students to the strategically important mine planning software solutions. As mergers and acquisitions are always a possibility, the mine planning software providers will understand their market impact so that they can develop appropriate decision-making strategies for future growth. Mining consultancy companies will also benefit from this research as the methodology provided through this thesis will help them to make the right choices of software solutions so that they can better serve their clientele.

1.4 Objectives of the research

Following on from the research question in Section 1.3, software usage in the mining industry is impacted by several reasons which include diversity, efficiency, capability and cost. Therefore the main objective of the research was to develop a methodology to evaluate mine planning software utilisation and construct broad decision-making

strategies relevant for the South African mining industry. Although Katakwa et al. (2013) initiated a new approach to determine the extent of mine planning software use in South Africa by establishing an online database called the “Mine Planning and Peripheral Software Database”, at http://db.mining.wits.ac.za, initially a lack of participation from CAE Mining, which is one of the main software providers in South Africa, as well as the limited functionality of the database (i.e. unable to produce proven techniques to predict future use of mine planning software) caused the online database to achieve limited success. This limited success formed a need to bring a new approach to the question of the evaluation of mine planning software utilisation for enhanced decision-making strategies in South Africa. In this approach, using the online database as a baseline, first CAE Mining was added to the database and then existing software providers’ datasets updated in April, 2014. Furthermore, a framework was developed to measure the level of software utilisation within the mining industry by establishing a logical relationship between commodity, functionality software (six different stages of mine planning namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation) and the time-factor.

Furthermore, the second objective of this research was to predict the future use of the mine planning software utilisation. After calculating the software utilisation index, consequently the database was prepared to perform time-series analyses (mine planning software market trend analysis) related research (which demonstrated how the utilisation of the mine planning software changed over time). This was achieved by using machine learning, specifically supervised learning techniques. Due to the fact of the high bias, as well as the incomplete nature of the data, and by considering the future development of the database as it is an evolving database, machine learning techniques were employed to alleviate the problem. Furthermore, Artificial Neural Networks, a division of Machine Learning, are the most widely used network structure in time-series analysis and they are easy to configure and train.

In order to achieve these objectives, the following were set out:

- Design and implement scripting language for estimating mine planning software utilisation for enhanced decision-making strategies in South Africa;

- Design and implement a temporal software utilisation dataset for time-series analyses;
- Capture the data compiled to populate the temporal database; and
- Design and implement machine learning and statistical pattern recognition methods on temporal database to achieve enhanced decision-making strategies.

1.5 Research methodology

The database was released in September, 2012. In the initial release version (beta version) of the database, the time-stamp for the information about mining companies (operating mines, quarries and processing plants) which was obtained from the Department of Mineral Resources website (http://www.dmr.gov.za) was the year 2010. Software vendors' information was up to date, with a September 2012 date-stamp. In order to achieve the objectives of this research stated in Section 1.4, the following tasks were carried out:

- CAE Mining data obtained;
- Existing software providers' datasets updated;
- Database re-analysed and a framework was developed to measure the level of software utilisation within the mining industry;
- A logical relationship between commodity and the functionality software was established by using the newly developed framework; and
- A framework was proposed for time-series analyses on an evolving dataset by considering the fact that the database will be continually updated in the future.

Most current data in the database were obtained from three main sources, namely:

- Software providers;
- Department of Mineral Resources; and
- Consulting firms.

After completion of the data collection, new data was added to the database with the appropriate date-stamp, i.e. April 2014, and the initial dataset was labelled as September 2012. Then the database had two different date-stamps for data entry and the database was prepared for time-series analyses. Machine learning and statistical pattern recognition techniques were applied on this temporal dataset to perform trend analysis. Furthermore, supervised machine learning techniques were used to generate a functional form of future software utilisation.

1.6 Overview of how stakeholders can use the methodology developed

The developed methodology can be used by a number of beneficiaries. These beneficiaries include mining companies, educational institutions, mining consulting companies and software providers. This information can then be used further too, to better inform decision-making strategies on software utilisation by these stakeholders to meet their different perspectives.

1.6.1 Mining companies

The South African mining companies will be the main beneficiary of these developments. By using this newly developed framework, utilisation of the various mine planning software were measured. Now, the South African mining industry has a framework that can be utilised in order to ultimately enhance decision-making strategies for utilisation of the mine planning software. The methodology can be used by existing software users who are reviewing their software combinations or those intending to purchase new software and would like to estimate the comparative attractiveness of a certain software package.

1.6.2 Educational institutions

The South African educational institutions play a very crucial role when it comes to preparing tomorrow's engineers, geologists, etc. to the industry. Due to this vital fact, it is essential for the educational institutions to expose their students to the most utilised software solutions. They can now make use of the framework to choose which software solutions would be of best interest to them.

1.6.3 Mining consulting companies

Mining consulting companies add value to the mining operations by bringing the expertise required to improve a number of key factors. Consequently, they can now advise their clients more effectively by using the developed framework.

1.6.4 Software providers

Software providers are the main custodians behind mine planning software development. The fact is that due to the size of the available market for the software providers being very limited, it is important for the software providers to compare their positions so that they can develop informed decision-making strategies accordingly.

1.7 Structure of thesis

This thesis is divided into seven chapters. Chapter 1 discusses the importance of the mining industry in South Africa, and elaborates some of the facts contributing to this statement. Then it links this importance to the mine planning software that is being used in South Africa, and talks about the significance of determining the estimation of mine planning software utilisation for enhanced decision-making strategies in South Africa. Chapter 2 is a literature survey and presents background information about the mine planning software that is being used in South Africa. It starts by giving details about some of the leading mine planning software that is being used in South Africa and then it expands on to the significance of the thesis by categorising mine planning software.

Chapter 3 discusses the online database at http://db.mining.wits.ac.za and gives further details about the temporal nature of the database. Chapter 4 discuses a new methodology to define and measure mine planning software utilisation in the South African mining sector within an evolving dataset framework. Chapter 5 explains the concept of the Artificial Neural Networks, machine learning and the statistical pattern recognition techniques, and motivates the use of these techniques for time-series analyses.

Chapter 6 covers the use of computational methods detailed in Chapter 5 to perform time-series analyses and proposes a framework to predict future software utilisation. Conclusions and recommendations are made in Chapter 7.

1.8 Summary

This chapter endeavoured to introduce the reader to some background information, the main research question, motivation and structure of the thesis. Furthermore, it explained why it was necessary to highlight the importance of evaluating the utilisation of mine planning software and listed several other parties, besides the South African mining industry, who will be the other beneficiaries through this research. The next chapter itemises the findings of a survey of literature on topics relevant to this research.

2 LITERATURE REVIEW

2.1 Introduction

Mine planning software plays a vital role in every stage of mining operations, even before the actual mining commences. Gibbs (1994) stated that mine planning software had changed radically in the five years prior to 1994. This statement is still valid today, twenty years after it was made.

Since the early days of mine planning software development, modelling methods, optimization methods, and many of the basics have been proven to work. Development in recent years has focused more on refinements, such as solid modelling, sophisticated visualization of models and mine designs, and better graphical user interfaces.

The dependence of a mining operation on a reasonably accurate 3D geological model is growing for all scales of operations, from small quarries to large open pit and underground mines. This trend can be attributed to an increasing demand for efficiency in mining with managers expecting reliable information on a regular basis. All requests relating to the resource or reserve ultimately lead back to the geological model on which planning and production activities are based (Johnstone, 2003).

Although there are several mine planning software available in the market, Table 2.1 shows some of the most commonly used mine planning software in South Africa. It is important to note that some of these software were not captured in the database as they did not participate and these are typed italics for identification purposes. The software shown in Table 2.1 is not just specific to South Africa. However, other mine planning software used elsewhere in the world include:

- MineSight
- MicroMine
- Trimble Open Pit Software

Table 2.1: Most commonly used mine planning software in South Africa

Software	Description
acQuire	acQuire is the complete user geoscientific data manager to make informed decisions (acQuire, 2013).
ArcGIS	ArcGIS is the most complete desktop GIS (ArcGIS, 2013).
CADSMine	CADSMine supports the creation of 3D mining elements to represent a mining framework for the extraction of an orebody (CADSmine, 2013).
CAE NPV Scheduler	NPV Scheduler - the most versatile open pit mine planning tool on the market (CAE NPV Scheduler, 2013).
CAE Studio	Previously known as Datamine, is the advanced mining planning tool (CAE Studio3, 2013).
CAE Studio 5D Planner	CAE Studio 5D Planner is an advanced mine design and scheduling software (CAE Studio 5D Planner, 2014).
Deswik	Leaders in innovative mine planning software and consultancy, the company was formed by the developers of the Mine2-4D mine planning software (Deswik, 2013).
GEOVIA Gems	Well-known mine planning software (GEOVIA GEMS, 2013).
GEOVIA MineSched	Designed to improve scheduling output (GEOVIA MineSched, 2013).
GEOVIA Minex	GEOVIA Minex provides mine planning solutions for coal deposits (GEOVIA Minex, 2013).
GEOVIA Surpac	Surpac is very well-known mine planning software (GEOVIA Surpac, 2013).
GEOVIA Whittle	Whittle is well known surface mining optimisation tool (GEOVIA Whittle, 2013).
Mine2-4D	Mine2-4D complete tool for mine planning (Mine2-4D, 2013).
Minemax Scheduler	Minemax specializes in mine planning and scheduling solutions (Minemax, 2013).
MineSight	Geomodelling and Mine Planning Software (MineSight, 2013).
Rocscience	Rock engineering software (RocScience, 2013).
SABLE Data Warehouse	SABLE is a geological data management software (Sable, 2013).
Surfer	Surfer is a well-known tool used from mapping to 3D visualization (Surfer, 2013).
Vulcan	Vulcan is a mine planning software (Vulcan, 2013).
Xeras Financial Evaluation	RungePincockMinarco's XERAS is a well-known software for financial valuation of the mining projects (XERAS, 2014).
XPAC Mine Scheduling	RungePincockMinarco's XPAC is a business focused mine scheduling application for forecasting, reserve and mine scheduling management (XPAC, 2014).

Although it is difficult to distinguish the actual reasons behind the acceptance of a particular mine planning software as most software have similar features, the relationship between software vendors and customers (mining industry), as well as after-sales service (customer support) play a very important role to decide which

software to choose and use. The other factors that influence decision makers to choose which software to purchase are:

- User preference;
- Continuity of the current structure; and
- Other factors as discussed below.

The user preference is important as it plays a crucial role to decide which software to use. The staff turnover within the mining industry is historically high, hence competent mine planning software users tend to move around quite often. When there is a new mine planning software user, for example a mine planner is appointed, if the new appointee is not familiar with the existing mine planning software package, the mine may consider purchasing a new mine planning software package which the new appointee is familiar with.

On the other hand, continuity of the current structure is also very important. Although there are a number of initiatives afoot to bring file-exchange standards within mine planning software companies, there is no probable agreement insight as yet. This forces mine planning software users to continue using their existing software or upgrade to the newer versions of the existing software packages.

There are also other factors that influence which mine planning software package to purchase. For example a surface operation that eventually becomes an underground operation will need to switch software on transition. Opoku & Musingwini (2012) listed some mines which are on open pit to underground transition. Some of the mining packages are designed to handle either surface or underground mining operations only. The change in the mining type, whether it is surface or underground mining, will also change the mining method that is in use. Consequently, this will most probably require a change of software package in use, too.

2.2 Categories of the mine planning software

One of the objectives of this research was to check which software is used in the South African mines. There are a number of mine planning software available in the market which were listed in Table 2.1. Therefore it is important to understand how to categorize the software accordingly.

Currently, major mine planning software producers are trying to combine most of the commonly used software features to sell as bundled software. For example, some of the features of GEOVIA Surpac (2013) are listed below to support this observation:

- Exploration;
- Geological Modelling;
- Geology;
- Geophysics;
- Geostatistics;
- Geotechnical;
- Mapping;
- Mine Design and Planning;
- Reserve Evaluation;
- Statistics; and
- Surveying.

GEOVIA Surpac (2013) software is also capable of displaying the terrain in 3D. Figure 2.1 shows a 3D pit design with drillholes.

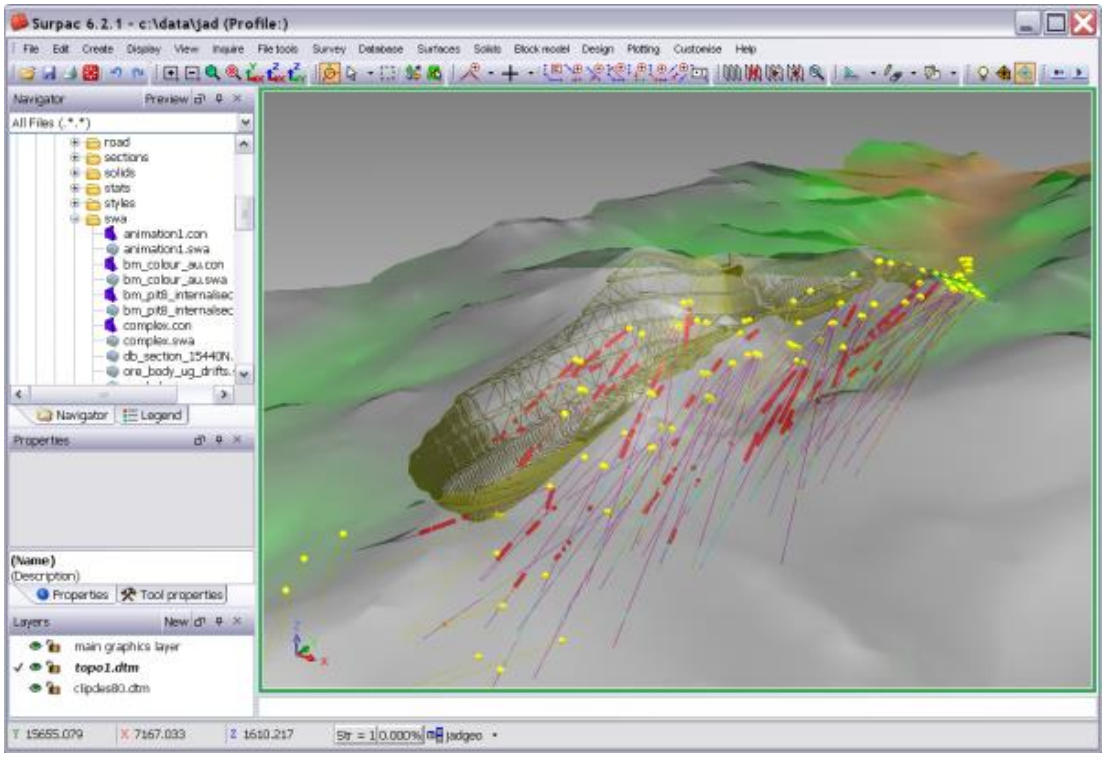


Figure 2.1: GEOVIA Surpac displaying drillholes (GEOVIA Surpac, 2013)

Being able to offer a range of functionalities improves the marketability of the software. Although the software companies are trying very hard to feature as much functionality as possible, in many cases the software that is being purchased by the mining company is not adequate for the intended purpose.

Every piece of software comes with limitations. To overcome some of the challenges faced by mining companies, software vendors started offering additional modules to the existing software. For example GEOVIA Surpac is not designed to do scheduling but by purchasing GEOVIA MineSched as an additional module, the scheduling part of the mining operation can also be handled by GEOVIA Surpac.

Based on Open Group's Business Reference Model (2010) which discusses mining methods and mine planning software functionalities, Katakwa et al. (2013, p. 499) listed these categories under eight main headings, namely:

- "Mapping;
- Geological data management;

- Geological modelling and resource estimation;
- Design and layout;
- Scheduling;
- Financial Valuation;
- Optimisation; and
- Other (Monitoring, ERP, etc.)".

Each of these categories is discussed in detail in the next section with an example of mine planning software being given for each section.

2.2.1 Mapping software

Geographic Information Systems (GIS) and products, including maps, and land information are designed to manage natural resources effectively. Mapping software plays a very important role in terms of understanding more about land information. By using mapping software, producing geological maps and providing technical and general information on the area of interest's various geological settings can be easily achieved. GIS technology provides a framework for gathering, organizing, sharing, and analysing spatial information. Esri's ArcGIS (2013) can create, analyse, and share maps and applications throughout an entire organization in such a way that Esri's ArcGIS can help integrating and analysing data for improved decision-making.

According to Latimer (2013, p. 1), Tom Gardener from Esri Australia stated that by using ArcGIS "a miner can follow their operation throughout the life of the project, from the moment they are mapping it out for environmental surveys through to drillhole locations, ore deposits, actual mining, and eventually, rehabilitation works". Figure 2.2 shows that by using ArcGIS scientists can create geological cross sections and map subsurface features in 3D.

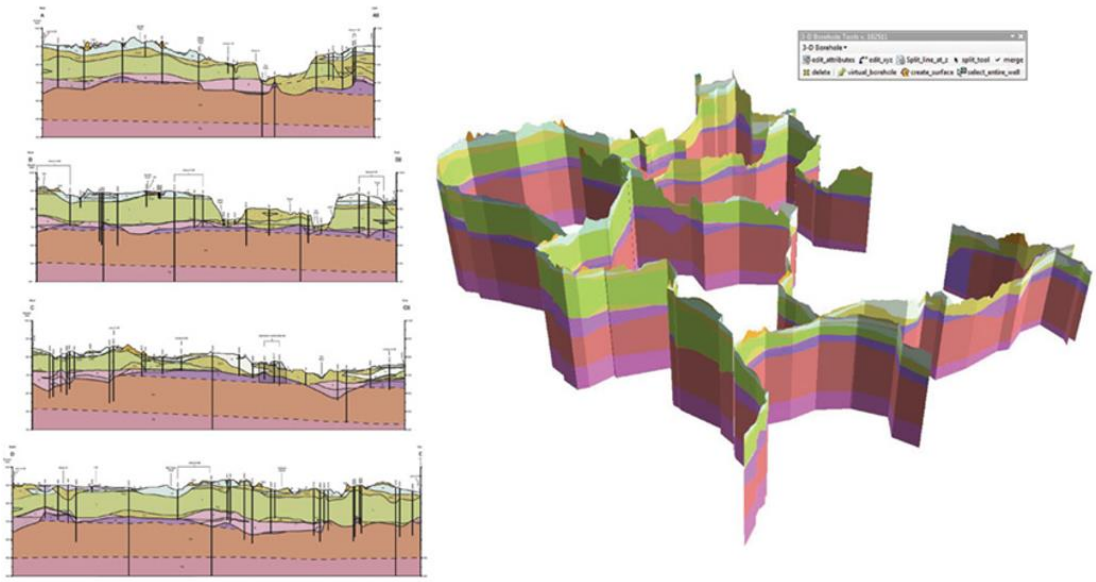


Figure 2.2: ArcGIS geological cross sections and map subsurface features in 2D and 3D (Esri, 2014)

2.2.2 Geological data management software

In any mining operation, understanding the geology of the orebody is vital for subsequent stages of a mining operation such as mine design and layout, scheduling, financial valuation and mineral processing. Better geological understanding will improve the profitability of the planned operation. Exploration data lays the foundation for better understanding of the resource and reserve calculations. For this reason, it is essential that geological data needs to be captured and modelled reliably.

Standardised Approach to Borehole Logging for Exploration (SABLE) is a South African geological data management company that was established in 1983. Since then SABLE has developed a number of software solutions to interpret the geological and sampling data from the drilling operation. SABLE Data Warehouse is its most advanced product to date (Sable, 2013). Figure 2.3 illustrates SABLE's logging of drillhole data.



Figure 2.3: SABLE Data Warehouse software logging drillhole data (Sable, 2013)

2.2.3 Geological modelling and resource estimation software

Geological modelling and/or resource estimation is also one of the most important stages of the mining operation. During this stage, the geological model of the orebody needs to be created. Both the 2D and 3D orebody models give a different perspective of the planned mining operation, hence this enables the operator to calculate the available resources (Bascetin & Sertabipoglu, 2010). Bascetin and Sertabipoglu (2010) broadly describe orebody modelling stages as including the following:

- Creating a database by using various data (maps, drilling, etc.);
- Compositing drillhole data along drillhole length;
- Creating cross-sections;
- Creating a block model;
- Estimating of values (grades, thickness, density, etc.); and
- Calculating the resources.

CAE's Studio software, previously known as Datamine software, is one of the leading software in orebody modelling. Some of the software features include (CAE Studio3, 2013, p. 2):

- "Ability to handle complex data sets including drillhole samples, channel samples, face maps, laser and image processed 3D models;
- Wire-framing and modelling tools that help define the structural geology quickly and reliably;
- Geotechnical modelling that provides a detailed understanding of ground stability and rock characteristics;
- Sophisticated block modelling with comprehensive sub-celling to enable detailed representation of grade and structure;
- Multi-parametric modelling capabilities that model geo-metallurgical and geotechnical parameters in addition to grade; and
- A Mineable Reserves Optimizer (MRO) that allows the stoping blocks to be identified to turn a geological resource into a mineable reserve".

Figure 2.4 shows a screen shot of the CAE's Studio3 software. In this screen shot, 3D model of the orebody, access roads and shafts are clearly visible.

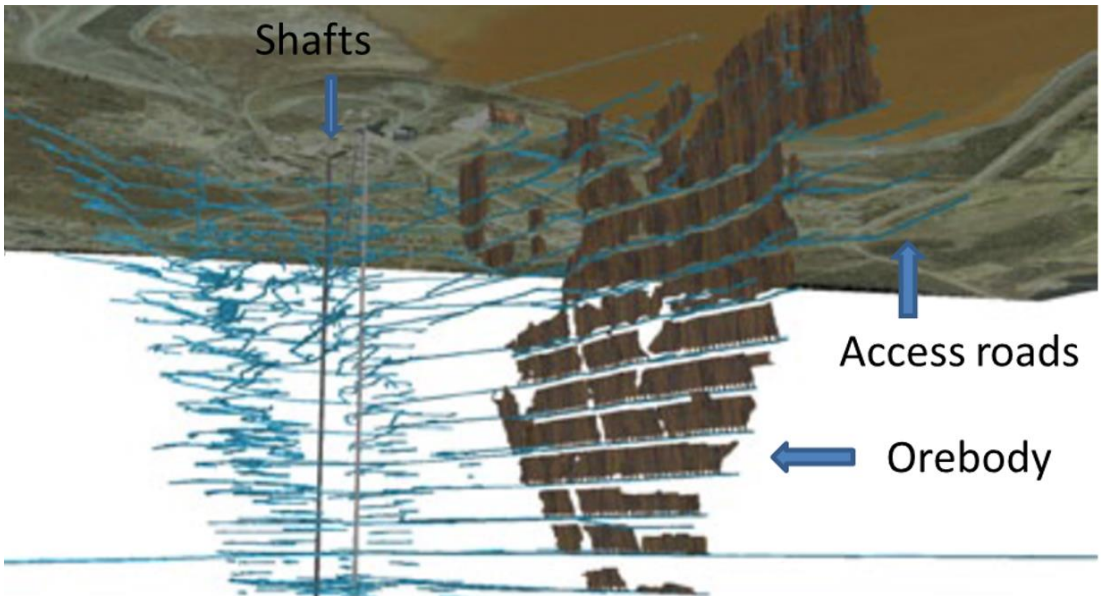


Figure 2.4: CAE's Studio software, showing access to the orebody (adopted from CAE Studio3, 2013)

2.2.4 Design and layout software

The mine design and layout software is one of the most important software in terms of the mine planning software that is being used in the South African mining industry. As previously mentioned in Chapter 1, mine planning software plays a very important role in every stage of the mining operation. The mine design software also helps users to determine the best possible mining methods as well as the initial economic feasibility of the potential operation. It also plays a crucial role in terms of creating a 3D orebody model, the mining method selection and designing and planning of access roads and extraction of the orebody.

Mine design is subject to a number of natural constraints. These constraints include the size and shape of the orebody, the chemical, physical and mechanical properties of the rocks, ore grade, impurities, groundwater, surrounding terrain, waterways, environment, and the weather. Many of these constraints need to be estimated or predicted from the various data samples obtained through drilling, and geophysical surveys or sampling (Johnson, 2011).

Johnson (2011) further stated that there is always uncertainty about what is under the surface, hence risk analysis must be performed. The relatively high start-up capital costs requires that the design life of a mine must extend sufficiently past the capital payback period and generate an adequate return on the overall investment. This long-term operating plan requires forecasting, albeit with associated uncertainty, of various economic conditions such as input costs, commodity prices and exchange rates. Modelling and interpretation of the orebody is the most critical task affecting the success of the mine design process, and the mine design and planning software plays an important role to successfully implement the best possible scenarios. The success of the mine planning and design will be dependent on not only the software itself, but also the user's expertise with the software.

Mine design software is necessary for most mining activities because of its multi-functionality. Maptek Vulcan is a multi-functional 3D modelling and mine design software that is capable of handling geology, block-modelling, resource estimation, mine design, scheduling as well as surveying. Some of the main functions of the Maptek Vulcan software are (Vulcan, 2013):

- Advance pit, dump and ramp design;
- Pit management and optimisation;
- Able to manipulate existing underground designs; and
- Run multiple scenarios and generate mineable shapes.

Although the Maptek Vulcan mine design and layout software is capable of designing both surface and underground mines, Figure 2.5 illustrates Maptek Vulcan mine design software displaying access roads on a pit.

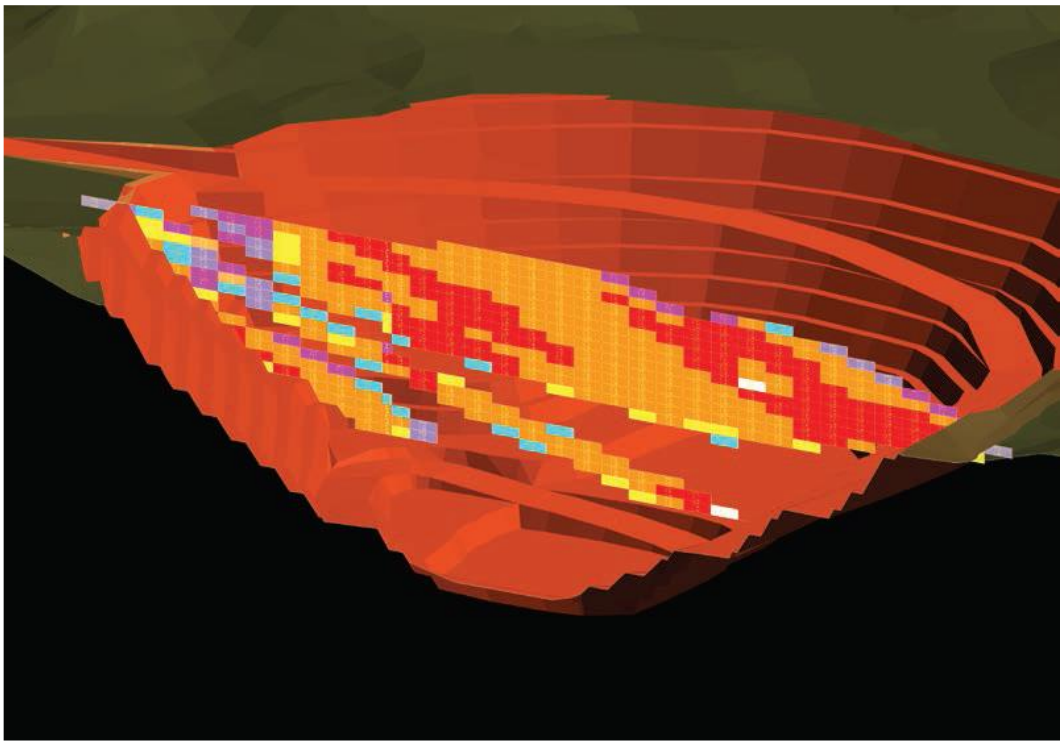


Figure 2.5: Maptek Vulcan software pit design (Vulcan, 2013)

It is also important to acknowledge other software providers in this area. The most widespread mine planning software are (Katakwa, et al., 2013):

- CADSMine;
- Minemax;
- GEOVIA Surpac;
- CAE Studio;
- MineScape;
- Maptek Vulcan;

- GEOVIA Gems;
- MineSight;
- GEOVIA Whittle;
- Mine2-4D;
- GEOVIA Minex; and
- Xpac.

2.2.5 Planning and scheduling software

Mine planning and scheduling is one of the deciding factors behind some of the key issues affecting mining operations such as the life of a mine (LOM), run of mine (ROM), etc. Depending on the financial planning of the whole operation, by using mine planning and scheduling software certain production assumptions can be made to determine cut-off grade and corresponding tonnage; hence fleet capacity as well as size of the beneficiation plant.

GEOVIA MineSched is a scheduling software designed for any kind and/or size of operation, regardless of it being an open pit or underground mine; according to the mining operations requirement, it can provide long and short term schedules. GEOVIA MineSched can also improve productivity and profitability. Figure 2.6 shows GEOVIA MineSched software creating scheduling activities.

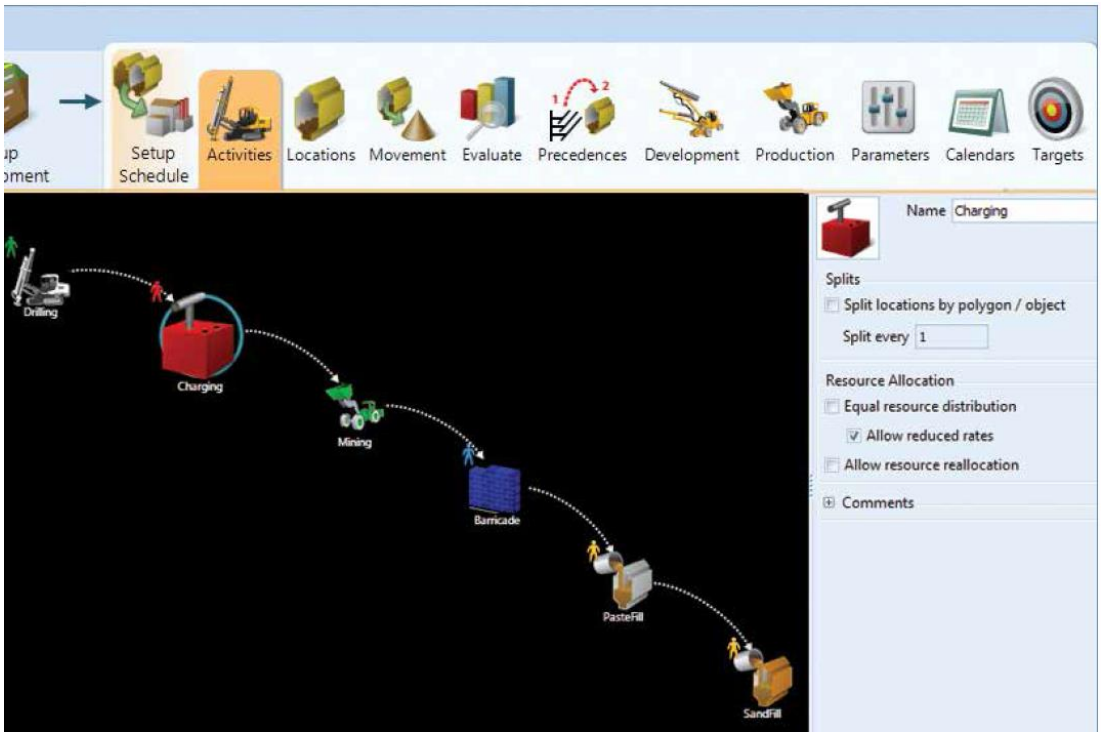


Figure 2.6: GEOVIA MineSched creating activities (GEOVIA MineSched, 2013)

2.2.6 Financial valuation software

Financial evaluation software plays a very important role to determine the profitability of planned mining operations. Financial valuation software can be used in the early stages of pre-feasibility, the calculation of LOM, ROM; financial, sensitivity, and risk analyses and high-level strategy development. It can also be used to calculate the net present value (NPV) of a mineral project.

RungePincockMinarco's XERAS strategic financial valuation solution software is designed for use in the initial stages of pre-feasibility, the life of a mine valuation and complex strategy development as discussed above. Financial valuation of mine plans and possible changes to those plans can easily be handled by XERAS financial planning. Figure 2.7 shows RungePincockMinarco's XERAS logo. Some of the features of the software are (XERAS, 2014, p. 1):

- "Forecast cash flows for multiple mining scenarios;
- Develop the most practical and valuable strategic mining approach;

- Arrive at reliable key economic indicators such as cash flow and net present value; and
- Assess the true value of the ore deposit".



Figure 2.7: RungePincockMinarco's XERAS logo (RungePincockMinarco's XERAS, 2013)

2.2.7 Mine optimisation software

One of the most important sections of mine design and planning is to improve the value of the mining operations. Whittle and Burks (2010) and Smit and Lane (2010) showed that the use of the right software solutions can help mine optimisation when used in conjunction with mine planning software such as CADSMine and Mine2-4D. Furthermore, (Dagdelen & Kawahata, 2007) developed a software called OptiPit and stated that through the use of OptiPit software, several companies derived improved value in their operations. (Trimble, 2014) also developed a software solution called Trimble Open Pit Design (TOP) to address the possibility of changing pit design by running multiple scenarios, and their solution was endorsed by Kumba Iron Ore during the SAIMM's Mine Planning School's (2014) presentation.

Choosing mine design and planning software will likely be dependent on which commodity is going to be mined as well as whether it is going to be an open-cast or underground mine. Mine planning software producers usually cater for both surface and underground operations. One of the most widespread open pit (surface mining) planning software is GEOVIA Whittle. GEOVIA Whittle is a well-known open pit mine planning software to optimise the value of the operations. On the other hand CAE Studio, formerly known as Datamine software (Datamine was bought by CAE Mining on 19 April, 2010) has a wide range of solutions that can be implemented in every stage of the mine planning operations. CAE NPV scheduler claims to be the most versatile open pit mine planning tool on the market and has users in the South African

mines. Figure 2.8 shows the CAE NPV Scheduler's mine flow optimiser module while Figure 2. 9 shows the GEOVIA Whittle core module.

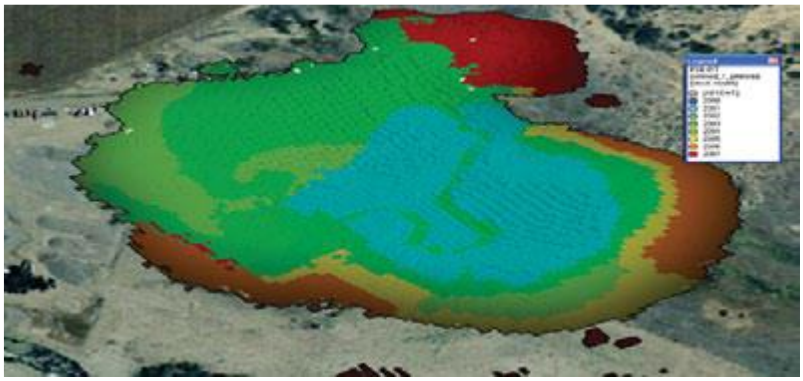


Figure 2.8: CAE NPV Scheduler – Mine flow optimiser (CAE NPV Scheduler, 2013)

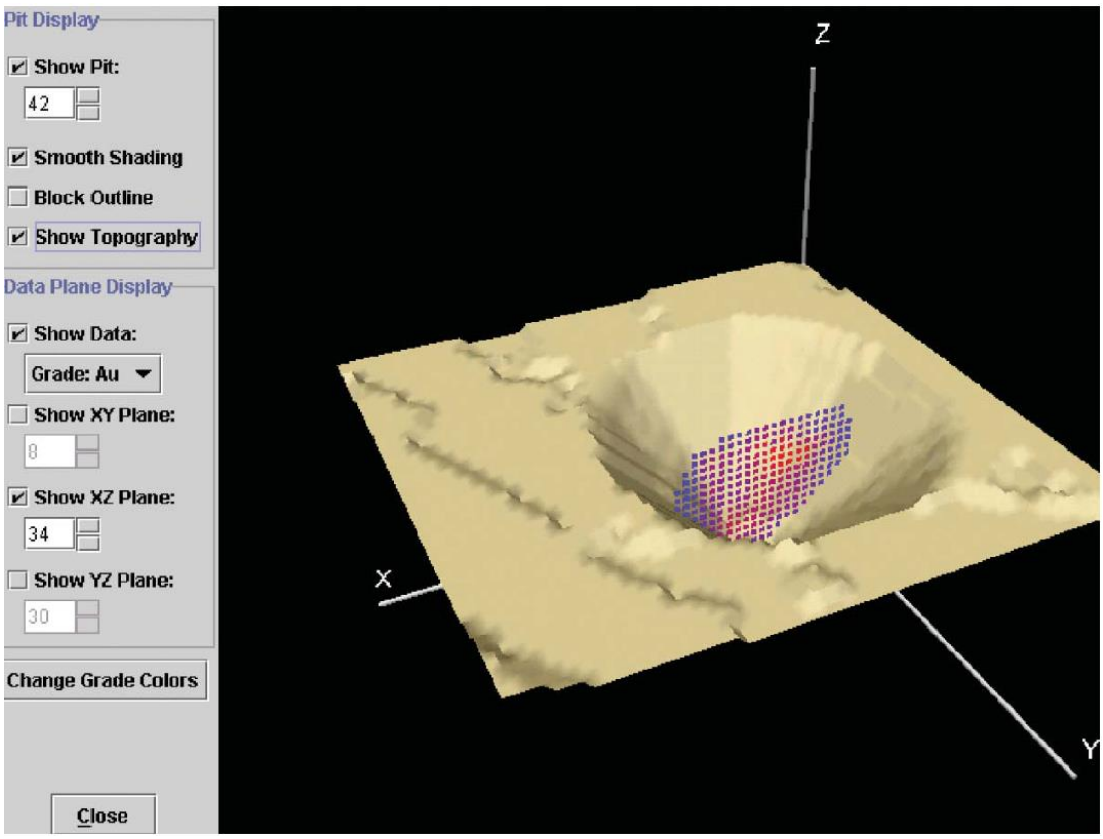


Figure 2.9: GEOVIA Whittle - Core module application (GEOVIA Whittle, 2013)

2.2.8 Other software (i.e. ERP, monitoring, reporting, etc.)

There are a number of other software solutions being used in the South African mining industry that does not fit in with the previous seven categories. Some of these software solutions are:

- Enterprise resource planning (ERP) software;
- Monitoring and reporting;
- Simulation; and
- Virtual reality (VR).

Each of these categories are discussed in detail, and one leading example of each software solution is given for each section.

2.2.8.1 ERP software

For many years mining companies were looking for better ERP solutions to manage their operations. ERP software incorporates all stages of running a business from manufacturing, distribution to financing and ERP software can be packaged in many forms such as supply chain management, and point-of-sale software (Wolf, 2013). Although an out-of-the-box solution is applicable for some mining operations, depending on the size and complexity of the operations, these solutions can differ from one company to the next. Market leaders in this field are SAP, Oracle and Microsoft Dynamics (Kimberling, 2009).

2.2.8.2 Monitoring and reporting software

Mining equipment is usually very expensive to purchase and maintain, and are among the biggest expense items in the capital expenditure cost. Similar to any other business practice, mining operations also require a fleet management solution to monitor their assets. An example is Modular Mining Systems' DISPATCH software for open-pit mines is one of the commonly known fleet management software in the mining industry. It uses a number of sensors to monitor mining operations such as (Modular's DISPATCH, 2013, p. 1):

- “Equipment allocation;
- Shift change;
- Refuelling; and
- Equipment downtime events”.

Modular Mining Systems’ DISPATCH system is not only capable of real-time positioning of the equipment by the use of a GPS system; it can also produce health and production reports. Other Dispatch software solutions in this field include Wenco’s Dispatching Solutions, Karjer Dispatch System and GPS-based Intelligent Vehicle Scheduling Management System for Open Pit Mines. Figure 2.10 shows Modular Mining Systems’ DISPATCH software monitoring equipment movement in a mine.

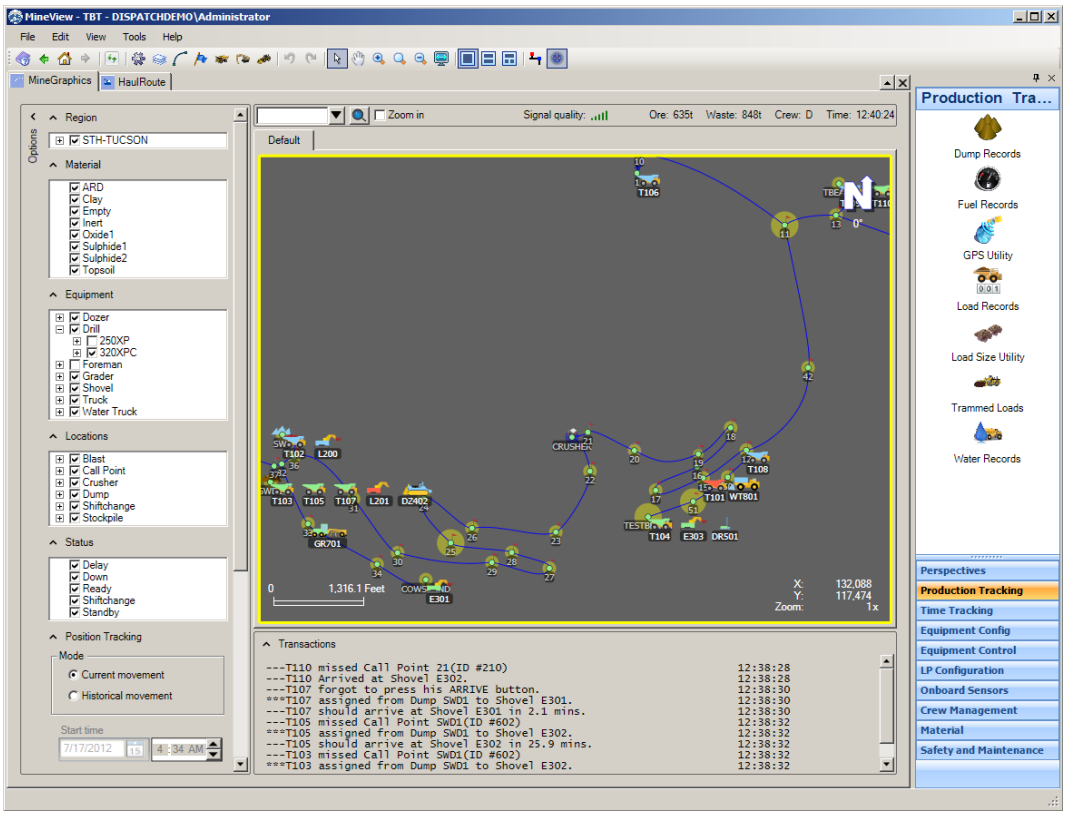


Figure 2.10: Modular Mining Systems’ DISPATCH software monitoring equipment movement (Modular’s DISPATCH, 2013)

2.2.8.3 Simulation software including simulators

As previously stated, mining equipment is one of the biggest expense in the capital expenditure cost. From this viewpoint, it is essential to train a machine operator properly long before the actual equipment is being put in use at the mine. It is now becoming industry standard to train machine operators using mining training simulators before the operator gets trained on real equipment. Most of these simulators are designed for open pit as well as underground equipment training. The expected outcomes from the simulator training include better handling of the mining equipment, improved productivity, as well as safety. Although certain mining equipment manufacturers have partnered with mining equipment training simulator providers, Immersive Technologies is perceived as the leader in simulation software. Figure 2.11 shows a 3-D model of the Immersive Technologies advance equipment simulator for coal mining. Furthermore, there are a number of computer simulation software specifically designed to address different simulation related problems. These include Arena, GPSS, Extend, and Crystal Ball.



Figure 2.11: Immersive Technologies CM360-B simulator for coal mining (Immersive Technologies, 2013)

2.2.8.4 Virtual Reality software

Virtual reality (VR) is the result of a technology that allows one to see something that is not really there – hence “virtually” real (RocScience Software, 2013). A number of initiatives are afoot to bring VR simulations in mining education. The School of Mining Engineering at the University of New South Wales has already established in their undergraduate programme to train their students in this new VR environment so that the students can use 3D visualisation and training where the underground or surface mine environments can be experienced through simulation (ViMINE, 2013).

VR is used in a wide variety of fields from work place activities to safety related training. To curb the high level of accidents in the mining industry in South Africa, as a part of a Safety in Mines Research Advisory Committee (SIMRAC) initiative, CSIR Miningtek conducted research to evaluate the possibility of using VR to train miners in the gold and platinum mining sectors in South Africa. The research concluded that VR has the capability to improve safety in mines (Squelch, 2001).

Since then the mining industry in South Africa started using VR training widely in more sectors such as the coal and diamond sectors. The idea behind implementing VR supported training mechanism are not to replace the current training methodologies which are in use, but aims to complement existing systems to improve health and safety as well as productivity of the mine workers. Figure 2.12 shows ViMINE VR software illustrating a 3D orebody where the user can virtually position the orebody using on-screen tools such as depth and dip of the orebody.

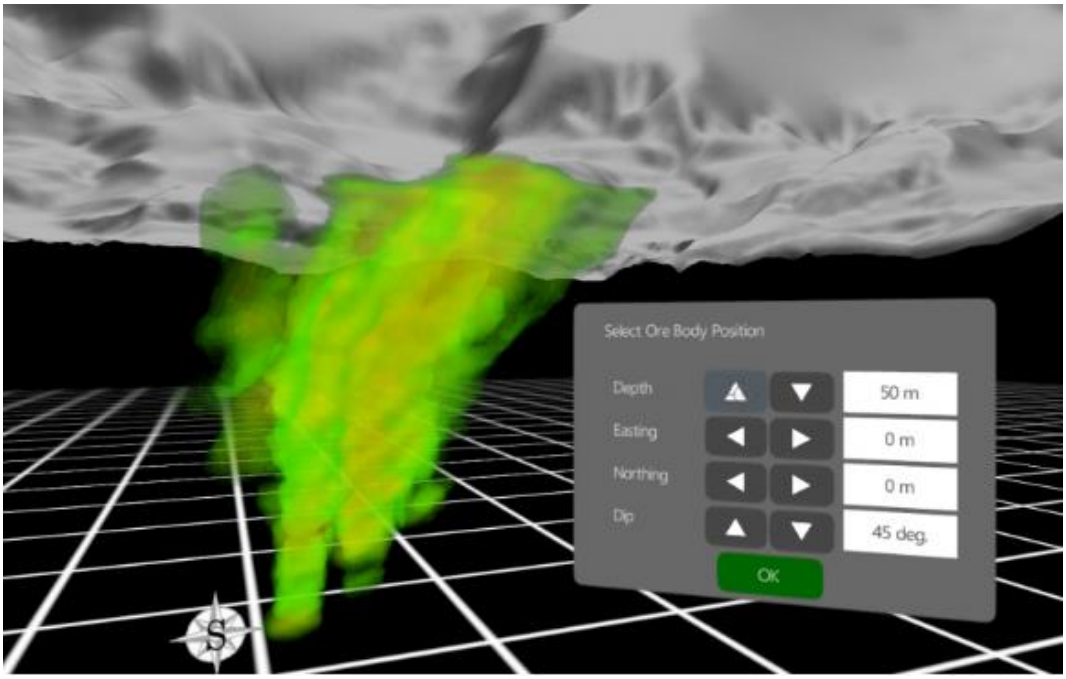


Figure 2.12: ViMINE VR software virtually showing 3D orebody (ViMINE, 2013)

2.3 Summary

This chapter listed and discussed some of the leading mine planning software that is being used in South Africa. Mine planning software is divided into eight different categories according to its functionalities, namely; Mapping, Geological data management, Geological modelling and resource estimation, Design and layout, Scheduling, Financial valuation, Optimisation, and Other (Monitoring, ERP, etc.) and each of these categories are discussed in detail. Chapter 3 looks at the first established online database at http://db.mining.wits.ac.za and discusses possible software utilisation strategies.

3 ONLINE SOFTWARE DATABASE

3.1 Introduction

As mine planning software plays a vital role in ensuring successful mining operations, it is also important to understand the level of the mine planning software utilisation in South African mines. Katakwa et al. (2013) initiated a new approach to determine the level of the mine planning software utilisation in South Africa by establishing an online database called the “Mine Planning and Peripheral Software Database”, and this web-based database is accessible to everyone at http://db.mining.wits.ac.za. This chapter outlines the structure and functionalities of this online database and discusses Katakwa et al.’s (2013) approach to determine the level of the mine planning software utilisation in South Africa.

3.2 Rationale for collating mine planning software

As there are several mine planning software available in the market, it is also important to find the current level of the mine planning software utilisation in South African mines. It is a well-known fact that most mine planning software users prefer using a combination of software to complete their task efficiently. In many cases the available software is capable of doing what is needed to be done but some users prefer using specific software for a certain task. For example while Mine2-4D software is able to do scheduling, according to Smit and Lane (2010) some users in the platinum sector prefer using AMPOT software for the same purpose. Using a combination of software to complete the required task will affect the level of utilisation on each of the software. Furthermore, it is also important to know the market impact of each particular software that is being used in South Africa so that decision-makers can make informed decisions regarding leading mine planning software solutions. By using a similar approach, other stakeholders can also benefit from this initiative. Other stakeholders include educational institutions, mining consulting firms and lastly software providers.

The School of Mining Engineering at the University of Witwatersrand established a research group called “The Mine Planning, Optimisation and Valuation Research Group” in 2010. Through this research group, an MSc research project started early

in 2012 to analyse the level of mine planning software utilisation in South Africa. This MSc research outcome was converted to an online database at http://db.mining.wits.ac.za as a beta version release (Katakwa, et al., 2013). This chapter presents the concept behind the database.

3.3 Mine planning and peripheral software database

The initial beta version of the database was completed towards the end of year 2012 as a result of an MSc research dissertation and later on a simplified version of the results were published by Katakwa et al. (2013). Since then the database is available online, it is freely accessible and compatible with any web browser.

To serve the online database, first of all, a server database engine was required. The world's most well-known open source database engine, MySQL, is not only freely available, but is also platform independent. Because of these great advantages, MySQL database engine was used for this task.

Secondly, to display the data in a web browser such as “Internet Explorer” a server side scripting language was needed. Server side scripting can be defined as a technology to generate dynamic web pages (Chaney, 2014). By using this technology, an interactive web site interface to the current database could be established. The main advantage to server side scripting is that it can be easily customised as per specific requirement. From a security point of view, as the web browser user cannot see the scripts that run on the server, this makes the database server more secure. Some server side scripting languages include (Chaney, 2014, p. 1):

- “ASP;
- ANSI C scripts;
- ColdFusion Markup Language (*.cfm);
- Java via JavaServer Pages (*.jsp);
- JavaScript using Server-side JavaScript (*.ssjs);
- PHP (*.php);
- SMX (*.smx);
- Lasso (*.lasso);

- WebDNA (*.dna,*.tpl); and
- Progress® WebSpeed® (*.r,*.w)".

From the list of available server side scripting languages, PHP was selected. Rakesh (2014, p. 1) listed some of the important features of PHP as:

- "Stable and secure – It is capable of developing powerful web applications that are scalable up to a very large number of users. PHP is stable and secure, robust enough for business-critical applications needing constant up-time and high security;
- Rapid and easy deployment – It is easy to understand and learn. In addition there are development frameworks such as QCubed which separate business logic from the user interface resulting in cleaner and reusable code;
- Cost – It is free;
- Portable – It is available for every operating system;
- Database integration – It includes the ability to connect directly to a MySQL database which removes the additional overhead when using Open Database Connectivity; and
- Support – It is supported by a vibrant community, estimated to consist of about 5 million developers that share code, evangelise about PHP, support each other to improve the language and increase its capabilities".

Similar to the MySQL, PHP is also free, very stable and secure, and more importantly it can connect directly to a MySQL database. Figure 3.1 is a diagrammatic representation of how the server-side scripting produces dynamic web pages.

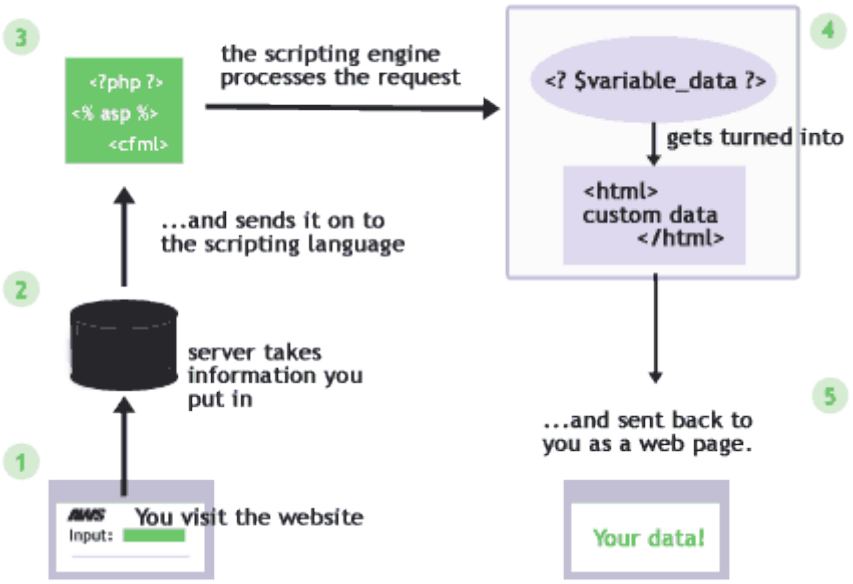


Figure 3.1: How the server-side scripting produces dynamic web pages (PHP and MySQL) (Bogotobogo, 2014)

3.4 Database structure

During the planning stage, the database was structured in such a way that users can easily browse the database, search the contents of the database and finally analyse the available information within the database. The database was categorised into six different sections, namely:

- Company profiles;
- Software functionalities;
- Mineral commodity types;
- Site types;
- Mine value chain stages; and
- Specialities.

Each of these categories has their own sub-sections to make browsing easier. For example the sub-section of the company profile is (Katakwa, et al., 2013, p. 499):

- “Specialist software provider;
- Mining company;
- Mineral processing company;

- Educational institution;
- Consulting company;
- Mineral Exploration Company; and
- Mining Equipment Provider".

Table 3.1 shows how the database was categorised and divided into sub-sections. Data in the database were obtained from three main sources, namely:

- Software vendors;
- Department of Mineral Resources; and
- Consulting firms.

In the initial release version (beta version) of the database, the time-stamp for the information about mining companies (operating mines, quarries and processing plants) which was obtained from the Department of Mineral Resources (DMR) website (http://www.dmr.gov.za) was the year 2010. Software vendors' information was up to date, with a September 2012 time-stamp.

Table 3.1: Categorisation of data (Katakwa et al., 2013)

Data	Categorisation	Data	Categorisation
Company profiles	• Specialist software provider • Mining company • Mineral processing company • Educational institution • Consulting company • Mineral Exploration Company • Mining Equipment Provider	Site types	• Mine • Quarry • Plant • Office • Agency • Campus
Software functionalities	• Mapping • Geological data management • Geological modelling and resource estimation • Design and layout • Scheduling • Financial Evaluation • Optimisation • Other (Monitoring, ERP, etc.)	Mine value chain stages	• Discover • Establish • Exploit • Beneficiate • Sell • Rehabilitate
Mineral commodity types	• Precious metals and minerals • Energy commodities • Nonferrous metals • Ferrous metals • Industrial minerals • Aggregate and sand • Processed minerals	Specialities	• Mining engineering • Geology • Survey • Rock engineering • Ventilation • Mineral processing • Finance • Human Resources • Environment

While the database was categorised and divided into sub-sections is important to represent data, it is also equally important to illustrate how categories and sub-sections within the database are linked. By doing so, one can easily follow the structure of the database. Figure 3.2 shows how the data is presented within the mine planning and related software database.

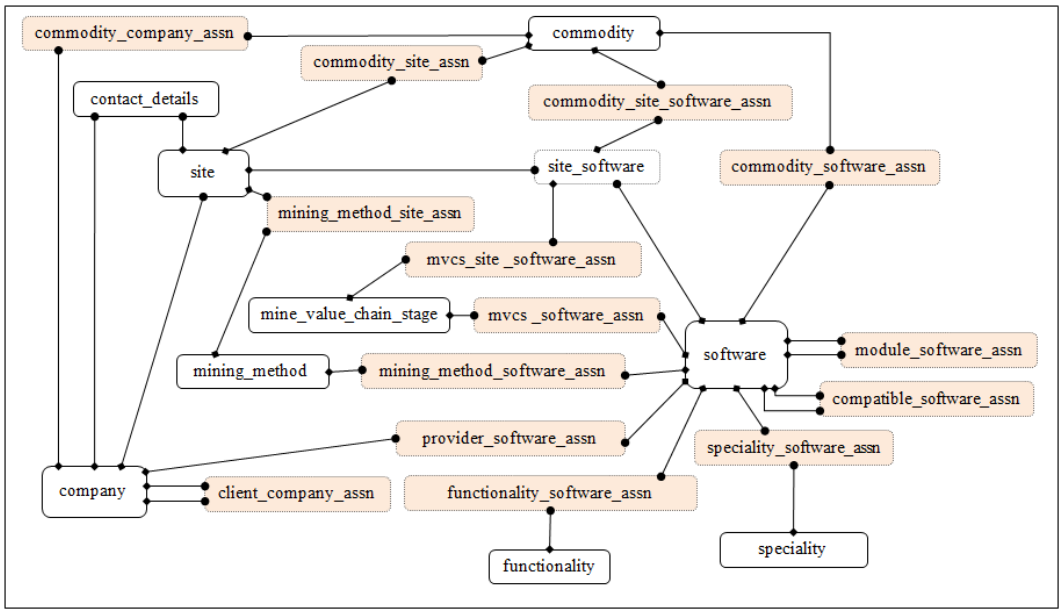


Figure 3.2: Summary of data representation within the mine planning software database (Katakwa, et al., 2013)

Katakwa et al., (2013) further developed a user-friendly web-based interface for accessing and analysing data within the database. The web-based interface was developed in such a way that pages were made as uniform as possible for a better and easier user interaction. Figure 3.3 shows how the website is structured to display the mine planning and peripheral software database.

QCubed, which is known as a PHP framework, was used to develop the website. To help developers, QCubed developed a number of frameworks to save the time for routine, repeatable tasks. Highcharts interactive JavaScript charts for the web-projects, is a charting library to create interactive effects within web browsers called JavaScript. Highcharts supports different chart types and Highcharts library were used to produce the graphs for displaying a number of analyses. Figure 3.4 shows the homepage of the mine planning and peripheral software database (Katakwa et al., 2013).

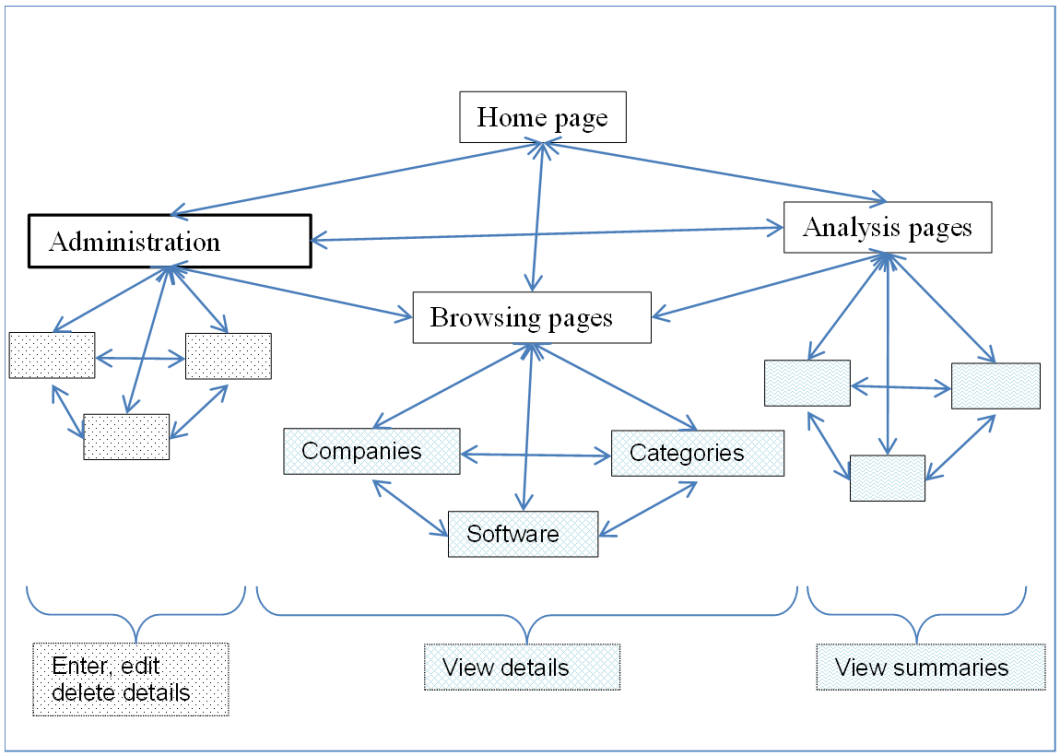


Figure 3.3: Website structure for the web-based interface of the database (Katakwa, et al., 2013)

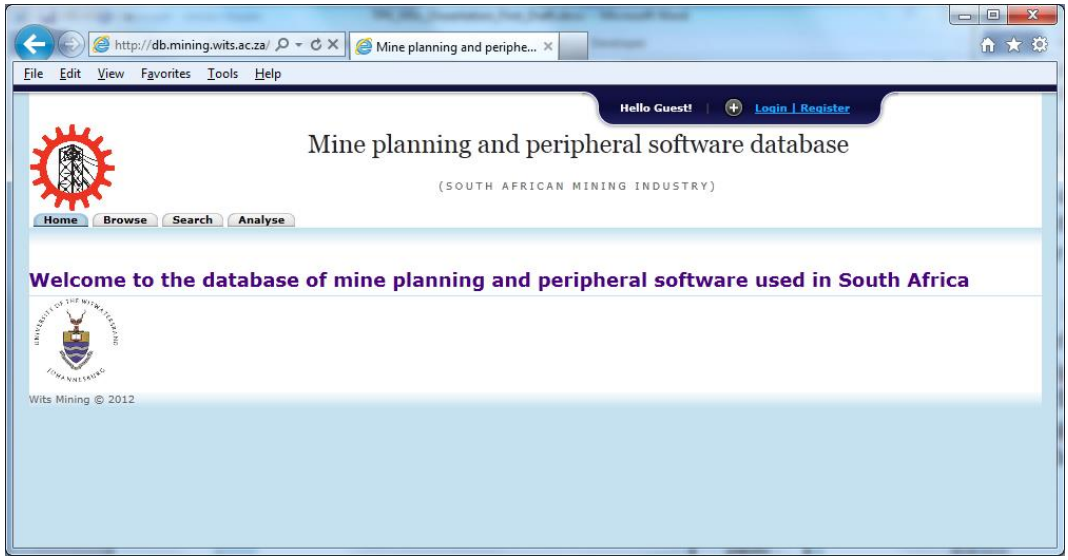


Figure 3.4: Homepage of the database (Katakwa et al., 2013)

3.5 Database features

Mine planning and peripheral software database is compatible with any web-browser and does not depend on the operating system. The user interface is easy to understand. Katakwa et al. (2013, p. 500) listed the functionalities of the database as:

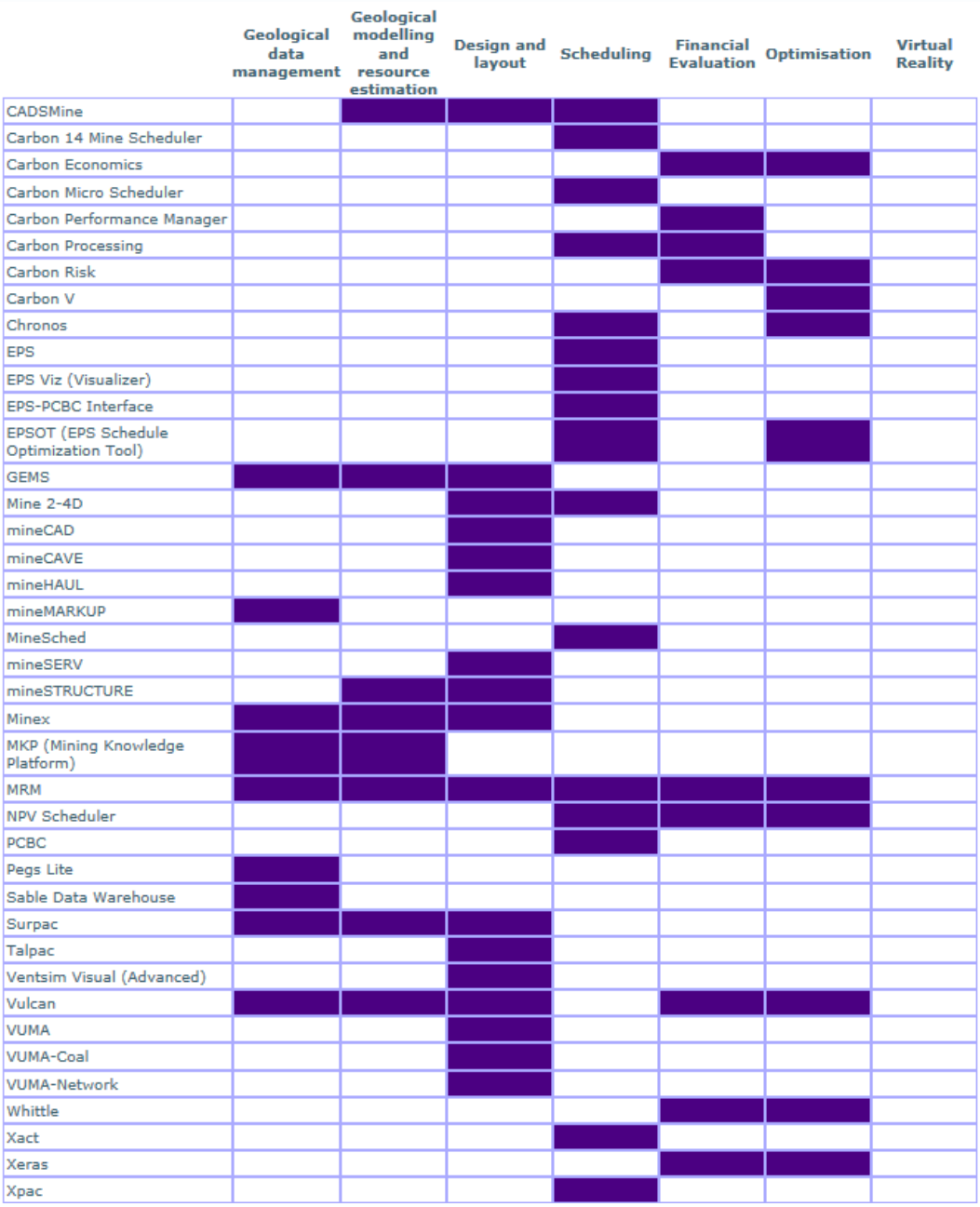
- "Data storage – The system stores information about companies, related sites and software installed at those sites. The software is categorised with respect to its functionality and related mineral commodities and mining methods. Companies are related as to whether they are developers, suppliers or clients of software;
- Data access – The system provides a main page which anyone surfing the web will be able to access. This page allows any user to browse or search for basic information on companies and software;
- Summary information – The system has a menu providing options such as Analyse market impact, Analyse software adoption, Analyse client distribution and Compare solutions available for certain mine value chain stages;
- Access levels – The system allows users to have different levels of privileges. A 'login' facility is provided to allow privileged users access to more advanced functionalities of the system. While all users can access all the summary information, only the privileged users are able to view certain confidential data and to enter or update specific data".

Figure 3.5 shows the main features of the web-based user interface. Using a navigation panel, different functionalities of the database can be used.



Figure 3.5: Main features of navigation and interface design for the web-based interface of the database (Katakwa, et al., 2013)

On the other hand, it is also possible that the database can easily display all the software that was registered within the system by a thorough search functionality. Figure 3.6 shows all the available software in the database with their functionality. As can be seen from Figure 3.6, there is no software associated with the virtual reality software (VR) because by the time the database was put online, efforts to put a VR lab in action were not complete in South Africa.



	Geological data management	Geological modelling and resource estimation	Design and layout	Scheduling	Financial Evaluation	Optimisation	Virtual Reality
CADSMine		■	■	■			
Carbon 14 Mine Scheduler				■			
Carbon Economics					■	■	
Carbon Micro Scheduler				■			
Carbon Performance Manager					■		
Carbon Processing				■	■		
Carbon Risk					■	■	
Carbon V						■	
Chronos				■		■	
EPS				■			
EPS Viz (Visualizer)				■			
EPS-PCBC Interface				■			
EPSOT (EPS Schedule Optimization Tool)				■		■	
GEMS	■	■	■				
Mine 2-4D			■	■			
mineCAD			■				
mineCAVE			■				
mineHAUL			■				
mineMARKUP	■						
MineSched				■			
mineSERV			■				
mineSTRUCTURE		■	■				
Minex	■	■	■				
MKP (Mining Knowledge Platform)	■	■					
MRM	■	■	■	■	■	■	
NPV Scheduler				■	■	■	
PCBC				■			
Pegs Lite	■						
Sable Data Warehouse	■						
Surpac	■	■	■				
Talpac			■				
Ventsim Visual (Advanced)			■				
Vulcan	■	■	■		■	■	
VUMA			■				
VUMA-Coal			■				
VUMA-Network			■				
Whittle					■	■	
Xact				■			
Xeras					■	■	
Xpac				■			

Figure 3.6: Available software in the database with their functionality (Katakwa, et al., 2013)

Some of the other charts can be reached by using the “Analysis” tab, where the user can request the following (Katakwa, et al., 2013, p. 502 & 503):

- “Distribution of commodities at mining sites;
- Adoption of software at mining sites;
- Relative market impact based on the criteria selected; and
- Distribution of clients by company profile for the different providers”.

Figure 3.7 to Figure 3.11 show some of the output based on the criteria selected under the ‘Analysis’ tab.

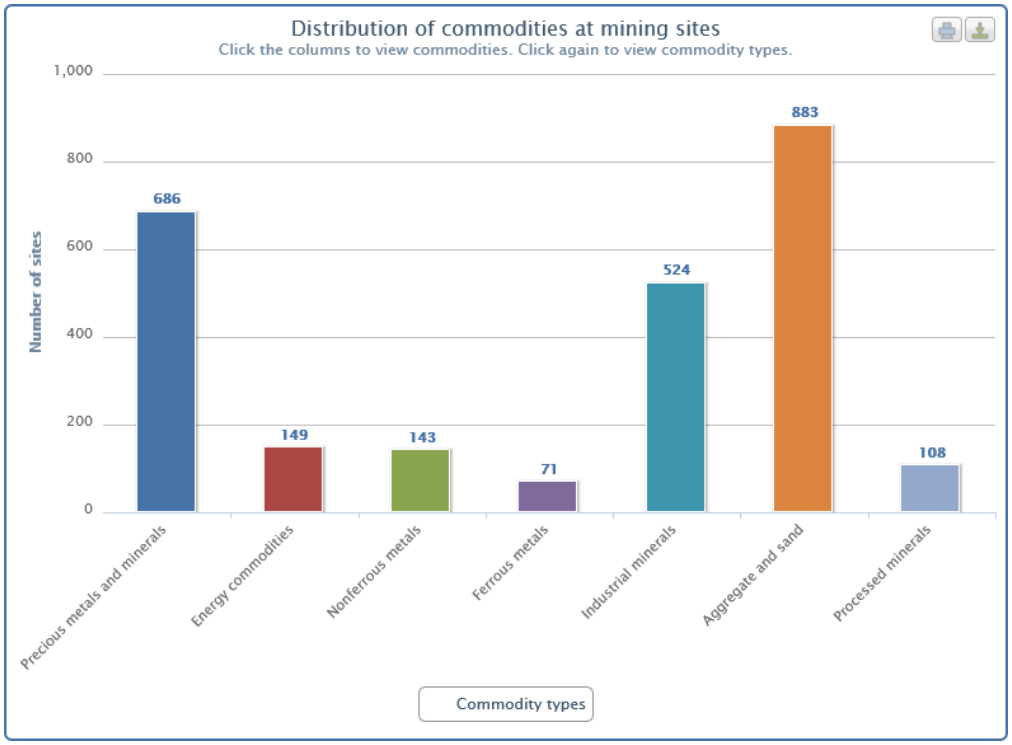


Figure 3.7: Distribution of commodities at mining sites

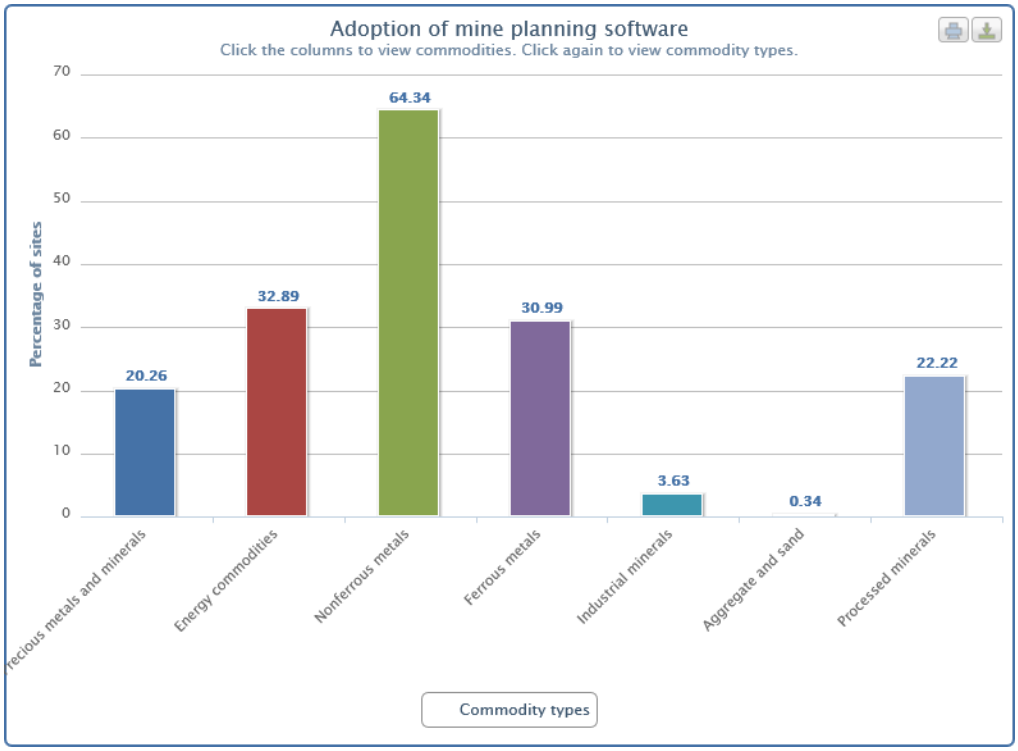


Figure 3.8: Adoption of mine planning software

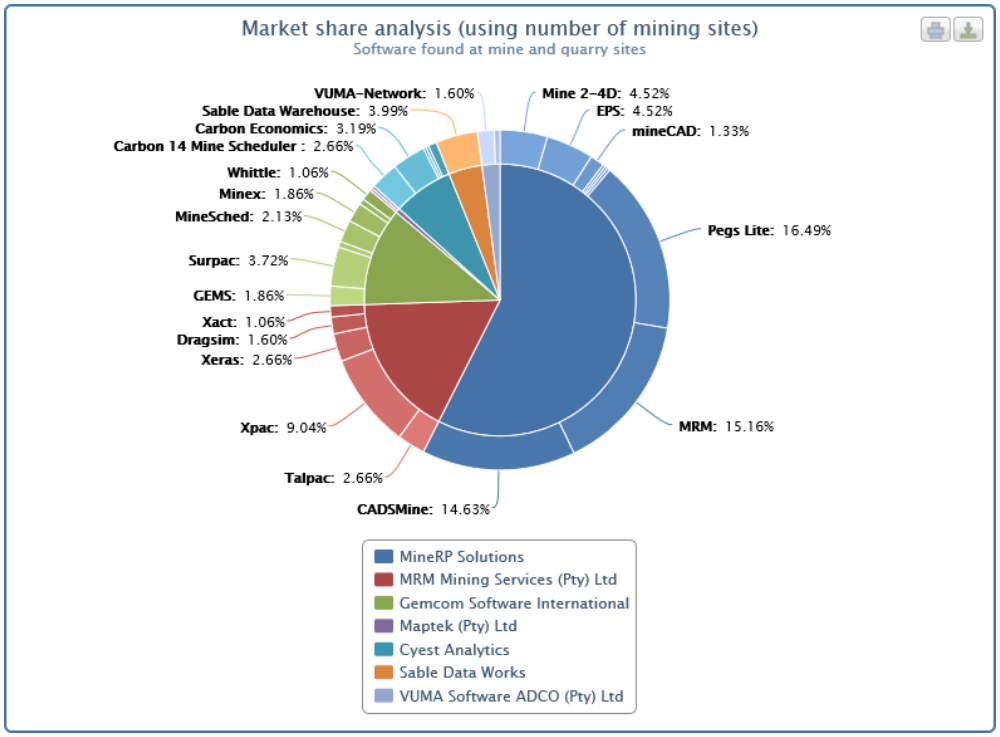


Figure 3.9: Market impact analysis (using number of mining sites)

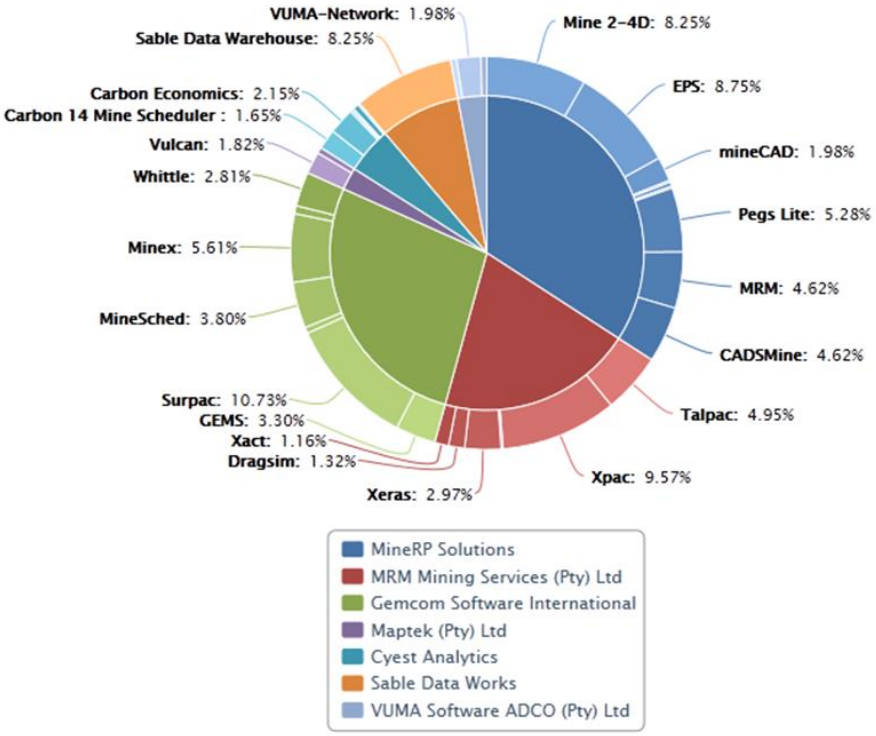


Figure 3.10: Relative market impact for software

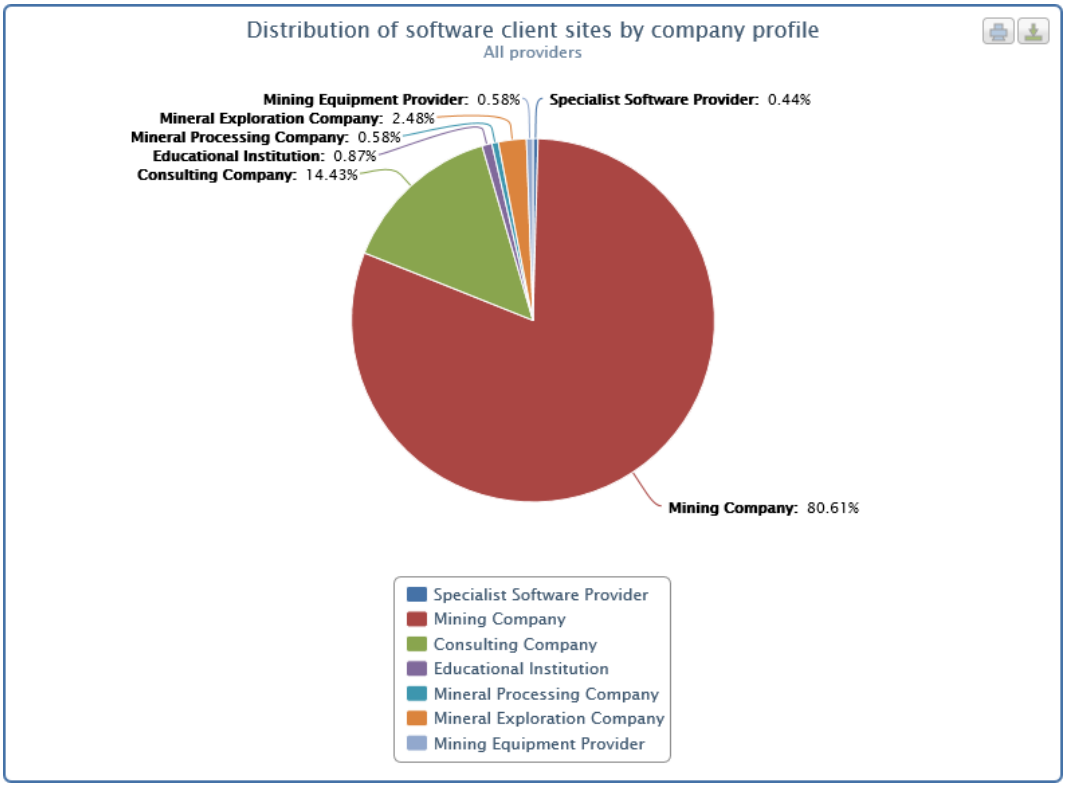


Figure 3.11: Distribution of software client sites by company profile

3.6 Discussion of the results

Figure 3.7 to Figure 3.11 show how the database illustrates different aspects of software usage by the mining industry in South Africa. The database clearly indicates the participating software companies' market impact, which mining sector uses software most, and the software usage share of consulting companies, as well as educational institutions. It needs to be noted that the time-stamp on the database was September, 2012.

There was a data gap within the database as a major software provider, CAE mine planning software solution had not yet provided the required information. CAE Mining data was made available in 2014.

One of the main reasons behind the establishment of the database was to indicate the extent of mine planning software utilisation within the mining industry in South Africa. The online database was the prototype of its kind as there was no similar research done anywhere else. Although the current database shows the market

impact analysis by the software vendors, lack of participation from some main software vendors in the study, as well as the limited functionality of the database (i.e. unable to produce proven techniques to predict future use of mine planning software) caused the online database to achieve limited success.

This limited success created a need to bring a new approach to the question of evaluation of mine planning software utilisation for informed decision-making strategies in the South African mining industry. In this approach, using the online database as a baseline, first CAE Mining was added to the database and then existing software providers' datasets updated in 2014. Furthermore, by considering the future development of the database as it is an evolving database, a framework, an index as defined in Section 4.3 was developed to measure the level of software utilisation within the mining industry by establishing the relationship between commodity type, the functionality software (six different stages of mine planning namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation) and the time factor.

Consequently the database was prepared to perform a time-series analysis (mine planning software market trend analysis) related research (which demonstrated how the utilisation of the mine planning software changed in time). This was achieved by using machine learning, statistical pattern recognition and data mining techniques which are explained further in Chapter 5.

3.7 Summary

While mine planning software in the South African mining industry continues to be more widespread, to address the level of software utilisation within the industry, an online database was established but the database had limited success because of lack of participation from some main software vendors in the study, as well as the limited software utilisation evaluation functionality provided by the database.

Chapter 4 answers the problems mentioned in this chapter and outlines the development of a new methodology to define and measure mine planning software utilisation in the South African mining industry. Using the updated dataset, a

methodology was developed and implemented to measure mine planning software utilisation to enhance decision-making strategies in the South African mining industry.

4 DEFINING AND DERIVING MINE PLANNING SOFTWARE UTILISATION

4.1 Introduction

As mentioned in Chapter 3, an initial dataset showing the mine planning software providers, their corresponding software solutions as well as the software capabilities and information on the number of licences were collected and compiled in 2012 in an online database. The database development and implementation was published in the *Journal of the Southern African Institute of Mining and Metallurgy* in 2013 (Katakwa, et al., 2013). In 2014 the dataset was updated with additional and new information. Using the updated dataset, a methodology was developed to measure mine planning software utilisation to provide information for enhancing decision-making strategies in South Africa. However, before discussing this methodology, it is essential to define the meaning of software utilisation.

Utilisation is a well-known word within the mining industry because of its ties with the level of productivity. Higher utilisation often leads to higher productivity, hence better profit margins. From this point of view, utilisation is an important factor regardless of the size of any operation including those in the mining sector. The root of the word of utilisation comes from the word “utilise”. “Utilise” means “make practical and effective use of” (Oxford English Dictionaries, 2014). By using this definition, software utilisation can be defined as the effective use of mine planning software in South Africa but in general, utilisation is associated with the overall equipment effectiveness which is one of the key performance-based metrics. It is important to understand the fundamentals behind these metrics.

4.2 Overall equipment effectiveness

In literature, utilisation is associated with time in a way that it can be defined as the measurement of time used to perform effective work. In the mining industry, regardless of whether it is a surface or underground operation, better equipment utilisation often leads to higher levels of productivity, hence profitability.

Although there are a number of ways to measure performance against the various metrics, the most widely used measure to determine performance against capability

of the equipment is Overall Equipment Effectiveness (OEE). OEE measurement is also commonly used as a key performance indicator (KPI) in Total Productive Maintenance (TPM) and Lean Manufacturing programs for measuring production efficiency (Vorne Industries, 2008).

There are six factors also known as the Six Big Losses, which are the main causes of production losses. TPM and OEE programs aim to control these six factors. Shirose (2013, p. 1) listed these six factors effecting equipment utilisation as:

- "Breakdown Loss;
- Setup and Adjustment Loss;
- Idling and Minor Stoppage;
- Reduced Speed Loss;
- Quality Defects and Rework, and
- Start-up Loss".

In the TPM model, Nakajima (1998) furthermore formulated utilisation using availability, performance rate and quality rate as shown in the following Eq. 4.1:

$$\textit{Equipment effectiveness} = \textit{Availability} \times \textit{Performance rate} \times \textit{Quality rate} \quad (4.1)$$

In this formula, equipment effectiveness defines the meaning of equipment utilisation and is calculated by multiplying equipment availability by performance rate and quality rate. Figure 4.1 shows time factors effecting equipment utilisation in a loading activity. In Figure 4.1, operation time is associated with total available time for a given period as this can be anything from a shift of the day, to a whole month. As shown in Figure 4.1, loading time can be calculated by deducting downtime from the operation time.

Availability can be calculated by dividing loading time by operation time. As the loading time calculation is already given, the availability equation then is (Shirose, 2013):

$$Availability = \frac{Operation\ time - Planned\ \ downtime - Unplanned\ \ downtime}{Operation\ time} \quad (4.2)$$

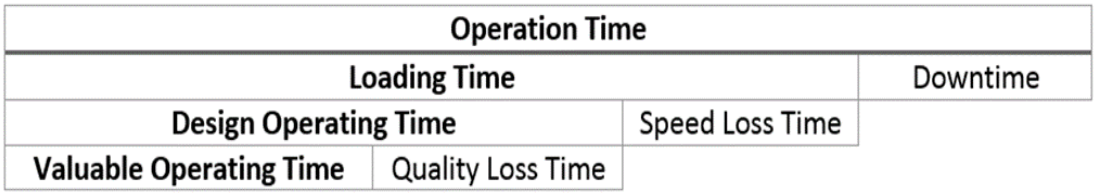


Figure 4.1: Time factors effecting equipment utilisation (adapted from Shirose, 2013)

Furthermore, speed loss time is the lost time caused by operating below the planned speed and can be calculated by using the actual time to make the production quantity minus the design time to make the same quantity as shown in the Eq. 4.3 (Shirose, 2013):

Speed loss time = Parts produced × (Design cycle time per part − Actual cycle time per part) (4.3)

Cycle time is the time taken to produce one part. Design cycle time is used to calculate the equipment's designed production rate, and actual cycle time is used to calculate the equipment's actual production rate. Design operating time is the time the equipment should have taken to produce the parts and is the difference between the loading time and the speed loss time. Performance rate is the ratio of the design operating time to loading time as shown in the Eq. 4.4 (Shirose, 2013):

$$Performance\ rate = \frac{Design\ operating\ time}{Loading\ time} = \frac{Loading\ time - Speed\ lost\ time}{Loading\ time} \quad (4.4)$$

Quality loss time is the time lost making nonconforming material. Valuable operating time is the time the equipment spends making conforming material. Quality rate is the ratio of conforming parts produced to total parts produced (Shirose, 2013) as shown in the Eq. 4.5:

Quality loss time = Nonconforming parts × Actual cycle time (4.5)

Although OEE is a very powerful tool to measure efficiency, hence utilisation; it is fundamentally designed for equipment utilisation which can be defined as *hardware utilisation*. The aim of this study is to define strategic software utilisation in the South African mining industry which can be defined as *software utilisation*, although OEE

gives some ideas regarding utilisation, it is not designed to establish a framework which can bring a new approach towards mine planning software utilisation.

4.3 Software utilisation

El-Ramly and Stroulia (2004) tried to explain software utilisation. They stated that there are a number of techniques available to understand how often the software is being used as well as in which degree it is being used. Many software systems collect or can be instrumented to collect data about how users use them, i.e., system-user interaction data. Such data can be of great value for program understanding and re-engineering purposes. Sequential data mining methods can be applied to identify patterns of user activities from system-user interaction traces (El-Ramly & Stroulia, 2004).

Despite the fact that user data may be available in some instances, using the data mining methods based on the user behaviour, to measure mine planning software utilisation is inappropriate when considering the size of the whole South African mining industry and user privacy. By selecting the number of targeted mining sites, limited research output could be possible but most probably would not be sufficient to satisfy the whole mining industry in South Africa.

To achieve a successful research initiative which covers the whole South African mining industry, a methodology was developed in such a way that utilisation of the various mine planning software could be measured. This measurement framework is defined below.

By using an analogy to the one given earlier, software utilisation can also be defined by associating many-to-many, one-to-many and many-to-one relationships between entity types. In this association, the relationship between software vendors, commodity, functionality and time factor were used to develop the following terminology:

$$\{C_i, F_l\} \rightarrow S_{k=\{i,\, l\}} \tag{4.6}$$

where "C_i" denotes commodity (*i)* and "F_l" denotes functionality (*l)*.

Furthermore, S_k is the software that performs tasks on commodity (*i*) and functionality (*l*). In the market usually there is more than one software specifically designed for commodity (*i*) and functionality (*l*). In order to identify and evaluate each particular software, a new index (*m*) is used so that $S_k^{(m)}$ is defined to represent a unique software whereas *k={i,l}* is an index which is a specific combination of *{i,l},* and:

$$m=1, 2, 3\ldots M$$

where *M* is the total number of software. For example, assuming there are three software companies namely X, Y and Z. Each of these three companies might have a number of software solutions, i.e. software company X has three types of software namely; X_1, X_2 and X_3, and company Y has only one type, namely Y_1, and company Z has two types of software, namely Z_1 and Z_2. Table 4.1 displays how to find *M.*

Table 4.1: Number of software solutions by company

Name of Software Company	Company X			Company Y	Company Z	
Name of Software Solution	X_1	X_2	X_3	Y_1	Z_1	Z_2
m	1	2	3	4	5	6

From Table 4.1, the total number of available software, *M* is 6.

Using a similar approach, the utilisation of the software can also be defined as such:

$u_{i,l}^{(m)}$ which is the utilisation of the software that performs a task on commodity (*i*) and functionality (*l*) by using software (*m*). Although there is no rigid definition of the software utilisation, it is defined as a numeric value which falls into the range between 0 and 1 inclusive, i.e.

$$u_{i,l}^{(m)} \in [0,1]$$

By doing so, performing further analytic development on the software utilisation can be accomplished. Furthermore, the utilisation formula can be extended by considering the time factor (*t*) to obtain the following equation:

$$u_{i,l}^{(m,t)} = f_{i,l}^{(m,t)} . w_{i,l}^{(m,t)} \quad (4.7)$$

where $f_{i,l}^{(m,t)}$ is a quantity factor that relates to the software that performs a specific task on commodity (*i*) and functionality (*l*) using software (*m*) at a specific time (*t*) and $w_{i,l}^{(m,t)}$ is the weighting factor, which will handle the missing data related issues and/or other factors such as market capitalisation of the companies. For instance $f_{i,l}^{(m,t)}$ can be defined as the total number of sites. For example, if the market capitalisation of the software companies X and Y are USD 1 million and USD 100 million respectively, but if both companies have a software solution having the same functionality, then the weighting factor of the small company will be higher than the other software company. Furthermore, the price of the mine planning software as well as support availability play an important role when considering the weighting factor.

As the software utilisation is already defined in a generic way previously, however, the software utilisation can also be defined in a specific way, i.e. the relative utilisation (*r*). Relative utilisation can be considered as a weighted software utilisation and can be formulated as:

$$r_{i,l}^{(m,t)} = \frac{u_{i,l}^{(m,t)}}{\sum_{n=1}^{M} u_{i,l}^{(n,t)}} \quad (4.8)$$

where $\sum_{n=1}^{M} u_{i,l}^{(n,t)}$ is total utilisation of all software which is used for normalisation.

Calculating relative utilisation leads to weighted market impact of the software utilisation. However, calculating relative utilisation, three variables were used to generate the results, namely:

- commodity (*i*);
- functionality (*l*); and
- time factor (*t*).

Although the calculation could be done on any commodity, the results presented in this chapter for two commodities (*i*) only. The choice of the two commodities was premised on them being among the most significant commodities in the South African

mining industry and to also provide perspectives on both soft-rock and hard-rock mining environment. In addition, analysis for a single commodity produces 12 datasets as explained later on in this chapter, Chapter 4. These commodities are:

- Coal
- Platinum Group Metals

In Chapter 2, under the Section 2.2, mine planning software was categorised under eight main functionalities. However, these eight functionalities were reorganised to reflect the capabilities of the database and relisted under six main functionalities; this was done by merging the "*Mapping"* functionality (Section 2.2.1) into the *Geological data management* functionality, and omitting "*Other (Monitoring, ERP, etc.)"* functionalities due to a lack of available software in these categories.

In addition, the calculation used six different functionalities (*I*) Katakwa et al. (2013, p. 499) namely:

1. "Geological data management;
2. Geological modelling and resource estimation;
3. Design and layout;
4. Scheduling;
5. Financial valuation; and
6. Optimisation".

Six functionalities listed by Katakwa et al. (2013) originated from the Open Group's Business Reference Model which categorises not only the functionalities of mine planning software, but also mine value chain stages and mining methods (The Open Group, 2010). The Open Group's Business Reference Model illustrates how the various software solutions interact with each other. Talpac software may seem to be an anomaly. However, its functionality as captured in the database was provided by the software supplier. In essence, this software is used for equipment selection. However, equipment selection is part of the mine design process. For example, size of excavators that are designed must also consider equipment to be used. In reality, some of the functionalities overlap/integrated functionalities. In fact, all the Design and Layout and Scheduling software are optimisation software, too. However, as can be

seen in Figure 4.2, only three software solutions are specifically designed for “Optimisation” functionality. Figure 4.2 shows the names of available mine planning software and their functionalities along the mining value chain.

The time (*t*) factor has two time-stamp indicators showing different data collection dates namely:

- September 2012, t=1
- April 2014, t=2

It is hoped the database will again be updated in 2016 in order to keep in line with an approximately two-year regular database update frequency. The two-year frequency is to allow for sufficient time for changes to have occurred in the South African mine planning software space, such as entry of new software providers and mergers and acquisitions of existing software providers.

	Software_ID	Geological Data Management	Geological Modelling and Resource Estimation	Design and Layout	Scheduling	Financial Valuation	Optimisation
Bentley Evaluation	109				■	■	
Bentley Scheduler	108				■		
BLOCK AGG	113	■			■		
CADSMine	99		■	■	■		
Carbon 14 Mine Scheduler	76				■		
Carbon Economics	77					■	■
Carbon Micro Scheduler	81				■		
Carbon Performance Manager	78					■	
Carbon Processing	80				■	■	
Carbon Risk	79					■	■
Carbon V	82						■
Chronos	69				■		
Dragsim	74				■		■
Enhanced Production Scheduler (CAE)	87				■		■
EPS (MineRP)	4				■		
EPS Viz (Visualizer)	20				■		
EPS-PCBC Interface	33				■		
EPSOT (EPS Schedule Optimization Tool)	116				■		■
GEMS	9	■	■	■			
Geological Data Management Solution	95	■					
HAULNET	114				■		
Interactive Short Term Scheduler	86			■	■		
LoM Economics	103					■	■
Maptek I-Site	100	■					
Maxipit	92					■	■
Mine 2-4D	2			■	■		
Mine Scenario Planning	101			■	■		
Mineable Layout Optimizer	90			■			
Mineable Reserves Optimizer (CAE)	89			■	■		
Mineable Shape Optimizer	88			■	■		■
mineCAD	5			■			
mineCAVE	117			■			
mineHAUL	37			■			
mineMARKUP	38	■					
Mineral Beneficiation	106					■	■
MineSched	12				■		
mineSERV	49			■			
mineSTRUCTURE	50		■	■			
Minex	13	■	■	■			
MKP (Mining Knowledge Platform)	72	■	■				
MRM	98	■	■	■	■	■	■
NPV Scheduler (CAE)	91				■	■	■
NPV Scheduler (MineRP)	6				■	■	
Open Pit Metals	111				■		
PCBC	14				■		
Pegs Lite	97	■					
Performance Diagnostics	107						■
Portfolio Modelling	104					■	
Qerent Modeller	110					■	■
Sable Data Warehouse	83	■					
Services and Logistics	102			■	■		■
Sirovision	93	■	■				
Strat 3D	94		■				
Studio 3 - Engineering	85			■			
Studio 3 - Geology	84		■				■
Studio 5D Planner	96			■	■		
Studio3 - Basics	1						■
Surpac	10	■	■	■			
Talpac	68			■			
Underground Coal	112				■		
Ventsim Visual (Advanced)	119			■			
Vulcan	68	■	■	■			
Whittle	15					■	■
Workforce Planning	120					■	■
Xact	75				■		
Xeras	73					■	■
Xpac	71				■		

Figure 4.2: Updated list of mine planning software including CAE Mining software and their functionalities along the mining value chain

By using all three variables, the weighted software utilisation, hence the market impact of each participating mine planning software was calculated. The dataset were extracted from the updated database and a script written to calculate mine planning software utilisation by using all three variables. The programming language *GNU Octave* was used to execute the scrip for the data analysis and the calculation of the software utilisation per functionality for the selected commodities using two different time-stamps as mentioned previously. Furthermore, the accuracy of the written script was validated by using a verification script to confirm the link between commodity, functionality and time-stamp. Both scripts used to calculate mine planning software utilisation by using all three variables and the validation script to confirm the results are given in Appendix 9.2.

It is important to note that if $f_{i,l}^{(m,t)}$ is 0; it means either the particular software being analysed does not support the specific functionality or the subject software does not support the specific commodity. Furthermore, when calculating $u_{i,l}^{(m,t)}$ and $w_{i,l}^{(m,t)}$, the value is set to 1 as at this stage of calculation as it was decided that the weighted software utilisation did not have any impact on the calculation of the relative software utilisation. The Software_ID is a unique identifier of each particular software.

4.4 Results for coal

In this section, mine planning software utilisation for the commodity (*i*) coal was calculated. Six functionalities (*l*) with two time-stamps (*t*) were used for the calculations and the results for each functionality with two time-stamps were presented as tables and figures, respectively. Accordingly, the total number of {6(*l*) x 2(*t*) = 12} tables for each commodity were created. This is the reason why in Section 4.3 it was necessary to present the results for only two commodities as each commodity analysis produces 12 datasets. According to the functionality list provided earlier, the first functionality, “Geological Data Management” was used with two different time-stamps to produce the first sets of two tables. After generating the tables, pie charts were created for each table for easy interpretation of the results. Consequently, using the functionality list, the remaining tables and figures were created in a similar manner.

Table 4.2 shows the names of the software providers currently captured in the database, as well as the location of their headquarters. It is important to note that data on CAE Mining was only made available in the April 2014 dataset. The results presented here do not cater for the mining methods nor the type of mine.

Table 4.2: Software providers and their headquarter locations

Software Providers	Location of headquarters
CAE Mining	Saint-Laurent, Canada
Cyest Corporation	Johannesburg, South Africa
Geovia	Paris, France
Maptek	Adelaide, Australia
MineRP Solutions	Centurion, South Africa
RungePincockMinarco	Brisbane, Australia
Sable	Johannesburg, South Africa

4.4.1 Geological Data Management software results for coal

Table 4.3 shows the market impact of the individual software for the commodity *coal* using the functionality, Geological Data Management, as at September, 2012. Table 4.3 results were converted into a diagram, and Figure 4.3 is a graphical representation of Table 4.3.

When comparing the results between the September 2012 and April 2014 time-stamps, it was noted that both time-stamps were identical, indicating that there were no changes between the two different datasets. This is the reason why there is only one table showing results for both time-stamps and similarly Figure 4.3 represents both time-stamps.

It is important to note that as $f_{i,l}^{(m,t)}$ has been defined as the total number of sites where the particular software has been utilised, in Table 4.3, m=3, Software_ID 13 has been utilised in 14 different sites out of a total number of 23 sites indicating that it has 61% utilisation.

When interpreting the obtained results, Minex software clearly is the market leader and the most utilised mine planning software when it comes to the functionality, Geological Data Management. Minex was followed by Sable Data Warehouse with a 22% market impact for the commodity coal.

Table 4.3: Geological Data Management functionality software utilisation for coal

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	9	0	1	0	0.0000
2	10	2	1	2	0.0870
3	13	14	1	14	0.6087
4	38	0	1	0	0.0000
5	68	1	1	1	0.0435
6	72	0	1	0	0.0000
7	83	5	1	5	0.2174
8	93	0	1	0	0.0000
9	95	0	1	0	0.0000
10	97	1	1	1	0.0435
11	98	0	1	0	0.0000
12	100	0	1	0	0.0000
13	113	0	1	0	0.0000

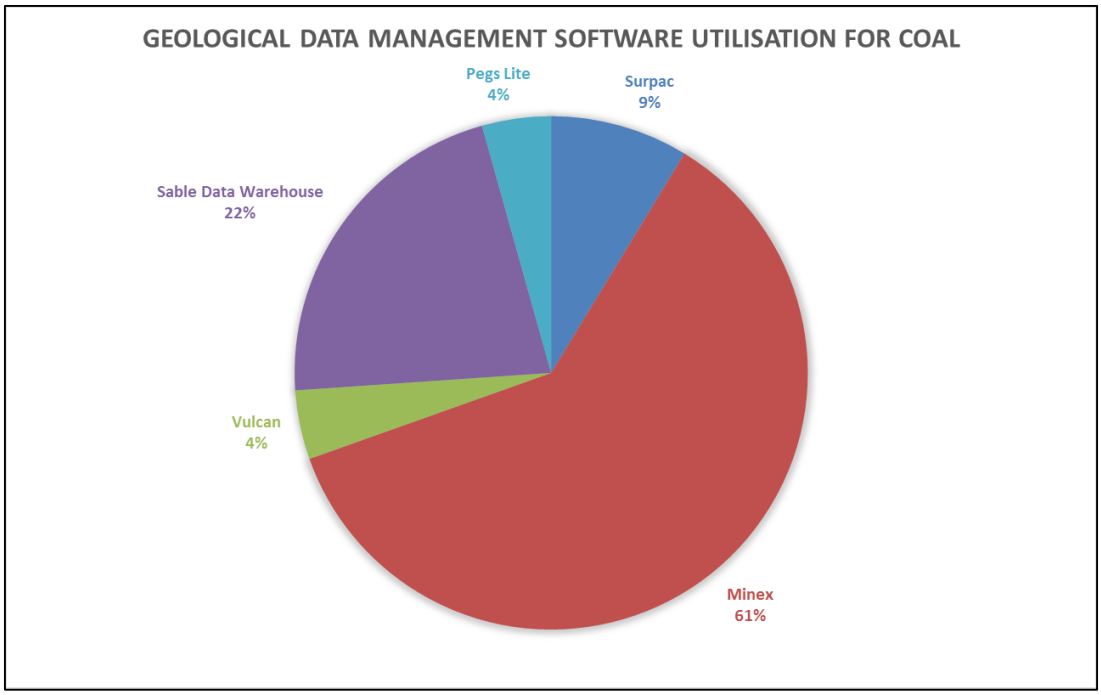


Figure 4.3: Geological Data Management functionality software utilisation for coal

4.4.2 Geological Modelling and Resource Estimation software results for coal

Geological Modelling and Resource Estimation software results for coal presented here were very similar to the previous section Geological Data Management software utilisation for coal, in terms of both time-stamps being identical. This is the reason why there is only one table (Table 4.4) showing results for Geological Modelling and Resource Estimation software and Figure 4.4 is the graphical representation of Table 4.4.

Table 4.4: Geological Modelling and Resource Estimation functionality software results for coal

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	9	0	1	0	0.0000
2	10	2	1	2	0.1176
3	13	14	1	14	0.8235
4	48	0	1	0	0.0000
5	68	1	1	1	0.0588
6	72	0	1	0	0.0000
7	84	0	1	0	0.0000
8	93	0	1	0	0.0000
9	94	0	1	0	0.0000
10	98	0	1	0	0.0000
11	99	0	1	0	0.0000

Geological modelling and resource estimation results for coal showed that Minex software once again is the market leader and the most utilised mine planning software in this functionality. While Minex is the leading software with 82%, it was followed by Surpac with 22% for the commodity coal.

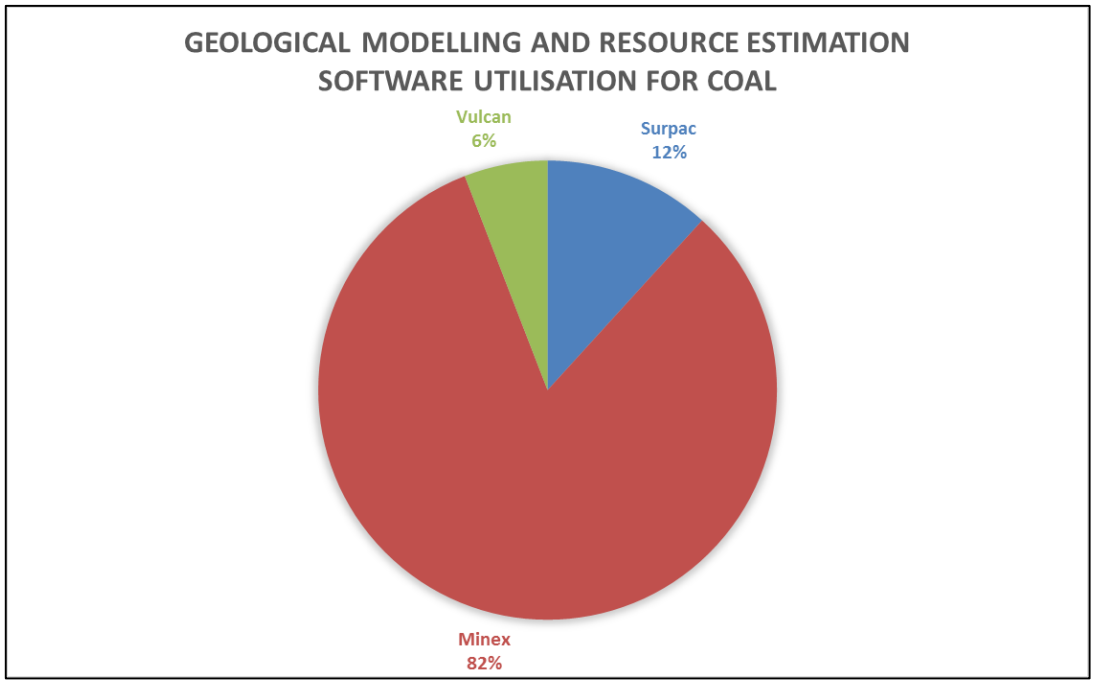


Figure 4.4: Geological Modelling and Resource Estimation functionality software results for coal

4.4.3 Design and Layout software results for coal

Table 4.5 shows the Design and Layout software results as at September 2012 for coal, while Table 4.6 shows results using a second time-stamp, i.e. April 2014. Figure 4.5 is a graphical representation of both tables.

When comparing both pie charts in Figure 4.5, the only difference between September 2012 and April 2014 is the inclusion of the Studio 5D Planner, a CAE Mine planning software in the 2014 dataset as CAE Mining made its dataset become the part of the Mine planning software database. Due to this addition, Studio 5D Planner has a 5% market impact in the April 2014 time-stamp. Nevertheless, Minex continues its dominance in the coal mining sector with a 64% market impact in April 2014 compared to 67% in September 2012, followed by Talpac software with 19% and 18% respectively, when Design and Layout functionality software results were evaluated.

Table 4.5: Design and Layout functionality software utilisation for coal as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	0	1	0	0.0000
2	5	0	1	0	0.0000
3	9	0	1	0	0.0000
4	10	2	1	2	0.0952
5	13	14	1	14	0.6667
6	31	0	1	0	0.0000
7	32	0	1	0	0.0000
8	46	0	1	0	0.0000
9	48	0	1	0	0.0000
10	49	0	1	0	0.0000
11	68	1	1	1	0.0476
12	70	4	1	4	0.1905
13	85	0	1	0	0.0000
14	86	0	1	0	0.0000
15	88	0	1	0	0.0000
16	89	0	1	0	0.0000
17	90	0	1	0	0.0000
18	96	0	1	0	0.0000
19	98	0	1	0	0.0000
20	99	0	1	0	0.0000
21	101	0	1	0	0.0000
22	102	0	1	0	0.0000

Table 4.6: Design and Layout functionality software utilisation for coal as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	0	1	0	0.0000
2	5	0	1	0	0.0000
3	9	0	1	0	0.0000
4	10	2	1	2	0.0909
5	13	14	1	14	0.6364
6	31	0	1	0	0.0000
7	32	0	1	0	0.0000
8	46	0	1	0	0.0000
9	48	0	1	0	0.0000
10	49	0	1	0	0.0000
11	68	1	1	1	0.0455
12	70	4	1	4	0.1818
13	85	0	1	0	0.0000
14	86	0	1	0	0.0000
15	88	0	1	0	0.0000
16	89	0	1	0	0.0000
17	90	0	1	0	0.0000
18	96	1	1	1	0.0455
19	98	0	1	0	0.0000
20	99	0	1	0	0.0000
21	101	0	1	0	0.0000
22	102	0	1	0	0.0000

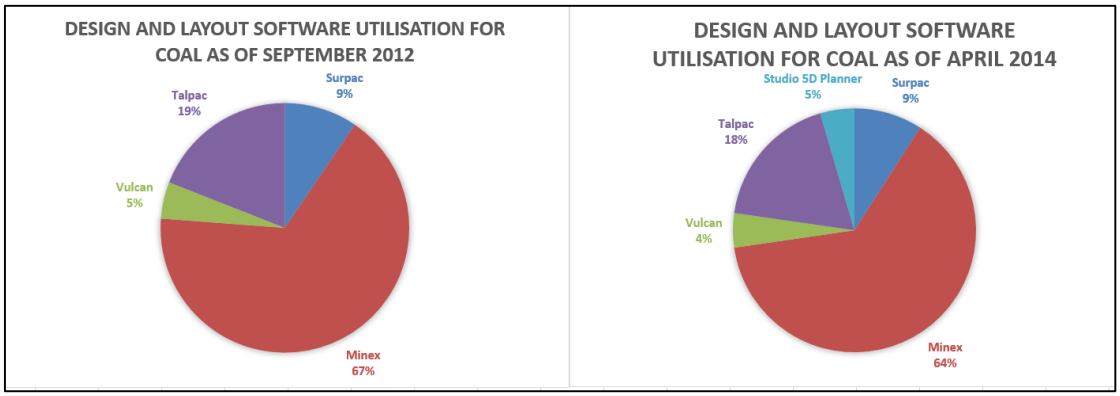


Figure 4.5: Design and Layout functionality software utilisation for coal for both time-stamp

4.4.4 Scheduling software results for coal

Table 4.7 shows the Scheduling software results as at September 2012 for coal, while Table 4.8 shows results using the second time-stamp, i.e. April 2014. Figure 4.6 is a graphical representation of both tables.

When comparing both pie charts in Figure 4.6, similar to the previous Design and layout section, the CAE Mine planning software presence is once again visible with Studio 5D Planner and Enhanced Production Scheduler (CAE) software. Both software have a 2% market impact in the April 2014 time-stamp, respectively. However, although it has lost 3% of the market impact between September 2012 and April 2014, Xpac is still the leader in the coal mining sector when it comes to Scheduling functionality software. Xpac is followed by Dragsim software with an 11% market impact on both the 2012 and 2014 time-stamps.

Table 4.7: Scheduling functionality software utilisation for coal as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	0	1	0	0.0000
2	4	0	1	0	0.0000
3	7	0	1	0	0.0000
4	12	2	1	2	0.0385
5	14	0	1	0	0.0000
6	20	0	1	0	0.0000
7	21	0	1	0	0.0000
8	33	0	1	0	0.0000
9	69	0	1	0	0.0000
10	71	41	1	41	0.7885
11	74	6	1	6	0.1154
12	75	3	1	3	0.0577
13	76	0	1	0	0.0000
14	80	0	1	0	0.0000
15	81	0	1	0	0.0000
16	86	0	1	0	0.0000
17	87	0	1	0	0.0000
18	88	0	1	0	0.0000
19	89	0	1	0	0.0000
20	91	0	1	0	0.0000
21	96	0	1	0	0.0000
22	98	0	1	0	0.0000
23	99	0	1	0	0.0000
24	101	0	1	0	0.0000
25	102	0	1	0	0.0000
26	108	0	1	0	0.0000
27	109	0	1	0	0.0000
28	111	0	1	0	0.0000
29	112	0	1	0	0.0000
30	113	0	1	0	0.0000
31	114	0	1	0	0.0000

Table 4.8: Scheduling functionality software utilisation for coal as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	0	1	0	0.0000
2	4	0	1	0	0.0000
3	7	0	1	0	0.0000
4	12	2	1	2	0.0370
5	14	0	1	0	0.0000
6	20	0	1	0	0.0000
7	21	0	1	0	0.0000
8	33	0	1	0	0.0000
9	69	0	1	0	0.0000
10	71	41	1	41	0.7593
11	74	6	1	6	0.1111
12	75	3	1	3	0.0556
13	76	0	1	0	0.0000
14	80	0	1	0	0.0000
15	81	0	1	0	0.0000
16	86	0	1	0	0.0000
17	87	1	1	1	0.0185
18	88	0	1	0	0.0000
19	89	0	1	0	0.0000
20	91	0	1	0	0.0000
21	96	1	1	1	0.0185
22	98	0	1	0	0.0000
23	99	0	1	0	0.0000
24	101	0	1	0	0.0000
25	102	0	1	0	0.0000
26	108	0	1	0	0.0000
27	109	0	1	0	0.0000
28	111	0	1	0	0.0000
29	112	0	1	0	0.0000
30	113	0	1	0	0.0000
31	114	0	1	0	0.0000

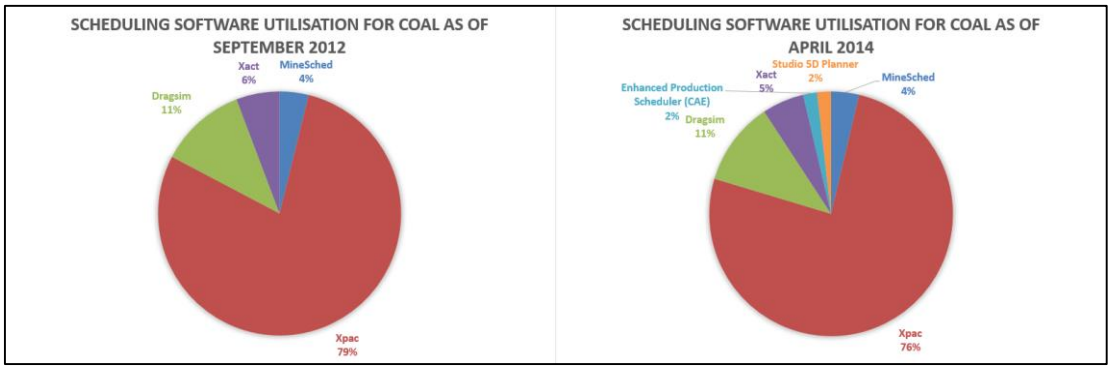


Figure 4.6: Scheduling functionality software utilisation for coal for both time-stamp

4.4.5 Financial Valuation software results for coal

Table 4.9 shows the Financial Valuation software results for coal. In this section, both time-stamps were found to be identical. This is the reason why there is only one table (Table 4.9) showing results for coal for the Financial Valuation software. On the other hand, Figure 4.7 is a graphical representation of Table 4.9.

It can be clearly seen that Xeras is the market leader in the coal mining sector when it comes to the Financial Valuation software with a 92% impact. Xeras is followed by Carbon Economics software with an 8% market impact.

Table 4.9: Financial Valuation functionality software utilisation for coal

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	7	0	1	0	0.0000
2	15	0	1	0	0.0000
3	73	11	1	11	0.9167
4	77	1	1	1	0.0833
5	78	0	1	0	0.0000
6	79	0	1	0	0.0000
7	80	0	1	0	0.0000
8	91	0	1	0	0.0000
9	92	0	1	0	0.0000
10	98	0	1	0	0.0000
11	103	0	1	0	0.0000
12	104	0	1	0	0.0000
13	105	0	1	0	0.0000
14	106	0	1	0	0.0000
15	109	0	1	0	0.0000
16	110	0	1	0	0.0000

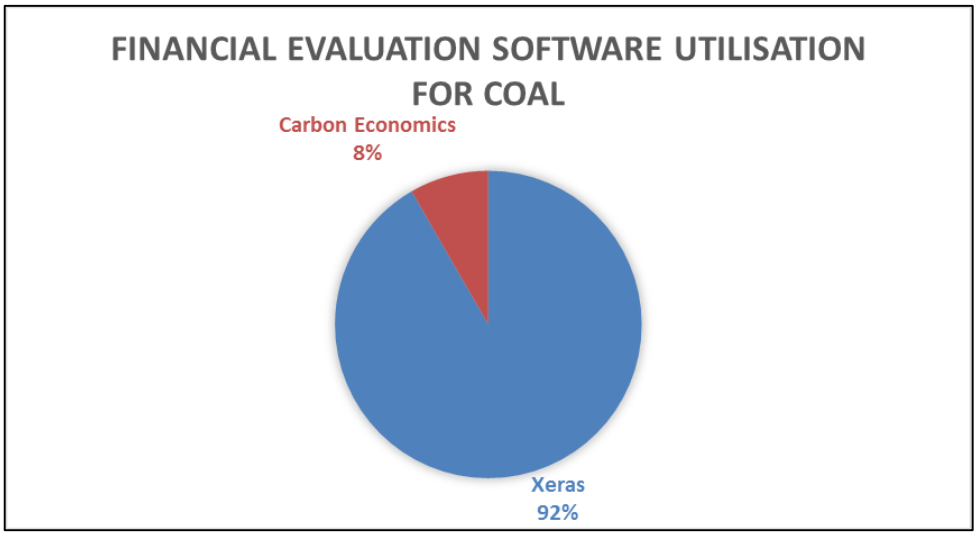


Figure 4.7: Financial Valuation functionality software utilisation for coal

4.4.6 Optimisation software results for coal

Table 4.10 shows the Optimisation software results as at September 2012 for coal, while Table 4.11 shows results using the second time-stamp, April 2014. Figure 4.8 is a graphical representation of both tables.

The graphical representation of the Optimisation software results is shown in Figure 4.8; the only difference between both pie charts is the presence of the Enhanced Production Scheduler (CAE) software in the April 2014 chart, with a 5% market impact. However, Xeras (58%) and Dragsim (32%) respectively continued their dominance in the coal mining sector during 2014 when it comes to Optimisation functionality software.

Table 4.10: Optimisation functionality software utilisation for coal as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	1	0	1	0	0.0000
2	15	0	1	0	0.0000
3	21	0	1	0	0.0000
4	73	11	1	11	0.6111
5	74	6	1	6	0.3333
6	77	1	1	1	0.0556
7	79	0	1	0	0.0000
8	82	0	1	0	0.0000
9	84	0	1	0	0.0000
10	87	0	1	0	0.0000
11	88	0	1	0	0.0000
12	91	0	1	0	0.0000
13	92	0	1	0	0.0000
14	98	0	1	0	0.0000
15	102	0	1	0	0.0000
16	103	0	1	0	0.0000
17	105	0	1	0	0.0000
18	106	0	1	0	0.0000
19	107	0	1	0	0.0000
20	110	0	1	0	0.0000

Table 4.11: Optimisation functionality software utilisation for coal as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	1	0	1	0	0.0000
2	15	0	1	0	0.0000
3	21	0	1	0	0.0000
4	73	11	1	11	0.5789
5	74	6	1	6	0.3158
6	77	1	1	1	0.0526
7	79	0	1	0	0.0000
8	82	0	1	0	0.0000
9	84	0	1	0	0.0000
10	87	1	1	1	0.0526
11	88	0	1	0	0.0000
12	91	0	1	0	0.0000
13	92	0	1	0	0.0000
14	98	0	1	0	0.0000
15	102	0	1	0	0.0000
16	103	0	1	0	0.0000
17	105	0	1	0	0.0000
18	106	0	1	0	0.0000
19	107	0	1	0	0.0000
20	110	0	1	0	0.0000

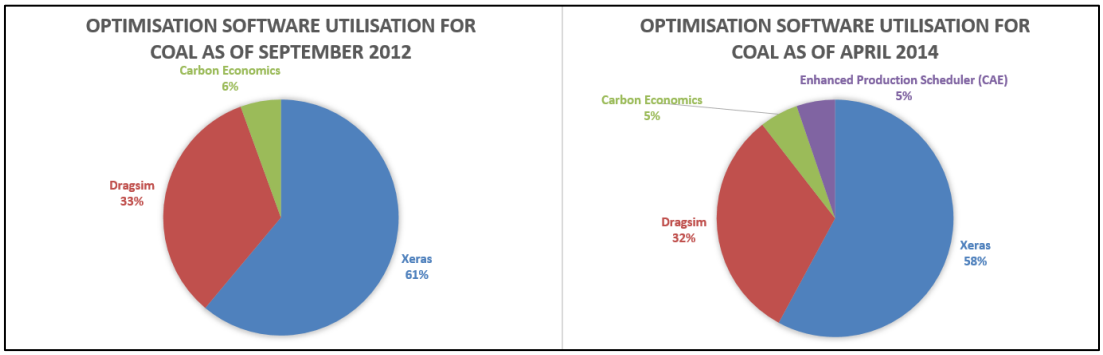


Figure 4.8: Optimisation functionality software utilisation for Coal for both time-stamp

4.5 Results for Platinum Group Metals

In this section, strategic software utilisation for the commodity (*i*) Platinum Group Metals (PGM) was calculated. Using the approach similar to the one taken in the previous results section for the coal, six functionalities (*l*) with two time-stamps (*t*) were used for the calculations and the results for each functionality with the two time-

stamps were presented as tables and figures, respectively. Accordingly, a total number of {6(*l*) x 2(*t*) =12} tables for each commodity were created.

4.5.1 Geological Data Management software results for PGM

Table 4.12 shows the market impact of the individual software for the commodity PGM using the functionality Geological Data Management as at September, 2012 while Table 4.13 shows results using the second time-stamp, April 2014. Figure 4.9 is a graphical representation of both tables.

When comparing the diagrams in Figure 4.9, there is a significant difference between the two pie charts; CAE Mining's Geological data management solution software with a 47% market impact in the April 2014 chart is clearly visible. Pegs Lite and MRM have each a 16% market impact in this field, respectively. A graphical representation of the Geological Data Management functionality software results is shown in Figure 4.9.

Table 4.12: Geological Data Management functionality software utilisation for PGM as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	9	2	1	2	0.0290
2	10	6	1	6	0.0870
3	13	4	1	4	0.0580
4	38	1	1	1	0.0145
5	68	1	1	1	0.0145
6	72	0	1	0	0.0000
7	83	12	1	12	0.1739
8	93	0	1	0	0.0000
9	95	0	1	0	0.0000
10	97	22	1	22	0.3188
11	98	21	1	21	0.3043
12	100	0	1	0	0.0000
13	113	0	1	0	0.0000

Table 4.13: Geological Data Management functionality software utilisation for PGM as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	9	2	1	2	0.0148
2	10	6	1	6	0.0444
3	13	4	1	4	0.0296
4	38	1	1	1	0.0074
5	68	1	1	1	0.0074
6	72	0	1	0	0.0000
7	83	12	1	12	0.0889
8	93	0	1	0	0.0000
9	95	64	1	64	0.4741
10	97	22	1	22	0.1630
11	98	21	1	21	0.1556
12	100	2	1	2	0.0148
13	113	0	1	0	0.0000

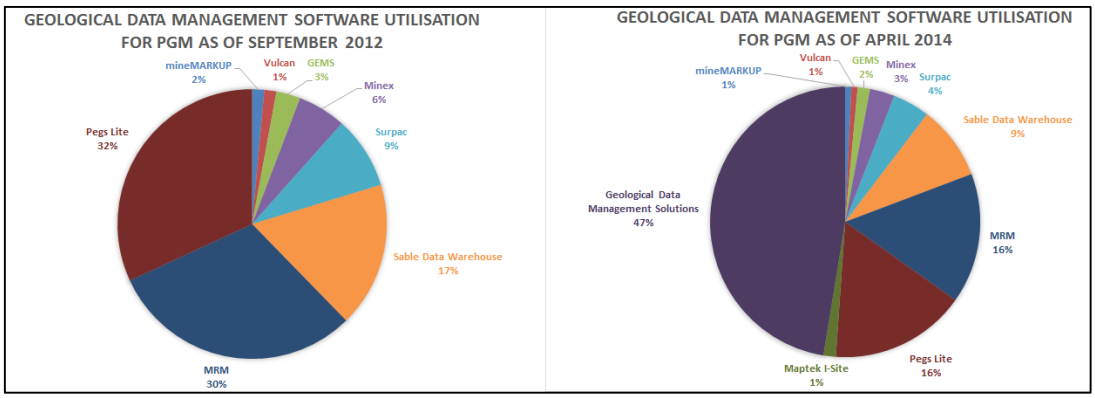


Figure 4.9: Geological Data Management functionality software utilisation for PGM

4.5.2 Geological Modelling and Resource Estimation software results for PGM

Table 4.14 shows the market impact of the individual software for the commodity PGM using the Geological Modelling and Resource Estimation functionality as at September, 2012 while Table 4.15 shows results using the second time-stamp, April 2014. Figure 4.10 is a graphical representation of both tables.

When comparing the diagrams in Figure 4.10, there is also another significant difference between the two pie charts; Studio 3 - Geology is the leading software with

a 50% market impact in the April 2014 chart. MRM and CADSMine software has a 19% and 18% market impact in this field, respectively.

Table 4.14: Geological Modelling and Resource Estimation functionality software utilisation for PGM as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	9	2	1	2	0.0370
2	10	6	1	6	0.1111
3	13	4	1	4	0.0741
4	48	0	1	0	0.0000
5	68	1	1	1	0.0185
6	72	0	1	0	0.0000
7	84	0	1	0	0.0000
8	93	0	1	0	0.0000
9	94	0	1	0	0.0000
10	98	21	1	21	0.3889
11	99	20	1	20	0.3704

Table 4.15: Geological Modelling and Resource Estimation functionality software utilisation for PGM as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	9	2	1	2	0.0183
2	10	6	1	6	0.0550
3	13	4	1	4	0.0367
4	48	0	1	0	0.0000
5	68	1	1	1	0.0092
6	72	0	1	0	0.0000
7	84	55	1	55	0.5046
8	93	0	1	0	0.0000
9	94	0	1	0	0.0000
10	98	21	1	21	0.1927
11	99	20	1	20	0.1835

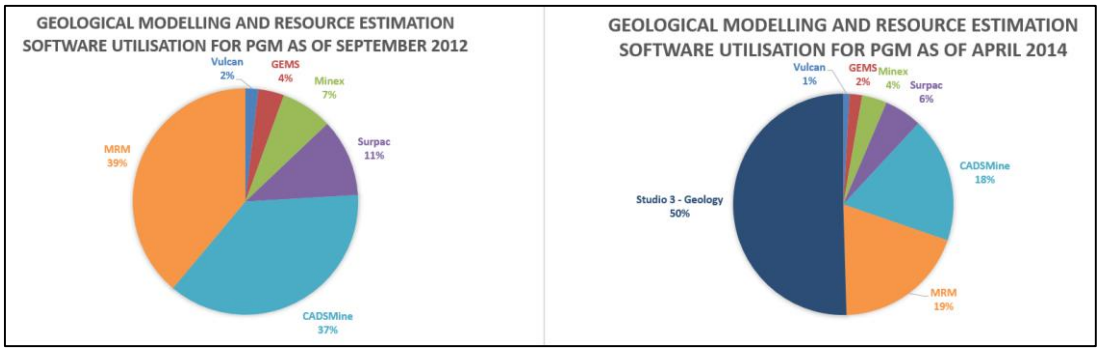


Figure 4.10: Geological Modelling and Resource Estimation functionality software utilisation for PGM

4.5.3 Design and Layout software results for PGM

Table 4.16 shows the market impact of the individual software for the commodity PGM using the Design and Layout functionality as at September, 2012 while Table 4.17 shows results using the second time-stamp, April 2014. Figure 4.11 is a graphical representation of both tables.

When comparing the diagrams in Figure 4.11, there are some differences between the two pie charts; although MRM and CADSMine software still have the biggest slices from the pie with a 25% and 24% market impact in the Design and Layout software field, respectively, Studio 3 – Engineering software is in the third place with a 13% market impact in the PGM sector using the Design and Layout functionality.

Table 4.16: Design and Layout functionality software utilisation for PGM as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	6	1	6	0.0857
2	5	2	1	2	0.0286
3	9	2	1	2	0.0286
4	10	6	1	6	0.0857
5	13	4	1	4	0.0571
6	31	0	1	0	0.0000
7	32	0	1	0	0.0000
8	46	0	1	0	0.0000
9	48	0	1	0	0.0000
10	49	0	1	0	0.0000
11	68	1	1	1	0.0143
12	70	2	1	2	0.0286
13	85	0	1	0	0.0000
14	86	0	1	0	0.0000
15	88	0	1	0	0.0000
16	89	0	1	0	0.0000
17	90	0	1	0	0.0000
18	96	0	1	0	0.0000
19	98	21	1	21	0.3000
20	99	20	1	20	0.2857
21	101	6	1	6	0.0857
22	102	0	1	0	0.0000

Table 4.17: Design and Layout functionality software utilisation for PGM as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	6	1	6	0.0706
2	5	2	1	2	0.0235
3	9	2	1	2	0.0235
4	10	6	1	6	0.0706
5	13	4	1	4	0.0471
6	31	0	1	0	0.0000
7	32	0	1	0	0.0000
8	46	0	1	0	0.0000
9	48	0	1	0	0.0000
10	49	0	1	0	0.0000
11	68	1	1	1	0.0118
12	70	2	1	2	0.0235
13	85	11	1	11	0.1294
14	86	0	1	0	0.0000
15	88	0	1	0	0.0000
16	89	1	1	1	0.0118
17	90	0	1	0	0.0000
18	96	3	1	3	0.0353
19	98	21	1	21	0.2471
20	99	20	1	20	0.2353
21	101	6	1	6	0.0706
22	102	0	1	0	0.0000

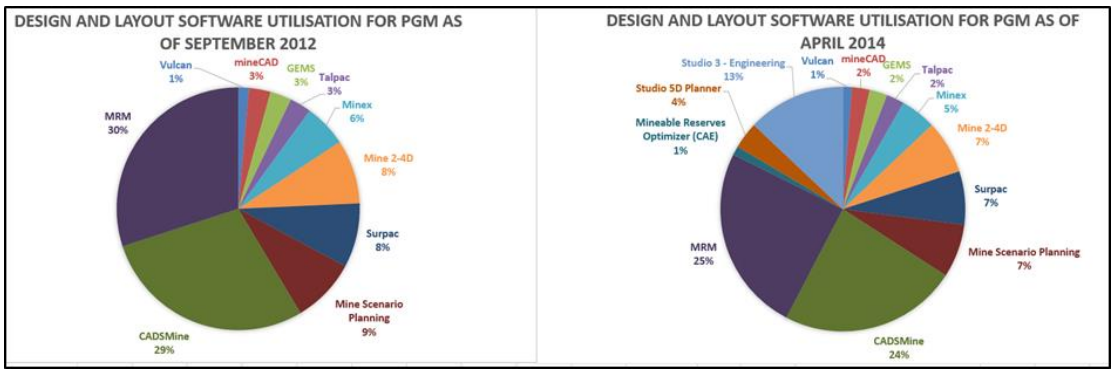


Figure 4.11: Design and Layout functionality software utilisation for PGM

4.5.4 Scheduling software results for PGM

Table 4.18 shows the market impact of the individual software for the commodity PGM using the Scheduling functionality as at September, 2012 while Table 4.19 shows

results using the second time-stamp, April 2014. Figure 4.12 illustrates the graphical representation of Table 4.17 while Figure 4.13 shows the graphical representation of Table 4.18.

When comparing Figure 4.12 and Figure 4.13, there is not much difference between the two figures; MRM and CADSMine software still have the biggest market impact with 17% and 16%, respectively in the Scheduling software field.

Table 4.18: Scheduling functionality software utilisation for PGM as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	6	1	6	0.0508
2	4	7	1	7	0.0593
3	7	0	1	0	0.0000
4	12	4	1	4	0.0339
5	14	2	1	2	0.0169
6	20	1	1	1	0.0085
7	21	0	1	0	0.0000
8	33	1	1	1	0.0085
9	69	1	1	1	0.0085
10	71	4	1	4	0.0339
11	74	0	1	0	0.0000
12	75	1	1	1	0.0085
13	76	11	1	11	0.0932
14	80	1	1	1	0.0085
15	81	2	1	2	0.0169
16	86	0	1	0	0.0000
17	87	0	1	0	0.0000
18	88	0	1	0	0.0000
19	89	0	1	0	0.0000
20	91	0	1	0	0.0000
21	96	0	1	0	0.0000
22	98	21	1	21	0.1780
23	99	20	1	20	0.1695
24	101	6	1	6	0.0508
25	102	0	1	0	0.0000
26	108	15	1	15	0.1271
27	109	15	1	15	0.1271
28	111	0	1	0	0.0000
29	112	0	1	0	0.0000
30	113	0	1	0	0.0000
31	114	0	1	0	0.0000

Table 4.19: Scheduling functionality software utilisation for PGM as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	2	6	1	6	0.0480
2	4	7	1	7	0.0560
3	7	0	1	0	0.0000
4	12	4	1	4	0.0320
5	14	2	1	2	0.0160
6	20	1	1	1	0.0080
7	21	0	1	0	0.0000
8	33	1	1	1	0.0080
9	69	1	1	1	0.0080
10	71	4	1	4	0.0320
11	74	0	1	0	0.0000
12	75	1	1	1	0.0080
13	76	11	1	11	0.0880
14	80	1	1	1	0.0080
15	81	2	1	2	0.0160
16	86	0	1	0	0.0000
17	87	3	1	3	0.0240
18	88	0	1	0	0.0000
19	89	1	1	1	0.0080
20	91	0	1	0	0.0000
21	96	3	1	3	0.0240
22	98	21	1	21	0.1680
23	99	20	1	20	0.1600
24	101	6	1	6	0.0480
25	102	0	1	0	0.0000
26	108	15	1	15	0.1200
27	109	15	1	15	0.1200
28	111	0	1	0	0.0000
29	112	0	1	0	0.0000
30	113	0	1	0	0.0000
31	114	0	1	0	0.0000

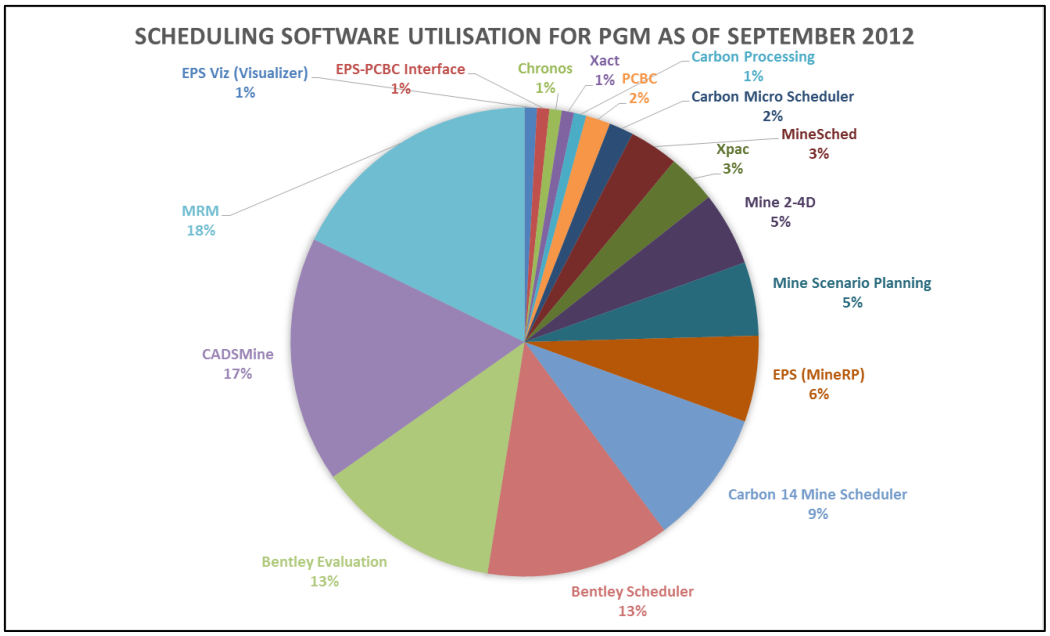


Figure 4.12: Scheduling functionality software utilisation for PGM as of September 2012

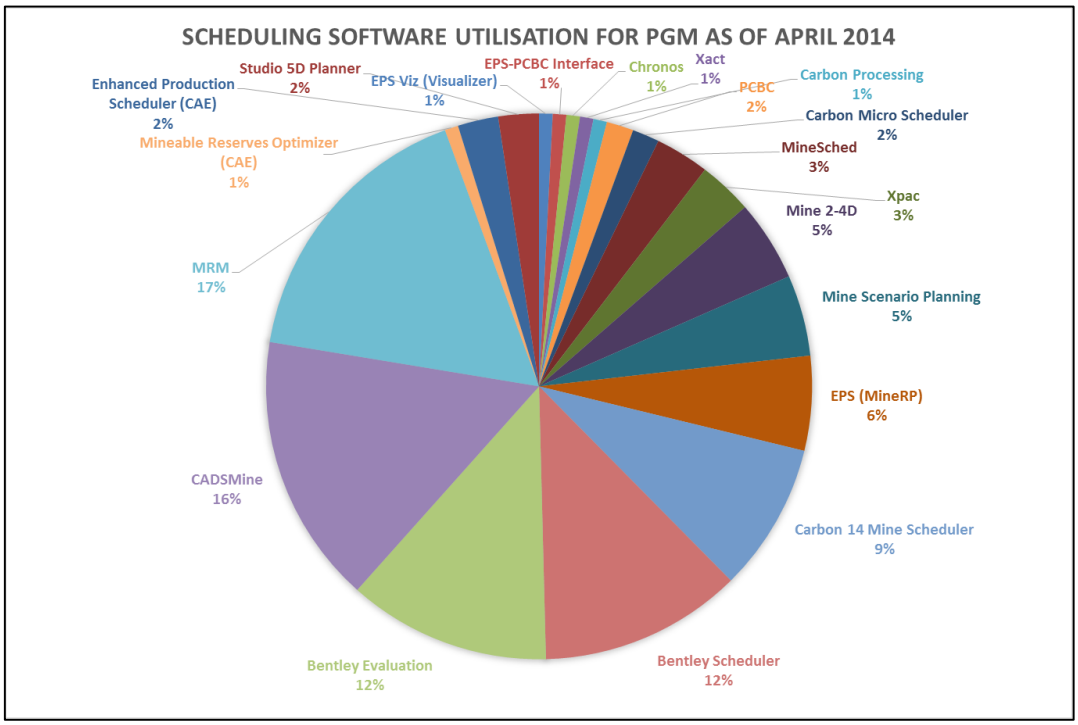


Figure 4.13: Scheduling functionality software utilisation for PGM as of April 2014

4.5.5 Financial Valuation software results for PGM

In this section, when comparing the results for the Financial Valuation functionality software between September 2012 and April 2014, it was noted that both time-stamps were identical, indicating that there were no changes between the two different datasets. Due to this, there is only one table, Table 4.20 showing the market impact of the individual software for the commodity PGM using the Financial Valuation functionality results for both time-stamps. Similarly, Figure 4.14 represents both time-stamps and illustrates the graphical representation of Table 4.20.

Table 4.20: Financial Valuation functionality software utilisation for PGM

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	7	0	1	0	0.0000
2	15	0	1	0	0.0000
3	73	1	1	1	0.0143
4	77	15	1	15	0.2143
5	78	1	1	1	0.0143
6	79	0	1	0	0.0000
7	80	1	1	1	0.0143
8	91	0	1	0	0.0000
9	92	0	1	0	0.0000
10	98	21	1	21	0.3000
11	103	0	1	0	0.0000
12	104	0	1	0	0.0000
13	105	0	1	0	0.0000
14	106	12	1	12	0.1714
15	109	15	1	15	0.2143
16	110	4	1	4	0.0571

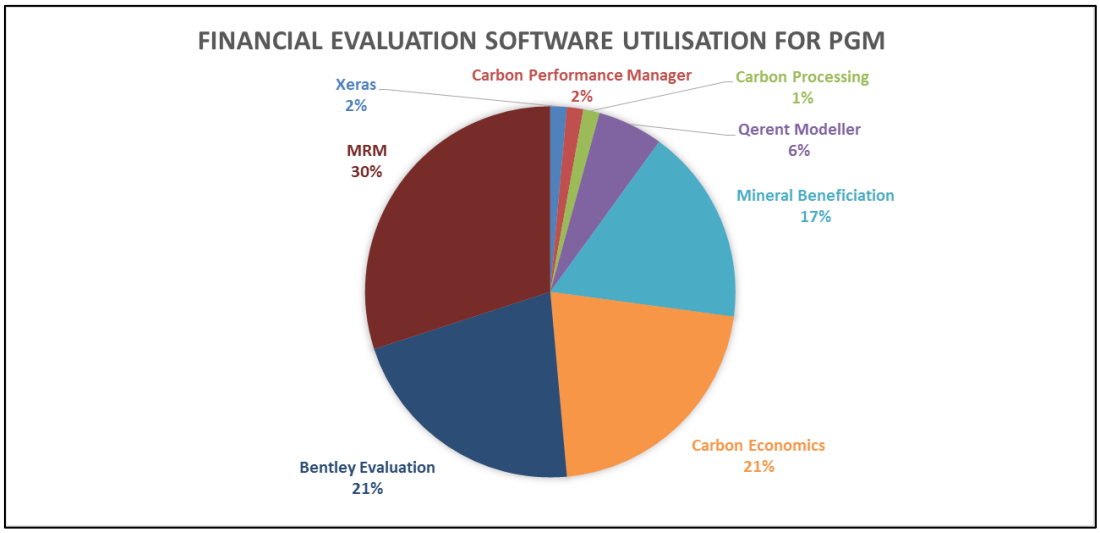


Figure 4.14: Financial Valuation functionality software utilisation for PGM

Figure 4.14 indicates that MRM is the leading software with a 30% market impact in the PGM sector when it comes to the Financial Valuation software. Bentley Valuation and Carbon Economics both have a 21% market impact in this section.

4.5.6 Optimisation software results for PGM

Table 4.21 shows the market impact of the individual software for the commodity PGM using the Optimisation functionality as at September, 2012 while Table 4.22 shows results using the second time-stamp, April 2014. Figure 4.15 is a graphical representation of Table 4.21 and Table 4.22.

Table 4.21: Optimisation functionality software utilisation for PGM as of September 2012

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	1	0	1	0	0.0000
2	15	0	1	0	0.0000
3	21	0	1	0	0.0000
4	73	1	1	1	0.0189
5	74	0	1	0	0.0000
6	77	15	1	15	0.2830
7	79	0	1	0	0.0000
8	82	0	1	0	0.0000
9	84	0	1	0	0.0000
10	87	0	1	0	0.0000
11	88	0	1	0	0.0000
12	91	0	1	0	0.0000
13	92	0	1	0	0.0000
14	98	21	1	21	0.3962
15	102	0	1	0	0.0000
16	103	0	1	0	0.0000
17	105	0	1	0	0.0000
18	106	12	1	12	0.2264
19	107	0	1	0	0.0000
20	110	4	1	4	0.0755

Table 4.22: Optimisation functionality software utilisation for PGM as of April 2014

m	Software_ID	$f_{i,l}^{(m,t)}$	$w_{i,l}^{(m,t)}$	$u_{i,l}^{(m,t)}$	$r_{i,l}^{(m,t)}$
1	1	2	1	2	0.0177
2	15	0	1	0	0.0000
3	21	0	1	0	0.0000
4	73	1	1	1	0.0088
5	74	0	1	0	0.0000
6	77	15	1	15	0.1327
7	79	0	1	0	0.0000
8	82	0	1	0	0.0000
9	84	55	1	55	0.4867
10	87	3	1	3	0.0265
11	88	0	1	0	0.0000
12	91	0	1	0	0.0000
13	92	0	1	0	0.0000
14	98	21	1	21	0.1858
15	102	0	1	0	0.0000
16	103	0	1	0	0.0000
17	105	0	1	0	0.0000
18	106	12	1	12	0.1062
19	107	0	1	0	0.0000
20	110	4	1	4	0.0354

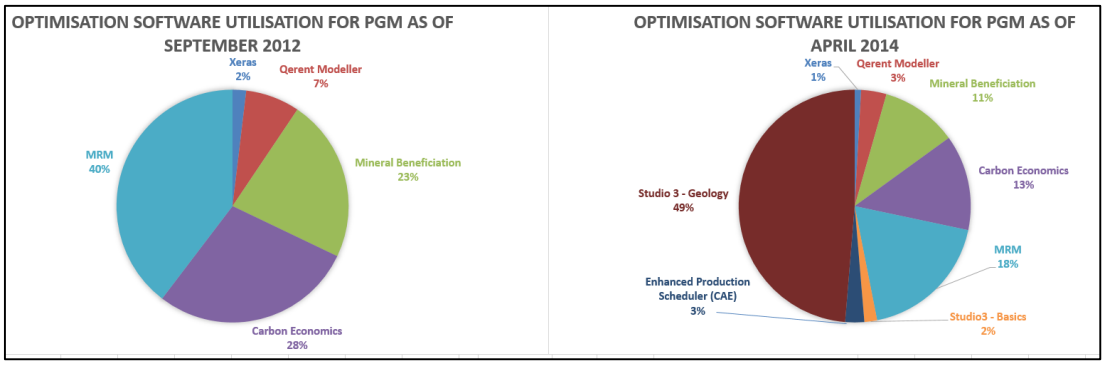


Figure 4.15: Optimisation functionality software utilisation for PGM

When comparing the diagrams in Figure 4.15, there is a noteworthy difference between the two pie charts; Studio 3 – Geology has emerged as a new leader with a 49% market impact in April 2014, which is followed by MRM with an 18% market impact in the Optimisation software field. Studio 3 – Geology may seem to be an anomaly with respect to Optimisation software. However, similar to the explanation provided for "Talpac" software, Studio 3 – Geology software's functionality as captured in the database was provided by the software supplier. Furthermore, Studio 3 – Geology which was previously known as "Datamine" is widely used in mining industry throughout Design and Layout, Scheduling and Optimisation stages.

4.6 Discussion of the results and indicative decision-making strategies

The results showed that the most widely used mine planning software in terms of number of licenses sold varies in the mining industry. Although there are some similarities when comparing the results, the market leaders in terms of mine planning software utilisation in South Africa differs, depending on the commodity that is being mined as well as the functionality that is being used. By using this newly developed framework, utilisation of the various mine planning software were measured. Now, the South African mining industry has a framework that can be utilised in order to ultimately enhance decision-making strategies for utilisation of the mine planning software. This information can then be used further too, to better inform decision-making strategies on software utilisation by software developers, educational institutions and consulting companies from different perspectives. South African mining companies could use this framework to choose the leading software combinations to improve productivity, the educational institutions could make use of

the framework to choose which software solutions they can expose their students to, consulting companies can advise their clients more effectively by using the framework and more importantly, software producers can compare their positions and develop more informed decision-making strategies accordingly.

As mentioned previously, data for CAE Mining was only made available in the April 2014 dataset. When comparing the results, the CAE Mining market impact in the South African mining industry is undoubtedly noticeable, especially in the PGM sector, specifically in the field of Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout and Optimisation.

As previously mentioned, mergers and acquisitions are a reality in the mine planning software space. More recently, after acquiring the “Datamine” mine planning software to enter the mining sector in 2010, CAE Mining announced that it was putting the mining unit up for sale in August, 2014 (Owram, 2014). Previously, Gemcom Software, a well-known mine planning software was acquired by Dassault Systèmes in July, 2012 and has become the new brand of GEOVIA. These developments show that similar activities can be expected to re-shape the mine planning software space hence the need to regularly update the database. Furthermore, there are other mine planning software solutions, which are trying to get into the South African mining sector. One is Trimble Open Pit solution software, which made a significant contribution towards Kumba Iron Ore’s mining operations which were mentioned by Trimble Software (Trimble, 2014) during SAIMM’s “Mine Planning School – 2014” presentation. These developments put forward a very important question: Is there a sufficiently large market available for so many different mine planning software solutions in South Africa? To answer this question, it needs to be remembered that while there are a number of software solutions available in the South African mining sector, each of these companies has an extensive product and consulting portfolio to get the balance right in this very competitive market.

4.7 Summary

In this chapter, a methodology for the evaluation of utilisation of mine planning software in the South African mining industry was developed. In this framework, three variables, namely, commodity (*i*), functionality (*l*), and time factor (*t*) were used to

calculate the results. Although calculations were done only on two commodities in this Chapter, Table 4.23 extended the calculation results to four commodities namely coal, diamond, gold and PGM; six functionalities namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation were applied on commodities using two different time-stamps (September 2012 and April 2014). By using this newly developed framework, utilisation of the various mine planning software were measured. Now, the South African mining industry has a framework in order to enhance decision-making strategies for utilisation of the mine planning software. This information can then be used to better inform decision-making strategies on software utilisation by software developers, educational institutions and consulting companies through purchase, training, financial returns, etc. Financial return can be measured by comparing year by year balance sheets.

It is important to note that data for CAE Mining was only made available in the April 2014 dataset. When comparing the results, the CAE Mining market impact is clearly visible, especially in the PGM sector, specifically in the functionalities of Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout and Optimisation. Regular updating of the database implies an evolving dataset necessitating time-series evaluation of the utilisation of the software. However, the market leaders in terms of mine planning software utilisation in South Africa differs, depending on the commodity that is being mined as well as the functionality that is being used.

One of the main objectives of this thesis is to predict future values of mine planning software utilisation within the South African mining industry by doing time-series analyses. Time-series analyses are conducted by the help of machine learning techniques i.e. Artificial Neural Networks. Although there are a number of techniques available to conduct the time-series analyses efficiently, Artificial Neural Networks stand out from the rest as one of the most used methodology to analyse the time-series data. In Chapter 5, Artificial Neural Networks are introduced and explained, and their use motivated.

5 PREDICTING SOFTWARE UTILISATION USING ARTIFICIAL NEURAL NETWORKS

5.1 Introduction

In the previous chapter, the software utilisation framework was defined and measured, and the required foundation to start time-series analysis was established. As the existing database will be continuously evolving, it is necessary to establish a framework so that it can be used for further analysis of the database. The idea behind this framework is to predict not only the next time-stamp i.e. (t+1), but also further time-stamps (t+2, t+3, … t+T_f where T_f is the last lag of desired time step) considering the fact that this evolution will continue in the future. Although there are a number of techniques available to conduct time-series analysis efficiently, as previously mentioned in Chapter 4, Artificial Neural Networks stand out from the rest as one of the most used methodology to analyse time-series data. In this chapter, Artificial Neural Networks are introduced and explained and their use motivated.

Recent development in the information technology field caused some huge advances in the science of Artificial Neural Networks (ANNs). Consequently, these advances especially in the area of pattern recognition such as object tracking and face recognition made ANNs a wide spread technique (Jain, et al., 1996). Additionally, ANNs can learn from the dataset and generalise things that were learned to produce meaningful solutions. Furthermore, ANNs are capable of modelling any kind of a system regardless of whether it is a linear or nonlinear system, as well as it can be applied to a wide range of fields such as engineering, medicine and finance.

5.2 What is ANN?

An Artificial Neural Network (ANN) consists of neurons and pathways, which connect to the neurons. The most important part of this connection is how the information is processed. This process entails connecting a vast number of interconnected processing elements (neurons) so that they can answer the specific, subject questions. ANNs can learn like people and can be configured for a specific application, such as function approximation or pattern recognition. This can be achieved through a learning process. This learning process is accomplished through

establishing the connections between the neurons (Stergiou & Siganos, 1996). Figure 5.1 shows a schematic drawing of a neuron.

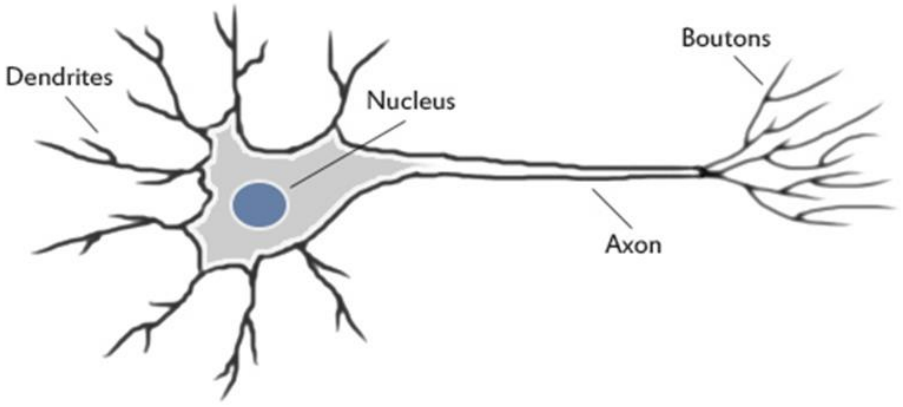


Figure 5.1: Schematic drawing of a neuron (Jacobson, 2013)

Some of the important developments regarding ANNs research started in the early 1940s (Jain, et al., 1996). Although so many researchers contributed towards the development of the ANNs, some of the pioneers who made the most impact for ANNs are:

- 1943 - McCulloch and Pitts: Formal representation of neuron
- 1949 - Hebb: Psychologist, the first learning rule
- 1958 - Rosenblatt: Perceptron convergence theorem
- 1960 - Widrow and Hoff: Least mean square (LMS) algorithm
- 1969 - Minsky and Papert: Limitations of a simple perceptron
- 1982 - Hopfield: Energy approach
- 1986 - Rumelhart, Hinton and Williams: Back-propagation learning algorithm

There are a number of important classes within ANNs which challenged researchers all over the world. Jain, et al., (1996, p. 2) furthermore listed these classes:

- “Pattern classification;
- Clustering/categorization;
- Function approximation;
- Prediction/forecasting;
- Optimisation;

- Content-addressable memory, and
- Control".

Over the years, researchers developed a number of models to solve problems associated with the above mentioned classes. Due to the successes of those developed models, ANNs are now one of the preferred subjects to solve these problems. The most commonly used tools are namely pattern classification, function approximation and prediction/forecasting, which are the most appropriate classes to predict time-series modelling done in this research. The other classes are not considered due to the fact that they are not designed to be used in forecasting.

5.2.1 Pattern classification

Pattern classification is a widely used tool used in ANNs to identify patterns. There is a wide variety of applications where pattern classification is often used, such as applications including handwritten digit recognition, object classification ((LeCun, et al., 1989); (Aha & Bankert, 1997) and (Murphey & Luo, 2002)), speech tagging and recognition ((Brill, 1994); (Even-Zohar & Roth, 2001)), bioinformatics ((Baldi & Pollastri, 2002), (Clare & King, 2001) and (Zhang, et al., 2005)), text categorization and information retrieval (Apte, et al., 1994). The pattern classification uses these applications as the input design to one of pre-determined classes (Figure 5.2 a). To be able to separate patterns from different classes, by using a set of training designs with pre-determined classes, the decision boundaries are constructed which can be a number of different shapes such as linear, non-linear or any other form (Jain, et al., 1996).

5.2.2 Function approximation

Function approximation is used often when it comes to approximate defined functions on a compact subset. According to Zainuddin & Ong (2007), function approximation finds the connection between input and output data in such a way that function approximation can make use of a given finite input-output data, and the function approximation can relate it to the number of classes such as prediction, classification and pattern recognition. As explained by Jain, et al. (1996), given a number of n

labelled input and output duos of learning patterns, $\{(x_1; y_1), (x_2; y_2),\ldots,(x_n; y_n)\}$, produced from an unidentified function *f*(x), the aim is to calculate the approximation, say, $\hat{f}(x)$ of unidentified function f(x) (Figure 5.2 b). The accuracy of the approximated function $\hat{f}(x)$ can be improved by adjusting the complexity of the trained dataset. Over-fitting is also important when it comes to the given noisy training data (Jain, et al., 1996).

According to MathWorks (2015) over-fitting is one of the problems that occur during the training of ANN. Over-fitting happens when ANN has the limited knowledge about training the neurons which are in the hidden layers. In another words, the ANN has memorised the training neurons but it has not learned to generalise. In this case, a "validation set" is used to ensure that there is no over-fitting of the learning process. The over-fitting problem usually occurs when training on large datasets and the remedy of validation is used to overcome the problem. However, in this study a small size data is used to perform prediction, thus the over-fitting problem is not an issue. Due to this reason, there is no need to validate the trained network.

5.2.3 Prediction/forecasting

Jain, et al. (1996) explained the prediction tool: Using a number of *n* models in a time-series, $\{y(t_1),y(t_2),\ldots y(t_n)|t_1 < t_2 < \ldots < t_n$, the aim is to forecast the sample $y(t_{n+1})$ at a forthcoming time t_{n+1}. It is an important part of daily life as it has been widely used in different types of fields from weather forecasting to stock market predictions (see Figure 5.2 c).

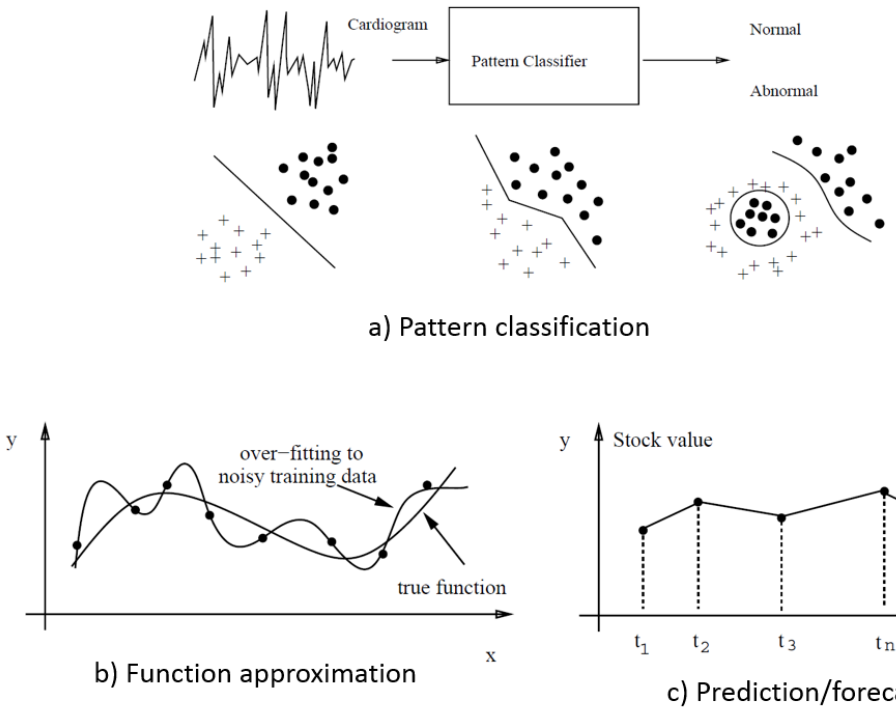


Figure 5.2: Commonly used neural networks tools: a) Pattern classification; b) Function approximation and c) Prediction/forecasting; (adopted from Jain et al., 1996)

5.3 ANNs Overview

In this section neurons with the network topology/architecture are presented and learning/training mechanisms are briefly explained. Several ANNs models are categorised by their topology and the training algorithm.

5.3.1 Neurons

The concept of neurons can be traced as far back as the 1940s where McCulloch & Pitts (1943) introduced mathematical neurons which become the milestone of ANNs research. Figure 5.3 shows McCulloch-Pitts's neuron. The neurons are the elements that are connected to each other to form the network.

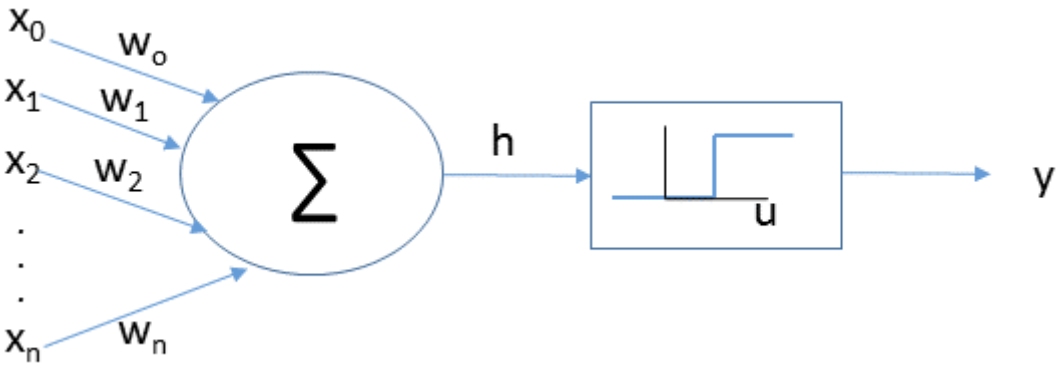


Figure 5.3: Neuron model (McCulloch and Pitts, 1943)

Jain et al. (1996) explained the behaviour behind this mathematical neuron stating that the mathematical neuron computes a weighted sum of its *n* input signals, assuming x_i, i = 1, 2, … , n be a set of independent or explanatory variables and let *y* be the nonlinear function. They are connected to a weight or strength and produce the output of "1" if the amount is higher than a known threshold W_0, and an output of "0" if otherwise. Scientifically,

$$y = g(\sum_{i=1}^{n} w_i x_i + w_0) \tag{5.1}$$

where g is a transfer function, W_i is a connection weight related to the i^{th} input. If the weight is positive, the related connection is excitatory. If negative, the connection is inhibitory.

Figure 5.4 shows four different types of activation functions which were generalised from the McCulloch-Pitts neuron approach. Jain et al. (1996) stated that the sigmoid function is a wide spread function in ANNs. Because of its asymptotic features, the sigmoid function displays smoothness. The sigmoid function is also known as the logistic function, and can be expressed by formula:

$$g(x) = \frac{1}{1+\exp(-\beta x)} \tag{5.2}$$

where β is the factor of the slope (Jain et al.,1996). Karlik & Olgac (2010) further stated that if ANN is trained by Backpropagation algorithms, using the sigmoid

function becomes very advantageous due to the fact that it is easy to distinguish, and this can improve the computation capacity for training.

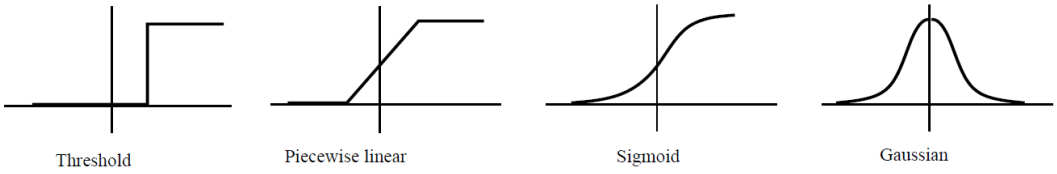


Figure 5.4: Different types of activation functions (McCulloch and Pitts, 1943)

5.3.2 Architecture/Topology of an ANN

An ANN consists of a number of neurons. How these neurons are connected to each other defines the neural network architectures type. Jain et al. (1996) grouped ANNs into two main categories based on the connection architecture, namely:

1. Feedforward Networks; and
2. Feedback (or Recurrent) Networks.

Jain et al. (1996) stated that there is no loop within Feedforward Networks as they pass the processed information forward but Feedback (or Recurrent) Networks consists of a number of loops where the network continuously feeds back the information processed. Figure 5.5 shows these two categories.

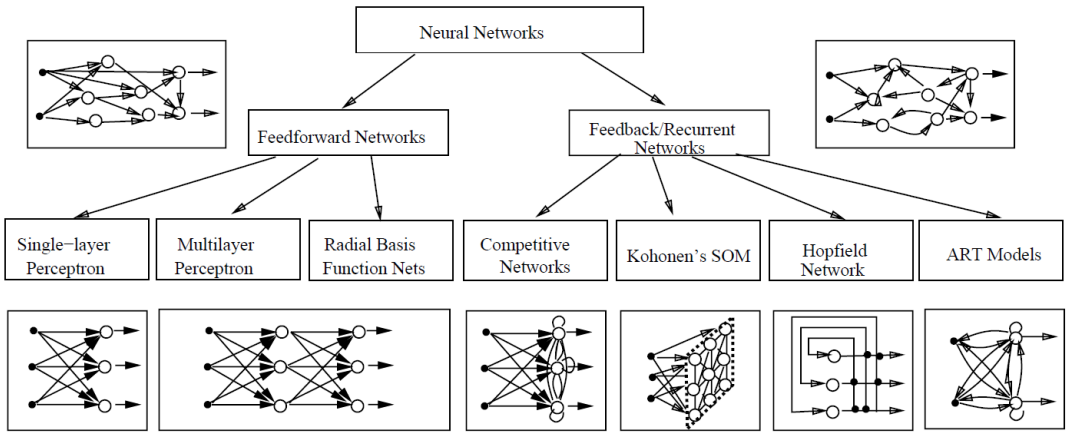


Figure 5.5: Categories of network architectures (Jain et al., 1996)

In this chapter, only Feedforward Networks are discussed because these layered networks are the most widely used network structure in time-series analysis and not only are they easy to configure and train, but they also produce very similar results compared to the most complex structures such as Feedback/Recurrent Networks.

Feedforward Networks are also known as Supervised Feedforward Networks as they fall under the Supervised Learning paradigm. It is important to understand what Supervised Learning is before discussing Feedforward Networks any further.

5.3.3 Supervised learning

Supervised Learning is a wide spread learning tool. Although there are three well-known learning paradigms that fall under Machine Learning as shown in Table 5.1, only Supervised Learning is discussed because of the previous motivation to use Feedforward Networks for the planned time-series analysis.

Table 5.1: Well-known learning algorithms (adopted from Jain et al., 1996)

Paradigm	Learning Rule	Architecture	Learning Algorithm	Task
Supervised	Error-correction	Single- or Multi-layer Perceptron	Perceptron learning algorithms Back-propagation ADALINE & MADALINE	Pattern classification Function approximation Prediction, control
	Boltzmann	Recurrent	Boltzmann learning algorithm	Pattern classification
	Hebbian	Multi-layer Feedforward	Linear discriminant analysis	Data analysis Pattern classification
	Competitive	Competitive	Learning vector quantization	Within-class categorization Data compression
		ART Network	ARTMAP	Pattern classification Within-class categorization
Unsupervised	Error-correction	Multi-layer Perceptron	Sammon's projection	Data analysis
	Hebbian	Feedforward or Competitive	Principal component analysis	Data analysis Data compression
		Hopfield Net	Associative memory learning	Associative memory
	Competitive	Competitive	Vector quantization	Categorization Data analysis
		Kohonen SOM	Kohonen SOM	Categorization Data analysis
		ART networks	ART1, ART2	Categorization
Hybrid	Error-correction & Competitive	RBF network	RBF learning algorithm	Pattern classification Function approximation Prediction, control

The idea behind Supervised Learning is to sample a specific distribution where input and the expected output is provided. This is achieved by calculating the weights in such a way that the output can be similar to the known correct answers. According to

Jain et al. (1996, p. 20), there are four kinds of learning rules as shown in Table 5.1 that falls under Supervised Learning, namely:

1. "Error-correction;
2. Boltzmann;
3. Hebbian; and
4. Competitive learning".

Error-correction rule is one of the most widely used methodology that falls under Supervised learning. Rosenblatt (1958) developed the perceptron convergence theorem using the error-correction rule to advance his ideas. In this chapter only error-correction rule Supervised Learning is discussed as it was used to predict models discussed later in Chapter 6.

5.3.4 Error-correction rule

As previously mentioned, Supervised Learning involves sampling a specific distribution where input and the expected output is provided. While training the network it can be noticed that the expected output *d* usually differs from the actual output, *y*. Due to this difference, the error-correction rule aims to reduce the difference between the error signals (*d* - *y*). This is achieved by adjusting the connection weights. As explained in Chapter 4, $u_{i,l}^{(m)}$ is the utilisation of the software that performs task on commodity (*i*) and functionality (*l*) by using software (*m*) with time factor (*t*) and

$$u_{i,l}^{(m,t)} \in [0,1]$$

Furthermore, the aim is to predict future software utilisation by performing time-series analysis and

$$\hat{u}_{i,l}^{(m,t)}$$

is the predicted software utilisation which will be discussed further in the next chapter, Chapter 6 where

$$d = \hat{u}_{i,l}^{(m,t)} \tag{5.3}$$

and

$$y = u_{i,l}^{(m)} \tag{5.4}$$

based on the notation used.

One of the important facts about error-correction rules is that the perceptron learning rule (Rosenblatt, 1958) is built on this logic. A perceptron made of a single neuron which has adjustable weights, W_i; i = 1, 2, …, n, and threshold W_o, was shown in Figure 5.3.

The input vector x = $(x_1, x_2, \ldots, x_n)^t$, the net input to the neuron is (Jain et al., 1996):

$$v = \sum_{i=1}^{n} w_i x_i - w_o \tag{5.5}$$

“The output y of the perceptron is +1 if v > 0, and 0 otherwise. In a two-class classification problem, the perceptron assigns an input pattern to one class if y = 1, and to the other class if y = 0” (Jain, et al., 1996, p. 21). “The linear equation

$$\sum_{i=1}^{n} w_i x_i - w_0 = 0 \tag{5.6}$$

defines the decision boundary which divides the space into two halves” (Jain, et al., 1996, p. 21).

To be able to determine the weights and threshold in a perceptron, Rosenblatt (1958) developed learning procedures using a set of training patterns. The following procedure describes this learning method:

1. “Initialize the weights and threshold to small random numbers.
2. Present a pattern vector(x_1, x_2, …, x_n)^t, and evaluate the output of the neuron.
3. Update the weights according to

$$W_i(t+1) = w_i(t)+\eta(d-y)x_i \quad (5.7)$$

where d is the desired output, t is iteration number, and η is the learning rate whereby η ($0.0 < \eta < 1.0$) is the gain (step size)” (Jain, et al., 1996, p. 21).

Learning only happens if the perceptron caused an error. Rosenblatt (1958) proved that famous perceptron convergence theorem works. According to Rosenblatt (1958) “if the training patterns are drawn from two linearly-separable classes, then the perceptron learning procedure will converge after a finite number of iterations” (Jain, et al., 1996, p. 21). However, during training, it is not known if the patterns are linearly separable. Nevertheless, a single layer perceptron can only separate linear patterns when it is used in conjunction with a monotonic activation function such as a Gaussian function, which could form non-linear decision boundaries. The back-propagation algorithm is also built on this code. Error-correction rule is widely used for training Feedforward Networks (Jain, et al., 1996). Learning stops when the error nearly stabilises.

5.4 Feedforward Networks

Feedforward Networks are named after their operational character which continuously feeds the information forward. As it consist of layers, it is also known as layered network. These three layers are the input, output, and hidden layer(s). The neurons are organised into layers where each neuron in one layer is connected to all the neurons of the following layer. Each of these connections shows different network behaviours. However, Feedforward Networks only create one set of output values for every given input value, not a sequence of values. Furthermore, the reply to a Feedforward Network input does not depend on the previous state of the network

(Jain et al., 1996). Karimi, et al. (2010) further stated that Feedforward Networks gained popularity within ANN due to the fact that they are very efficient to solve complex problems. Although there are three different classes of network architectures which fall under Feedforward Networks as is shown in Figure 5.5, only the first two are discussed due to the fact that Multilayer Feedforward Networks can sufficiently produce desired outcomes in terms of time-series analysis.

5.4.1 Single-layer Feedforward Networks

Single-layer Feedforward Network is a very basic network which contains the input and output layer. This network is only capable of handling a linear regression model. This limitation i.e. it can only separate linearly separable patterns, led to the development of Multilayer Feedforward Networks.

5.4.2 Multilayer Feedforward Networks

Multilayer Feedforward Networks are the upgraded version of Single-layer Network and can handle both linear and non-linear models. Compared to Single-layer Network, it has an additional feature called hidden layers (neurons, also known as units) where all the neurons between layers are connected to each other as shown in Figure 5.6. Minsky & Papert (1969) recognised that Multilayer Feedforward Networks are able to establish arbitrarily complex decision boundaries.

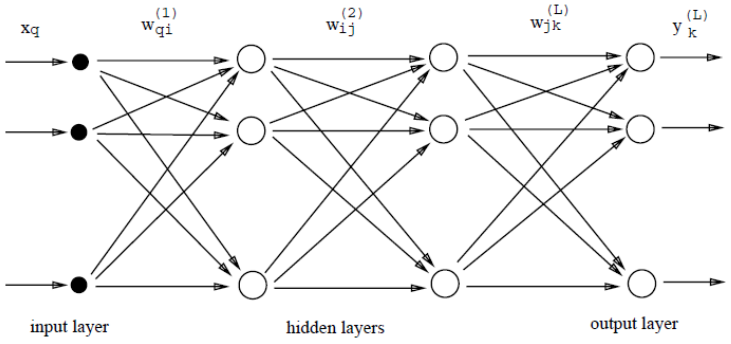


Figure 5.6: Feedforward Network architecture with three-layers (Jain et al., 1996)

Figure 5.6 shows a typical Feedforward Network architecture with three-layers. It needs to be noted that, in this network, the layers are connected to each other in such

a way that, there are no feedback connections between layers, also neurons in the same layer are not connected to each other.

Multi-Layer perceptrons (MLP) is a well-known class which was originated from Feedforward Networks where each unit uses either the sigmoid or the thresholding function. This type of network is generally trained with the algorithm called Backpropagation learning algorithm, which was developed by Rumelhart and McClelland (1986). This development helped to determine weights in a MLP and made Backpropagation learning algorithm more wide spread among ANNs users. As illustrated in Figure 5.6, w_{ij} shows the weight on connection between two units i.e. from the i^{th} unit in layer (l - 1) to j^{th} unit in layer l.

Assume that $\{(\mathbf{x}^{(1)},\mathbf{d}^{(1)}), (\mathbf{x}^{(2)},\mathbf{d}^{(2)}),\ldots,(\mathbf{x}^{(p)},\mathbf{d}^{(p)})\}$ given as a number of p training designs, where the input vector is defined as $\mathbf{x}^{(i)} \in R^n$ in the n-dimensional pattern space, and $\mathbf{d}^{(i)} \in [0,1]^m$ in the m-dimensional hyper-cube. Here m can be classified as the number of classes. The squared-error cost function, which is very often used in the ANN literature, is defined as (Jain, et al., 1996):

$$E = \frac{1}{2}\sum_{i=1}^{p} \| y^{(i)} - d^{(i)} \|^2 \tag{5.8}$$

The Backpropagation algorithm (Rumelhart, et al., 1986) is given below. It is often used to lessen the squared-error cost function in Eq. 5.6.

1. "Initialize the weights to small random values;
2. Randomly choose an input pattern $x^{(\mu)}$
3. Propagate the signal forward through the network;
4. Compute δ_i^L in the output layer ($o_i = y_i^L$)

$$\delta_i^L = g(h_i^L)[d_i^\mu - y_i^L], \tag{5.9}$$

where h_i^l represents the net input to the i^{th} unit in the l^{th} layer.

5. Compute the deltas for the preceding layers by propagating the errors back-wards;

$$\delta_i^L = g(h_i^l)\sum_j w_{ij}^{l+1}\delta_j^{l+1}, \tag{5.10}$$

for $l=(L-1), \ldots, 1$.

6. Update weights using

$$\Delta\omega_{ji}^l = \eta\delta_i^l y_j^{l-1} \tag{5.11}$$

7. Go to step 2 and repeat for the next pattern until the error in the output layer is below a pre-specified threshold or a maximum number of iterations is reached" (Jain, et al., 1996, p. 31)

In the Backpropagation algorithm, $g(.)$ is the unknown activation function and is approximated by function approximation techniques. Furthermore, as mentioned previously, *d* and *y* represent target and actual output, respectively and η is a small constant called learning rate which will be explained in Chapter 6.

To explain the role of hidden units, Jain et al. (1996) produced a diagram showing a geometric interpretation of the hidden units with the help of the threshold activation function in Figure 5.7. As hyper-planes can approximate boundaries between pattern classes, in the pattern space, each unit in the first hidden layer forms a hyper-plane. Consequently, using the outputs of the first layer units, a unit in the second hidden layer forms a hyper-region; by performing an "AND" operation on hyper-planes, and then a decision region is obtained. Furthermore, by performing logical "OR" operations, using the units in the second hidden layer, the output-layer units combine

the decision regions. This explanation aims to express the role of hidden units, but after training they could behave differently. Contrary to what is shown in Figure 5.7, it is actually possible to form complex decision boundaries using a two-layer network. Moreover, MPL can form smooth decision boundaries with sigmoid activation functions (Jain, et al., 1996).

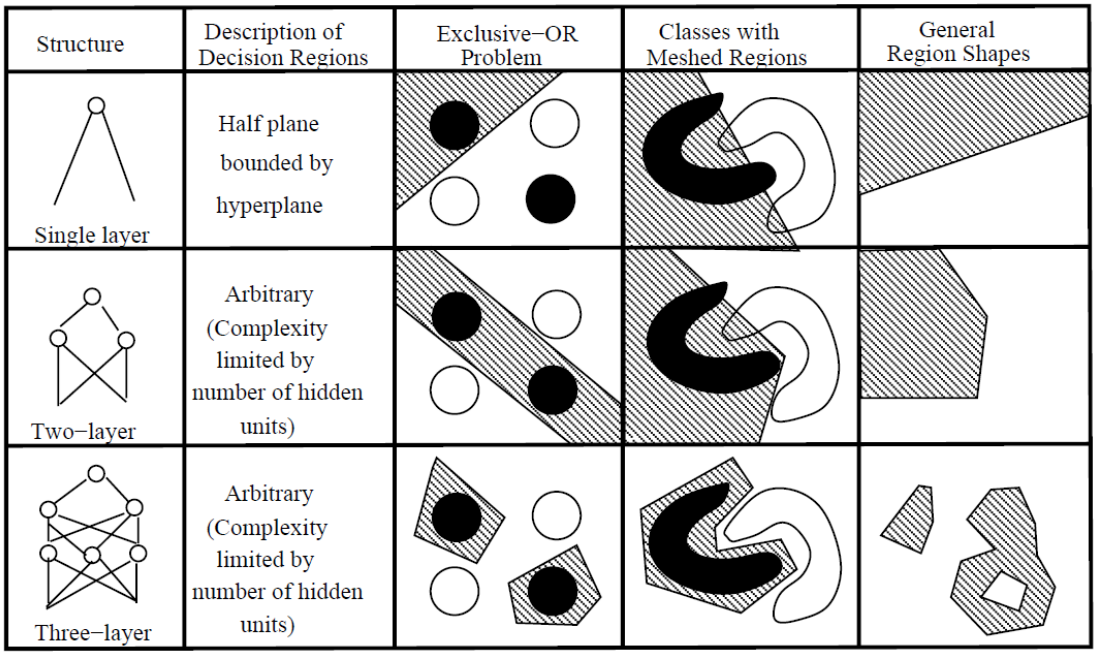


Figure 5.7: The role of hidden units (Jain et al., 1996)

Jain et al. (1996) listed some of the important factors to consider when designing Feedforward Networks, these are:

1. Sum of required layers for the task;
2. How many units per layer;
3. How is generalisation going to be performed on data outside training set; and
4. Size of training set be for “good” generalisation.

Backpropagation is one of the most widely used learning algorithm in time-series analysis that falls under Supervised Learning which is used to find weights in Multilayer Feedforward Networks. It is based on a gradient descent algorithm and it is often used for modelling of ANNs. Although as stated by Hertz, et al. (1991), the Backpropagation learning algorithm has been extensively used for pattern

classification, function approximation and forecasting, but many parameters need to be configured by learning from the mistakes. Current literature can only provide some guidelines for helping to find the correct parameters in practice (Jain et al., 1996). Due to these reasons, in Chapter 6, the Backpropagation learning algorithm is used to train the mine planning software dataset for the time-series analysis.

5.5 Training/Learning of ANNs

Training/Learning is a very important part of ANNs as the level of success depends on how well the dataset is trained. Training on the mine planning software dataset is done using the Backpropagation algorithm, which attempts to reduce the error in the network by adjusting the weights and is based on the error-correction rule. The objective of training is to determine the set of weights and thresholds such that for an input signal the ANN's output is the closest to the expected output. To start the training, first the initial weights are chosen randomly and then the training begins. It is assumed that the neural network has no prior knowledge about the problem before being trained which is a feature of Feedforward Networks as mentioned in Section 5.4.

The ANN's training/learning section is usually divided into three different categories namely; training, validation and testing. However, because of the temporal nature of the mine planning software database, the dataset used in time-series analysis is simulated data, hence the database does not require to be validated. Consequently, the dataset is split into two main parts:

1. a training set, to decide on the weights ;
2. a testing data set, independent sets of data, not used in training or validation and are used to verify the effectiveness of the stopping criterion and to estimate the expected performance of the ANNs.

While training the dataset, the predicted output is compared with the actual (expected) output and the root-mean-squared-error between the two are calculated. The root-mean-square error, also known as root-mean-square deviation is a widely used method in statistics to measure the differences between predicted values by an ANN model and the real observed values. The root-mean-square error function, which is a function of the weights in the network, is minimised. This is done efficiently by using

Backpropagation learning algorithm. This approach continues until an acceptable weight combination is found and the prediction error of the testing data is minimised. When this stage is reached, the ANN is considered trained. Supervised training, which is common in time-series analysis was used to train the mine planning software dataset.

In summary, the choice of ANN over the traditional methods is actually case specific. Depending on the structure of the data as shown in Figure 5.4 (there might be linear or nonlinear structures) a predictor amongst linear or nonlinear regression of different types can be chosen. However, it is difficult and time consuming to design a regressor which can perform prediction under a different data structure. ANN is a remedy for this problem as its generic architecture is able to replicate the performance of any type of regressor. Due to this reason, ANN is used as a choice of predictor to perform time-series analysis in this study.

5.6 Summary

Following the establishment of the software utilisation framework in Chapter 4, in this chapter ANNs were defined, the reasons behind choosing supervised learning paradigm with supervised Multilayer Feedforward Networks architecture were motivated. Furthermore, using pattern classification, function approximation and prediction/forecasting tools, it was decided to use the error-correction learning rule with Backpropagation learning algorithm to the dataset. Consequently, this background information has established the foundation to start time-series analysis. The next chapter, Chapter 6, outlines how the time-series analysis performed.

6 TIME-SERIES ANALYSIS WITH ARTIFICIAL NEURAL NETWORKS

6.1 Introduction

In the previous chapter, ANNs were defined, the reasons behind choosing supervised learning paradigm with supervised Multilayer Feedforward Networks architecture to conduct time-series analysis were motivated. In this analysis, the Feedforward Neural Networks were trained to approximate the mathematical functions generating the semi-chaotic time-series which represents the history of mine planning software utilisation indices in order to predict the values for the future. Furthermore, an ANN model is used to forecast a time-series. It is shown that the error-correction learning rule with Backpropagation learning algorithm method can be used to calculate standard errors and prediction limits for the forecasts.

As stated previously, the existing database, which consists of only two different time-stamps, will be continuously evolving, meaning that it is expected to grow further in time. Due to this reason, it is necessary to establish a framework so that it can be used for further analysis of the evolving database. In general, time-series data is usually collected on a daily, monthly or annual basis so that there is sufficient data available for analysis. Due to the temporal nature of the database, which consists of only two different time-stamps, it is necessary to create a simulated dataset so that time-series analysis could be conducted.

6.2 What is a time-series?

A time-series, also known as regular collection of a statistical dataset, can be found easily in many different applications such as engineering, medicine and finance. Feedforward Neural Networks is a well-known tool used for time-series analysis. Furthermore, there are many fields of application for the multi-layered Feedforward Neural Network, like classification, pattern recognition and forecasting. The following example illustrates a time-series prediction problem.

Let y_t, y_{t-1}, y_{t-2} ... be a stationary time-series and let y_{t+1} be the value to be predicted.

Then a set of *d* such values $y_t, y_{t-1}, y_{t-2}, \ldots y_{t-d+1}$ can be selected from the time-series

be the inputs to a Feedforward Neural Network, and the future value y_{t+1} used as the target for the output. By using a similar approach, mine planning software utilisation can be predicted from the following time steps, i.e.

$$u_{i,l}^{(m,(t+1))} = g\left(u_{i,l}^{(m,t)}, u_{i,l}^{(m,t-1)}, u_{i,l}^{(m,t-2)}, ..., u_{i,l}^{(m,t-T_f)}\right) \quad (6.1)$$

where T_f is the last step in the time-series and $g(.)$ is the unknown function which is approximated by the help of machine learning techniques, particularly using Feedforward Neural Networks. The following explains further time lags:

t+1 is the future time which is the desired future prediction time
t is the current time,
t-1 is the past time with lag as 1st step, and
t-2 is the past time lag as 2nd step of the analysis.
.
.
.
t-T is the time lag as Tth

The task is to predict $u_{i,l}^{m,t+1}$ or $u_{i,l}^{m,t+T_f}$ and *k*={*i* ,*l*}, where T_f is a desired future time step, from present and past samples such as $u_{i,l}^{(m,t)}, u_{i,l}^{(m,t-1)}, u_{i,l}^{(m,t-2)}, ..., u_{i,l}^{(m,t-T_p)}$ where T_p is the final time-stamp considered in prediction.

6.3 Time-series analysis using Backpropagation Feedforward Neural Networks

In this section, the performance of Backpropagation Neural Networks applied to the problem of predicting mine planning software utilisation index. The Feedforward Neural Networks were trained to approximate the mathematical functions generating the semi-chaotic time-series which represents the history of software utilisation indices in order to predict the values for the future. Since the evolving database is limited in size, where there are currently only two time-stamps, which makes it impossible to model the future values accurately. Due to this limitation without losing generality, the data used in time-series analysis is simulated data. However, the

prediction quality obtained is illustrated by presenting several simulation results. Furthermore, the presented results could be used as an indicator that this framework can be applied on a real database when the actual database is populated with sufficient time-stamps in the future.

For a Feedforward Neural Network, the derivatives of a root-mean-square error (RMSE) function with respect to the weights can be obtained efficiently by using Backpropagation. There are four main stationary points namely A, B, C, and D at which the local gradient of an error function can possibly disappear. Figure 6.1 illustrates this point (D) by showing the value of an error function E against a single weight *w*:

- Point A is called a local minimum;
- Point B is a local maximum;
- Point C is a flat region where most error functions end early;
- Point D is the global minimum, because the value of the error function is minimal at this point.

In order to avoid over-training of the dataset, the test data is used during the training process. With each iteration, RMSE of both the training and the testing sets are calculated and compared. The RMSE of the training set will typically decrease, and after a number of iterations, stabilise. The RMSE as a measure of performance can be determined by calculating the sum of squares of the error function. During this exercise, RMSE function should stop when it reaches the Point D where the value showing the error function is minimal (Figure 6.1). The remaining data is used to test the network's performance after training.

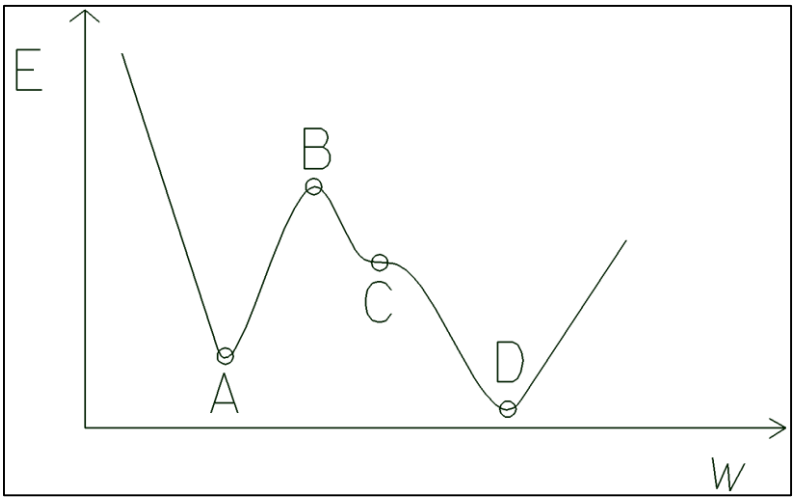


Figure 6.1: An error function with four stationary points (Bishop, 1996)

6.3.1 Data simulation

As previously stated, due to the lack of available data in the database, it was decided to simulate the dataset. In the remainder of this chapter, the notation $u_k^{(t)}$is interchangeably used in place of $u_k^{(m,t)}$ when *m* is fixed. The data $u_k^{(t)}$ is simulated according to the following formulation:

$$a_k^{(t)} = \sum_{m=1}^{M}(-1)^{\delta_{u_{m,1},u_{m,2}}}\, e^{-\left(t-u_{m,3}\right)^2/2\left(s+u_{m,4}\right)^2} \tag{6.2}$$

$$b_k^{(t)} = \frac{a_k^{(t)}-a_k^{min}}{a_k^{max}-a_k^{min}} \tag{6.3}$$

$$u_k^{(t)} = \frac{b_k^{(t)}}{\sum_{l=1}^{K} b_l^{(t)}} \tag{6.4}$$

where $u_{m,1}$, $u_{m,2}$, $u_{m,3}$, $u_{m,4}$ are uniformly distributed random variables in the closed interval of $[0,1]$, t is a discrete time index defined in the unit interval of $[0,1]$ and sampled uniformly with sampling rate of t_s which represents the period of the data acquisition rate, a_k^{min} and a_k^{max} are, respectively, the minimum and maximum values of the sequence $a_k^{(t)}$, s is a constant, K is the number software performing a specific commodity and functionality.

$$\delta_{u_{m,1},u_{m,2}} = \begin{cases} 1, u_{m,1} > u_{m,2} \\ 0, \text{otherwise} \end{cases} \tag{6.5}$$

Different realisations of the simulated dataset can be created by changing the values of variables M, K, k, initial seed value of random number generator and t_s; M defines the number of Gaussian distribution (also known as "bell-shaped curve", see Figure 5.4) with random parameters used in simulated data, and t_s is used to simulate the period of data collection (the shorter the period, the smoother the simulated data is). Different realisations of data simulation are shown in Figure 6.2, Figure 6.3 and Figure 6.4. The impact of sampling rate and Gaussian distribution to the data simulation process is shown in Figure 6.2, Figure 6.3 and Figure 6.4 by changing only two variables namely M and t_s, respectively; t_s sampling rate usually represents the frequency of sampling rate, i.e. daily, weekly, month, yearly, so on.

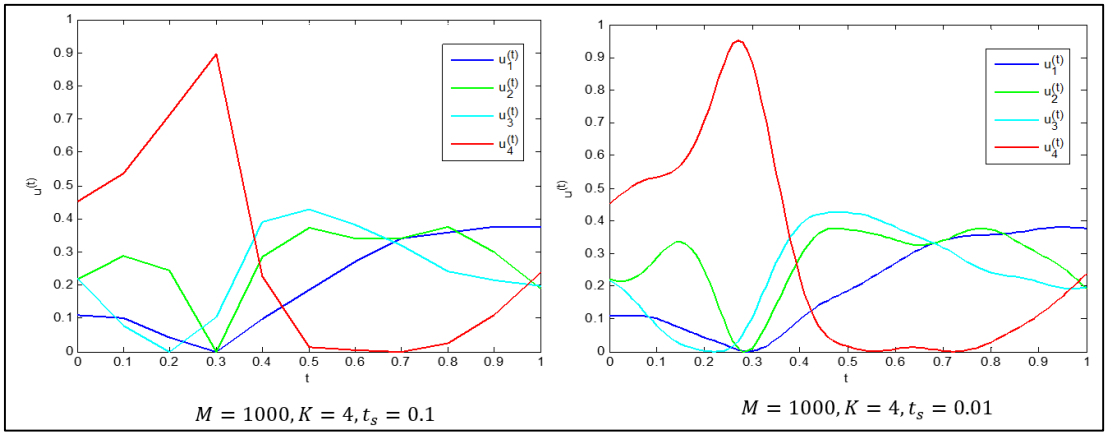


Figure 6.2: Impact of sampling rate to the data simulation process ($M = 1000$, $K = 4$ with variable t_s)

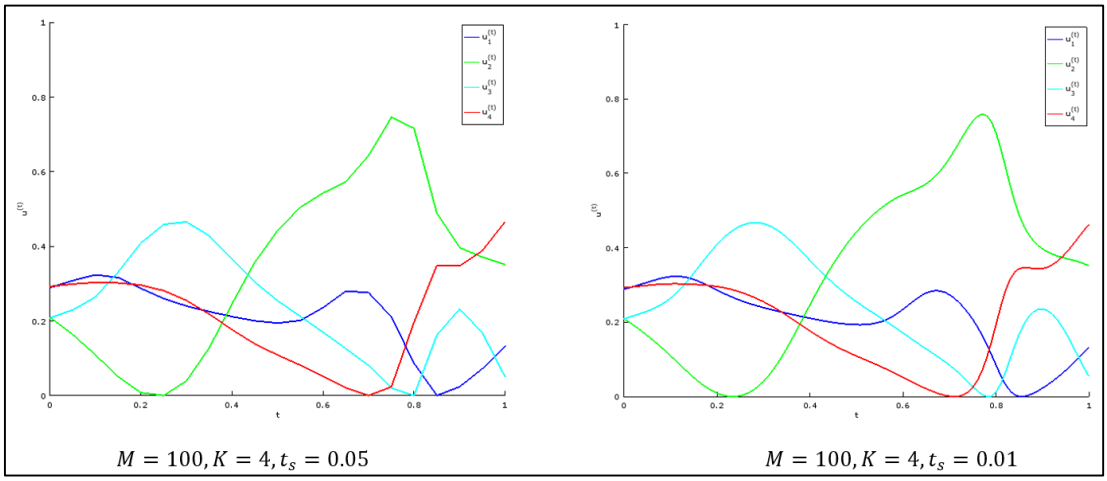


Figure 6.3: Impact of sampling rate to the data simulation process ($M = 100$, $K = 4$ with variable t_s)

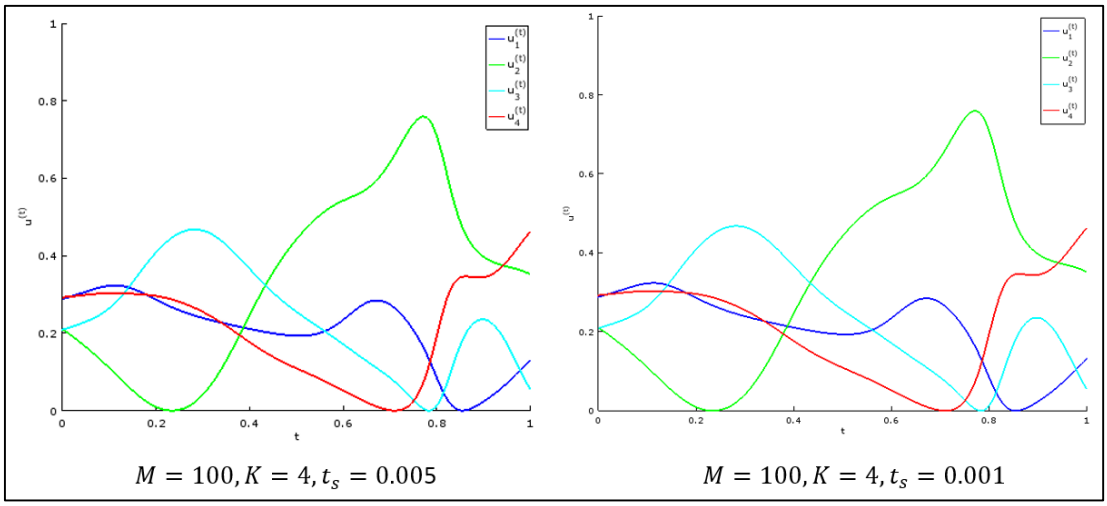


Figure 6.4: Additional sampling rate examples ($M = 100,\ K = 4$ with variable t_s)

Furthermore, Figure 6.2, Figure 6.3 and Figure 6.4 as K=4, $u_1^{(t)}$, $u_2^{(t)}$, $u_3^{(t)}$ and $u_4^{(t)}$ illustrate the level of mine planning software utilisation for four different software. Each of these software utilisation indicate different levels of market condition, i.e. software utilisation $u_4^{(t)}$ is the best performing showing the best performance in the market, followed by $u_2^{(t)}$. In the remainder of this chapter, the dataset simulated by using the following values: $M = 100,\ K = 4$ with variable t_s.

6.3.2 Data validation

The data validation process was discussed in detail in Chapter 5.5. What needs to be remembered regarding data validation is that, due to a lack of available data, it was decided to simulate the dataset. As this is not only an actual dataset, but also very small in size, it is not required to be validated. Hence, no data validation was conducted.

6.3.3 Performance evaluation

The performance of prediction for each time-series $u_k^{(t)}$ is measured according to RMSE, i.e.

$$\text{RMSE} = \sqrt{\sum_t \left(\hat{u}_k^{(t)} - u_k^{(t)}\right)^2} \tag{6.6}$$

where $\hat{u}_k^{(t)}$ is the predicted time-series from Feedforward Neural Networks. As previously mentioned, RMSE, also known as root-mean-square deviation is widely used method in statistics to measure the differences between predicted values by an ANN model and the real observed values. Consequently, the lower the RMSE, the better the predictions.

6.4 Backpropagating Feedforward Neural Network

As previously mentioned, Backpropagation ANNs contain a collection of inputs and processing units known as neurons. As shown in Figure 6.5, neurons are interconnected by weights (shown as W in Figure 6.5) as they carry information of a trained network. Furthermore the hidden and output layers are connected to a bias neuron which has a "+1" value. An ANN with one hidden layer with an adequate number of hidden neurons can approximate the sigmoid function. However, ANNs with one or two hidden layers are known to achieve a good outcome. Increasing the number of hidden layers can cause undesirable results such as over-fitting which leads to a poor network performance and it also increases the computation time.

Considering these observations, the following architecture was used as shown in Figure 6.5, where the output layer performs the forecasting of software utilisation $\hat{u}_k^{(t+1)}$ (T_f). Performance analysis for a different number of nodes on input and hidden layers is conducted to see the effect of the number of nodes on function approximation and/or prediction. The connections are only feedforwarded between the adjacent layers. For the transfer function, Backpropagation ANN employs the sigmoid function as mentioned previously.

The learning algorithm that controls Backpropagation is given below in a simple breakdown of the following five steps:

Step 1 - Start-up: Prepare connection weights and neurons to small random values.
Step 2 - Data feeding: Start the continuous data set of inputs and outputs to the backpropagation ANN.
Step 3 - Management of outputs: Calculate the outputs, and adjust the connection weights a few times by applying the RMSE.

Step 4 - Variations: Amend the activation thresholds and weights of the neurons in the hidden layer according to the Delta rule.
Step 5 - Repeat ***Step 2*** to ***Step 4*** to converge to the minimum error.

Furthermore, data was divided into two independent data sets namely training set and test set. The training set data was observed during training. The *testing data* in this section, refers to future predictions from ANN. Backpropagation Feedforward Neural Network was selected as the best network by identifying the minimal error found using the test data (Figure 6.5). In Figure 6.5, $w_{h,i}^{l,m}$ is a weight of the link which connects node "*m*" in input layer to node "*l*" in the hidden layer, and similarly $w_{o,h}^{l}$ is a weight of the link that connects node "*l*" in the hidden layer to the output node. Furthermore, nodes with index "0" on each layer are bias nodes that are all set to "+1". The number of input nodes defines the number of past points used to predict the future values which is also parametrical as T_p. The number of nodes at the hidden layer is selected according to $n_h=2n_i$ which is empirically determined.

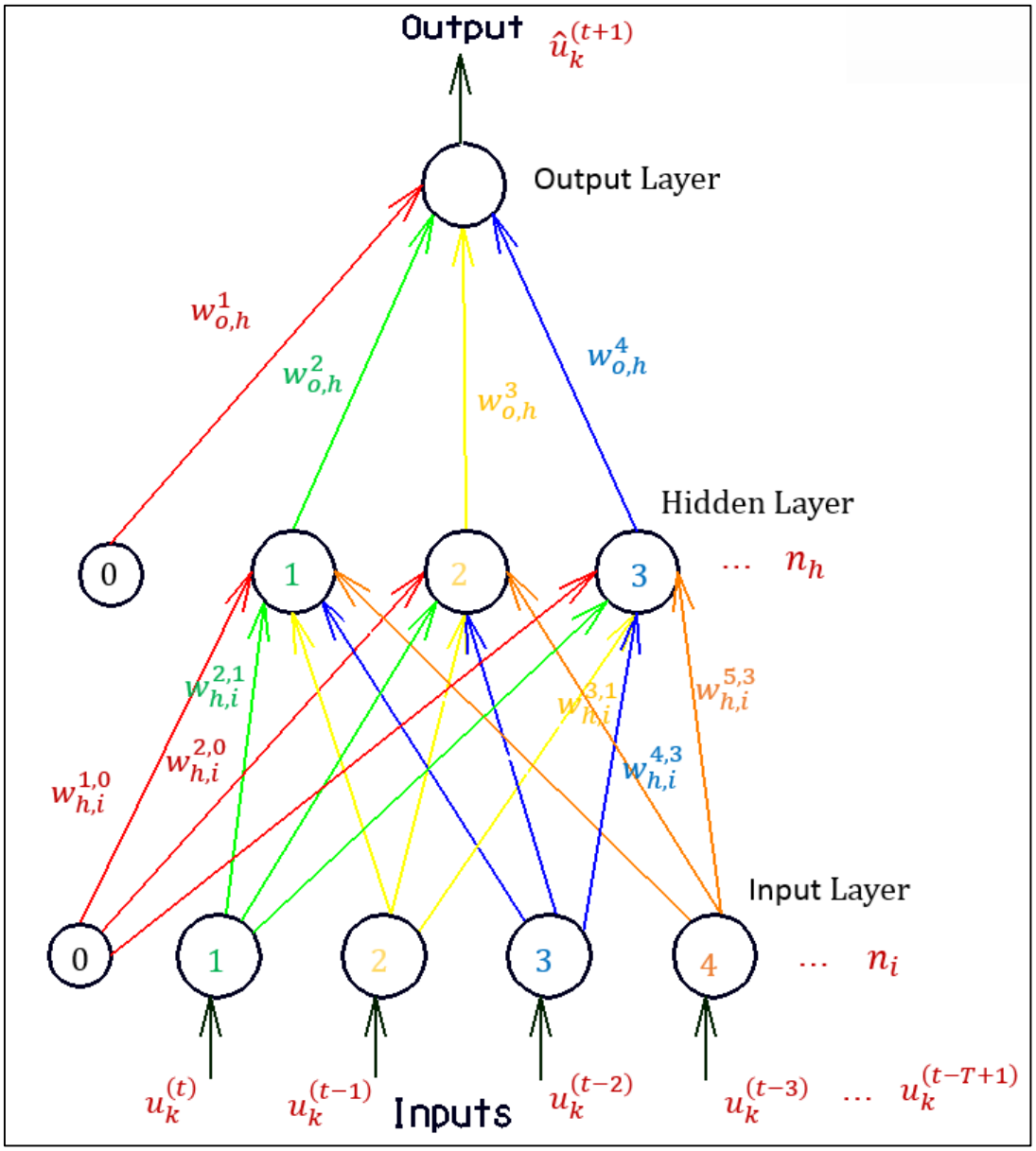


Figure 6.5: Architecture of Feedforward Neural Network considered in utilisation prediction

Learning of a neural network model takes place during training by means of the training algorithm. At the training stage, T_p samples were used of which the most recent value $u_k^{(t)}$ is predicted using the rest of data points as input feature to ANN. In other words, weights $w_{h,i}^{l,m}$ s and $w_{o,h}^{l,m}$ s are learned using $u_k^{(t)}$ as the target and

$$u_k^{(t-1)},\ u_k^{(t-2)}, u_k^{(t-3)}, \dots, u_k^{(t-Tp)}$$

as input features to ANN. Figure 6.6 shows how training is performed. In Figure 6.6, the input feature vector to ANN is $\overline{x}$

and ($\overline{x}$, y) →

$$\overline{x} = (u_k^{(t-1)}, u_k^{(t-2)}, u_k^{(t-3)}, \dots, u_k^{(t-Tp)}) \tag{6.7}$$

and

$$y = u_k^{(t)} \tag{6.8}$$

y, which is a target value that ANN tries to predict, and the task of learning algorithm is to minimise the square of error between output of ANN $\hat{u}_k^{(t)}$ and the real value of $u_k^{(t)}$.

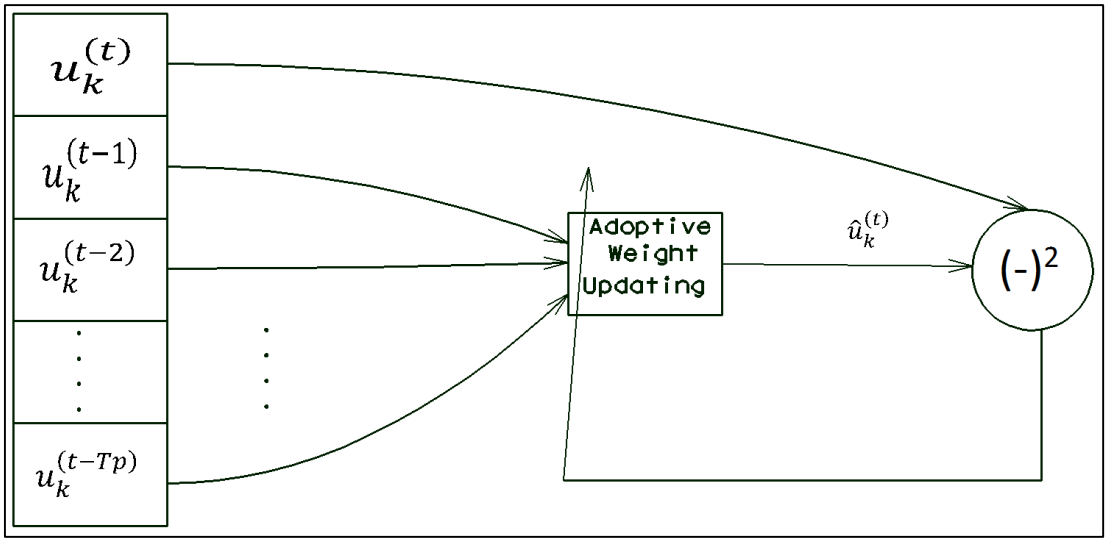


Figure 6.6: Training of Feedforward Neural Network

The iterative formulas given in Eq. 6.7 through Eq. 6.8 are derived using Backpropagation algorithm. The iterations are stopped when the maximum of the iterative differences between consecutive values of weights is smaller than a predefined threshold value. Figure 6.7 shows how prediction was done.

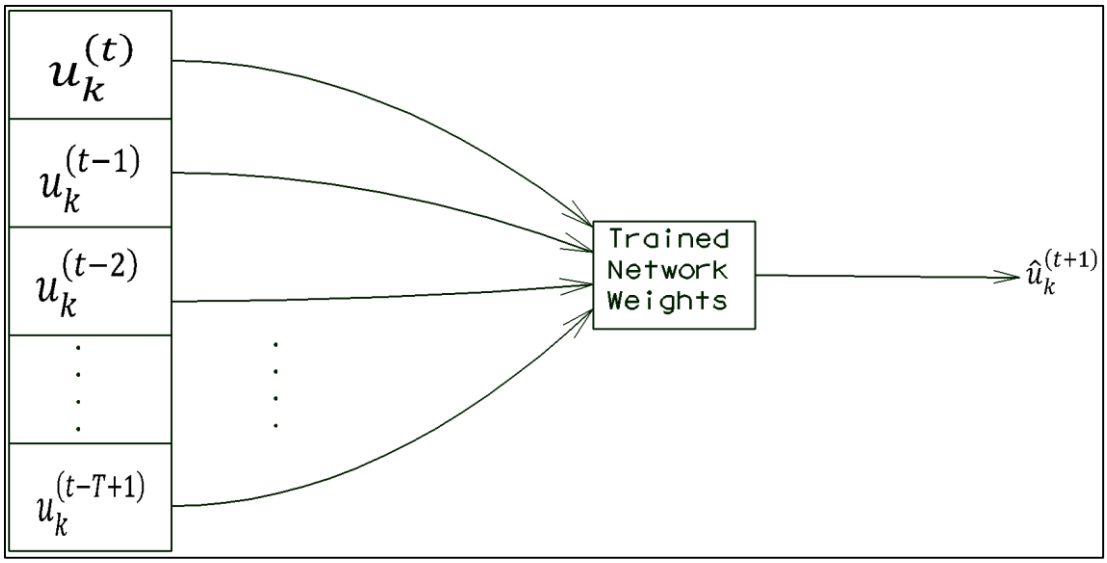


Figure 6.7: Obtaining software utilisation prediction of Feedforward Neural Network

6.5 Predicting software utilisation lags

Predicting future values using ANN is a crucial part of this chapter. Figure 6.8, Figure 6.9 and Figure 6.10 illustrate how prediction is performed. Figure 6.8 shows the learning of the Feedforward Neural Network (FNN) to predict $\hat{u}_k^{(t+1)}$software utilisation. FNN is a well-known ANN architecture which falls under supervise learning and it utilises the Backpropagation learning algorithm. Furthermore, through this learning, the network weights are also learned. Afterwards, Figure 6.9 shows how $\hat{u}_k^{(t+2)}$software utilisation is predicted. In this prediction, using the learned weights and obtained $\hat{u}_k^{(t+1)}$, software utilisation is used to predict the desired $\hat{u}_k^{(t+2)}$. In Figure 6.10, by using a similar approach of predicting $\hat{u}_k^{(t+2)}$, this time predicted $\hat{u}_k^{(t+2)}$ was fed into ANN, and then the next $\hat{u}_k^{(t+3)}$software utilisation prediction was obtained. Furthermore, by using this approach, the other software utilisation predictions ($\hat{u}_k^{(t+4)}, \hat{u}_k^{(t+5)}, \ldots, \hat{u}_k^{(t+Tf)}$) could be obtained. After configuring the correct combinations of the variables, finally, by running the forecast scripts, simulation results were generated. The next section discusses these results.

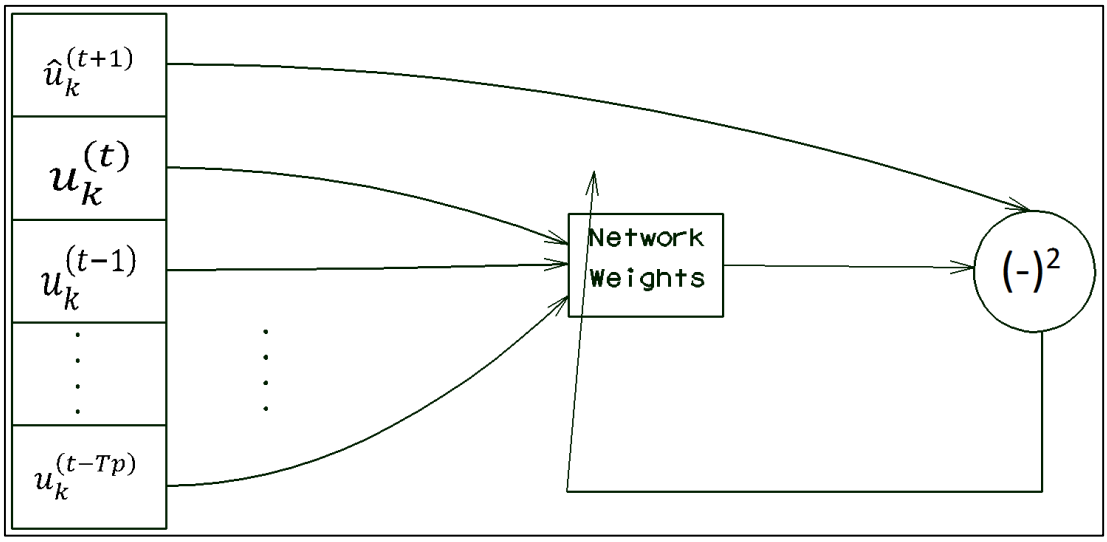


Figure 6.8: Training FNN to predict $\hat{u}_k^{(t+1)}$ software utilisation

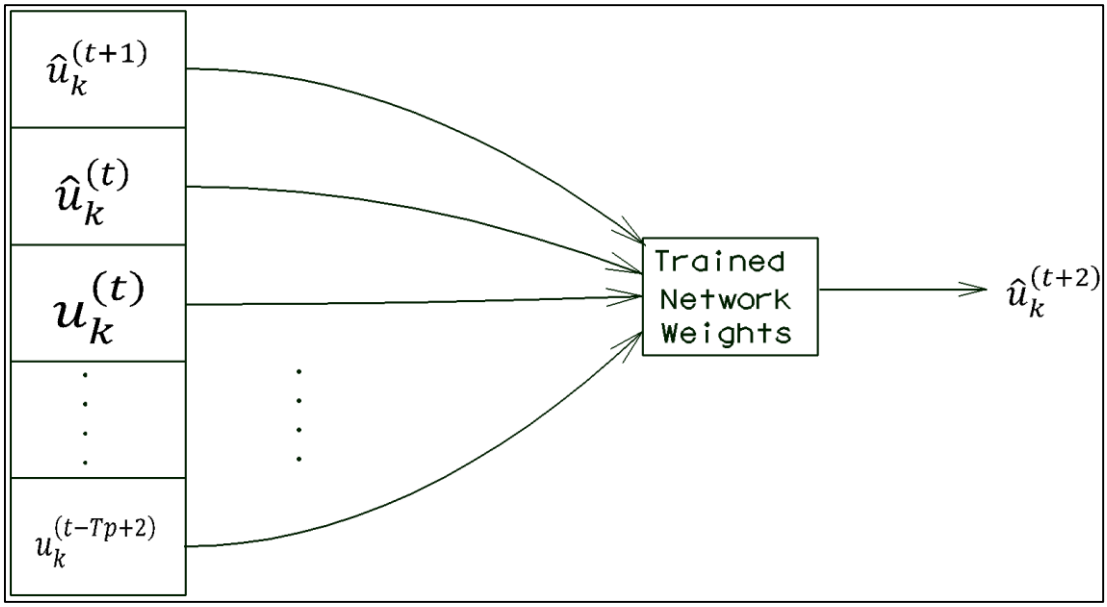


Figure 6.9: Predicting $\hat{u}_k^{(t+2)}$ software utilisation

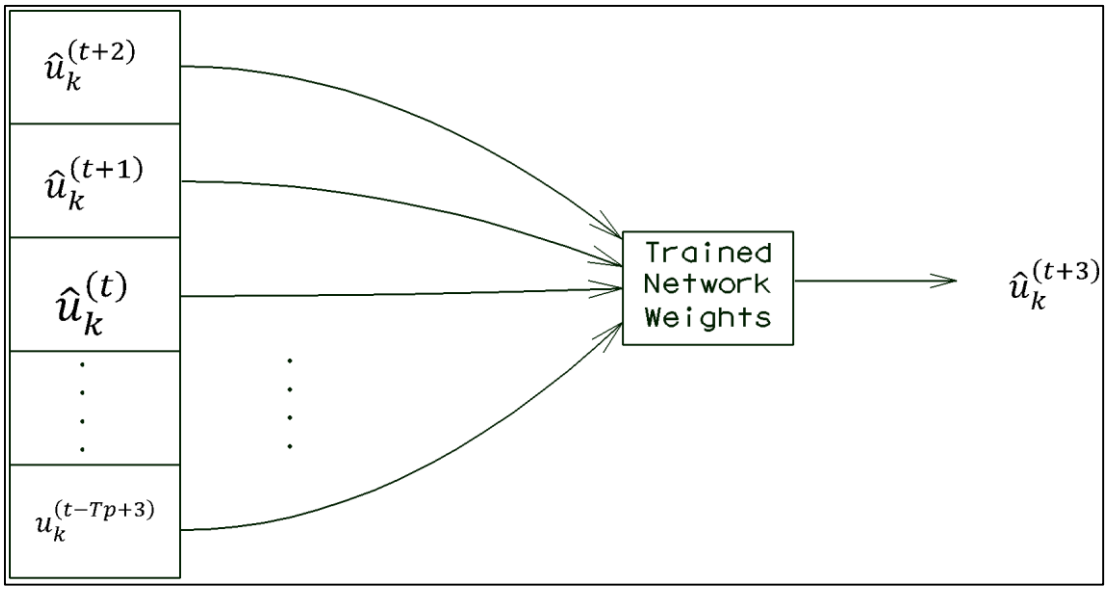


Figure 6.10: Predicting $\hat{u}_k^{(t+3)}$ software utilisation

6.6 Simulation results

A number of Octave scripts were written to model the dataset. These include scripts for Backpropagation algorithm, Feedforward ANN, forecasting, simulations, testing and training, respectively. However, the scripts are designed in a way that they all complement each other. The script details used to predict mine planning software utilisation time-series analysis are given in Appendix 9.3.

During dataset simulation, the *M*, *K* and t_s values are kept constant ($M = 100,\ K = 4$ and $t_s = 0.01$). Figure 6.11 shows how the RMSE values changes with the change of the T_f value, *k* which is an index, represents four different software utilisations. For each data point, its past T_f values are used together to train FNN, the then trained FNN is used to predict the very next feature, i.e. $\hat{u}_k^{(t+1)}$. Furthermore, Figure 6.11 indicates that up to window four, the prediction has a reasonable confidence level which is the highest (low RMSE), and the lowest (high RMSE) when T_f =10.

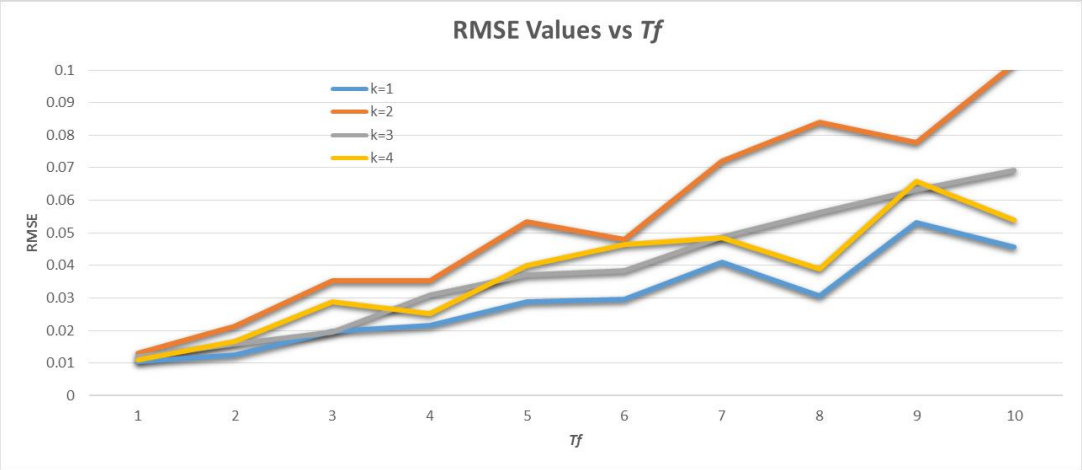


Figure 6.11: How RMSE values change with T_f

Consequently, Figure 6.12 shows how prediction results stand against the actual values shown. A similar outcome could be observed if two figures, Figure 6.11 and Figure 6.12 were compared. Figure 6.12 shows that the predicted (t+1) results are the closest compared to the Actual values. Furthermore, Figure 6.12 indicates that the predictions through 1 to 10 do not show much fluctuations when compared to Figure 6.11. Since the pie charts were constructed on relative utilisation index values, the error in predictions of the different utilisation values are very close to each other which diminishes in the relative utilisation index calculation. Table 6.1 shows the Actual and Predicted values that was shown in Figure 6.12.

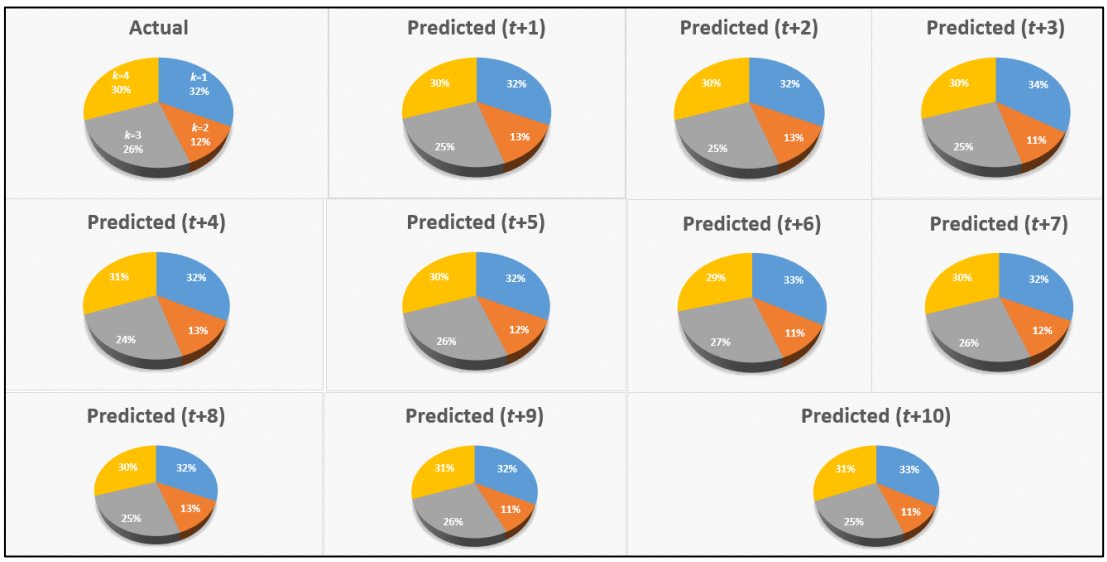


Figure 6.12: Actual value comparison with different prediction stages

Table 6.1: Data showing Actual and Predicted values

	***k*=1**	***k*=2**	***k*=3**	***k*=4**
Actual	0.32057	0.121710	0.254900	0.302810
t+1	0.33095	0.137378	0.261383	0.312186
t+2	0.33095	0.132033	0.261364	0.312223
t+3	0.33095	0.111140	0.241706	0.292998
t+4	0.33095	0.129593	0.244844	0.312454
t+5	0.33095	0.124412	0.262580	0.311953
t+6	0.33095	0.113010	0.264500	0.292886
t+7	0.33095	0.127237	0.264719	0.311760
t+8	0.33095	0.134916	0.265184	0.312199
t+9	0.33095	0.108819	0.269266	0.313078
t+10	0.33095	0.108603	0.247038	0.312775

Consequently, the following results show the prediction quality obtained by presenting a few simulation results where each software utilisation is demonstrated. Although the simulation could be continued up to any desired lag, in this framework, the simulation carried out up to 10 times (*t*+10) where 10 future predictions were made. Figure 6.13 displays four different (*K*=4) software utilisations where the prediction was carried out. The prediction is carried out on each *k* index, where the *k* index is identified by colour. The first prediction results were shown in Figure 6.14 and Figure 6.15 where Figure 6.14 shows six future predictions of the software utilisation (from *t*+1 to *t*+6) and Figure 6.15 shows the remaining four lags of prediction (from *t*+7 to *t*+10) as well as RMSE value changes with the prediction lag (*T*=10). A similar approach was taken to predict future software utilisation for *k*=2, *k*=3 and *k*=4 and these results were presented in Figure 6.16, Figure 6.17, Figure 6.18, Figure 6.19, Figure 6.20 and Figure 6.21, respectively. It was noted that the prediction quality was the highest level in $\hat{u}_k^{(t+1)}$ (99% accuracy) and it was the lowest level in $\hat{u}_k^{(t+10)}$ (94% accuracy) which was an expected outcome, and expected to decrease with an increase in T_f. The main reason behind the decrease in prediction quality is that for every future prediction made using the previous prediction, i.e. when $\hat{u}_k^{(t+2)}$ is predicted, $\hat{u}_k^{(t+1)}$ was used and similarly, when $\hat{u}_k^{(t+3)}$ is predicted, $\hat{u}_k^{(t+2)}$ was used. A similar approach was taken to predict future values such as $\hat{u}_k^{(t+10)}$, indicating that $\hat{u}_k^{(t+10)}$ demonstrates the lowest quality of prediction. What is noticeable from Figure 6.13 to 6.25 is that every prediction made is subsequently used in a succeeding stage. For example Figure 6.13 shows the initial

software utilisation before prediction. This is then used for a 10-step prediction in Figures 6.14 and 6.15 for four different software utilisations. The last graph in Figure 6.15 next to the legend then shows the RMSE prediction.

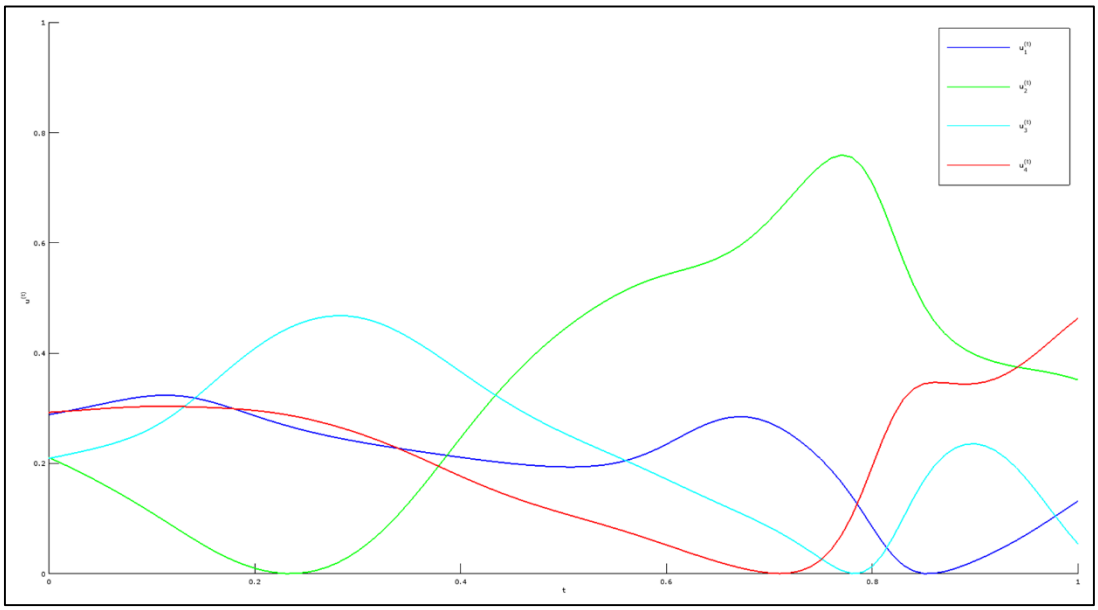


Figure 6.13: Initial software utilisation before prediction

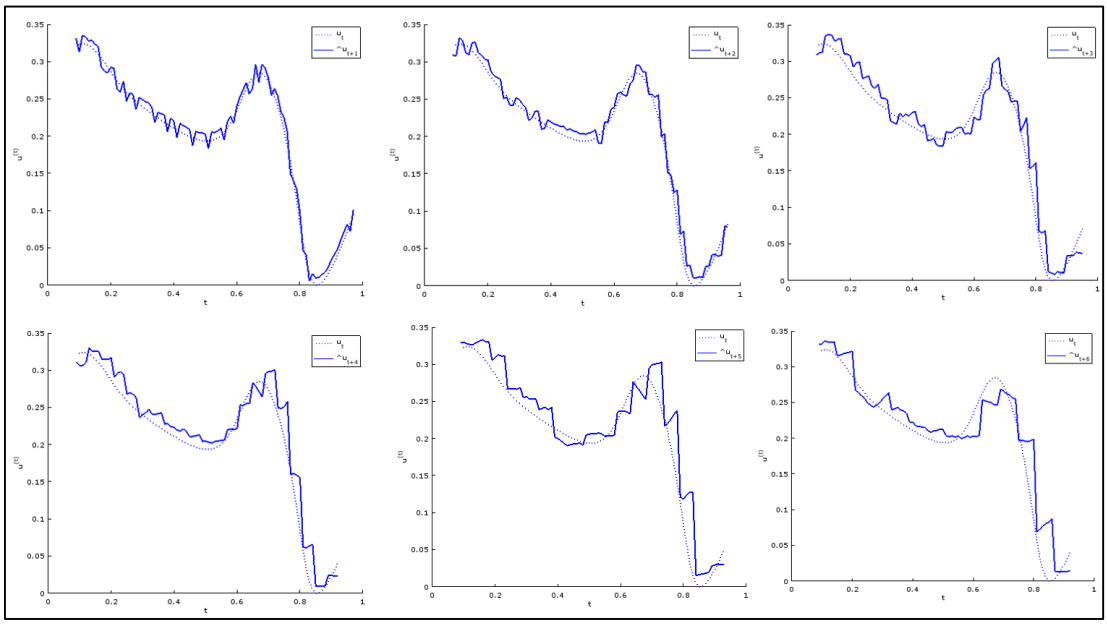


Figure 6.14: Software utilisation prediction for k=1 (from *t*+1 to *t*+6)

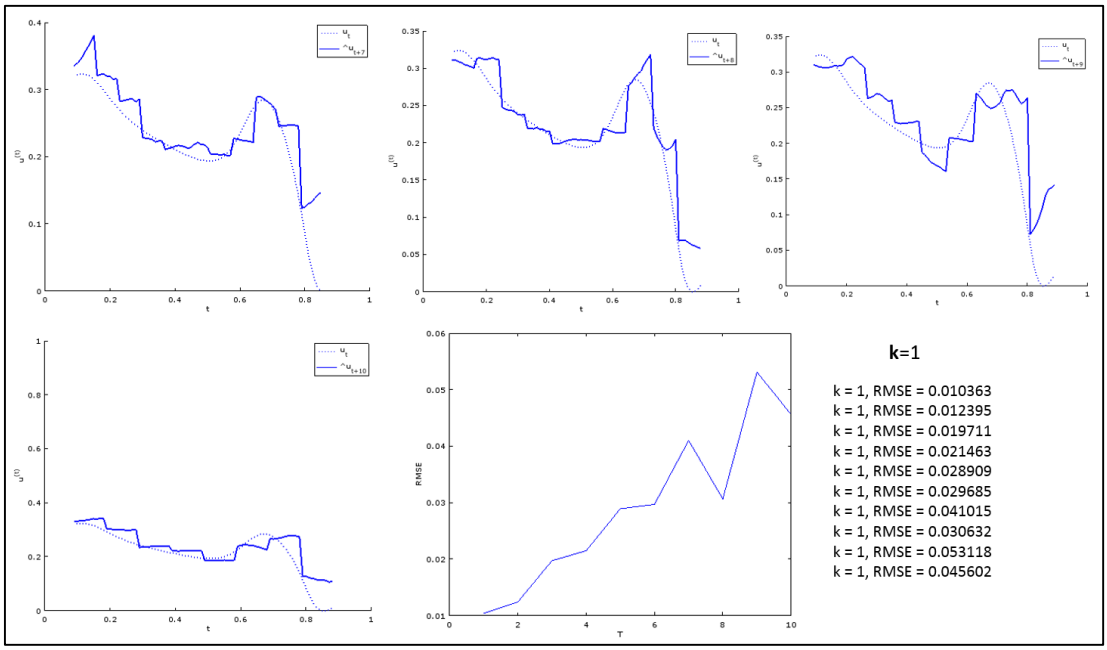


Figure 6.15: Software utilisation prediction for k=1 (from *t*+7 to *t*+10) with RMSE values

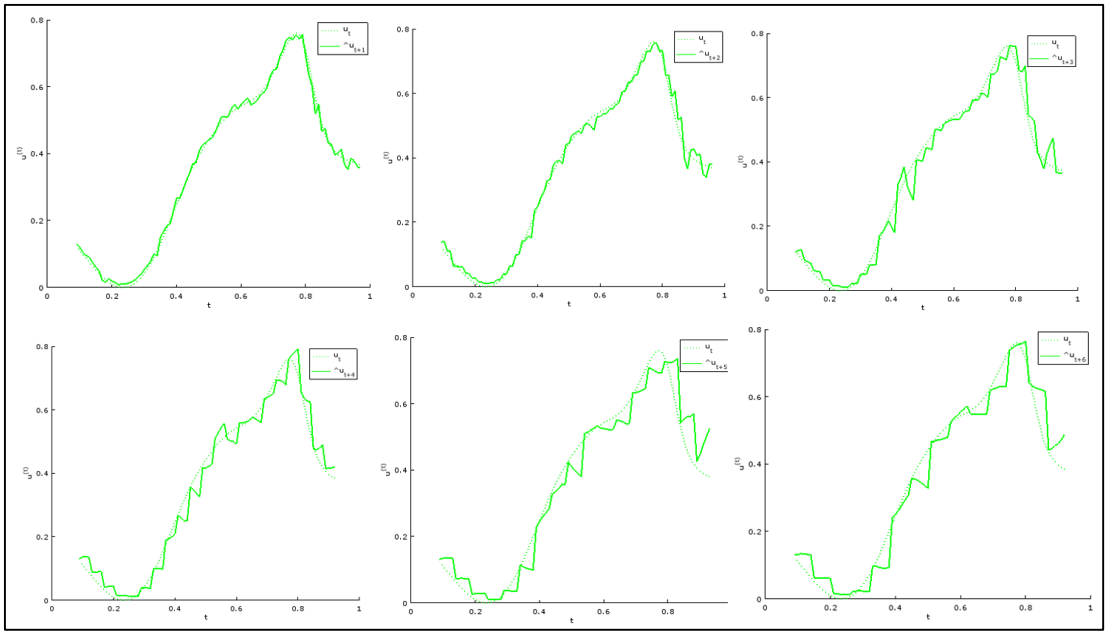


Figure 6.16: Software utilisation prediction for k=2 (from *t*+1 to *t*+6)

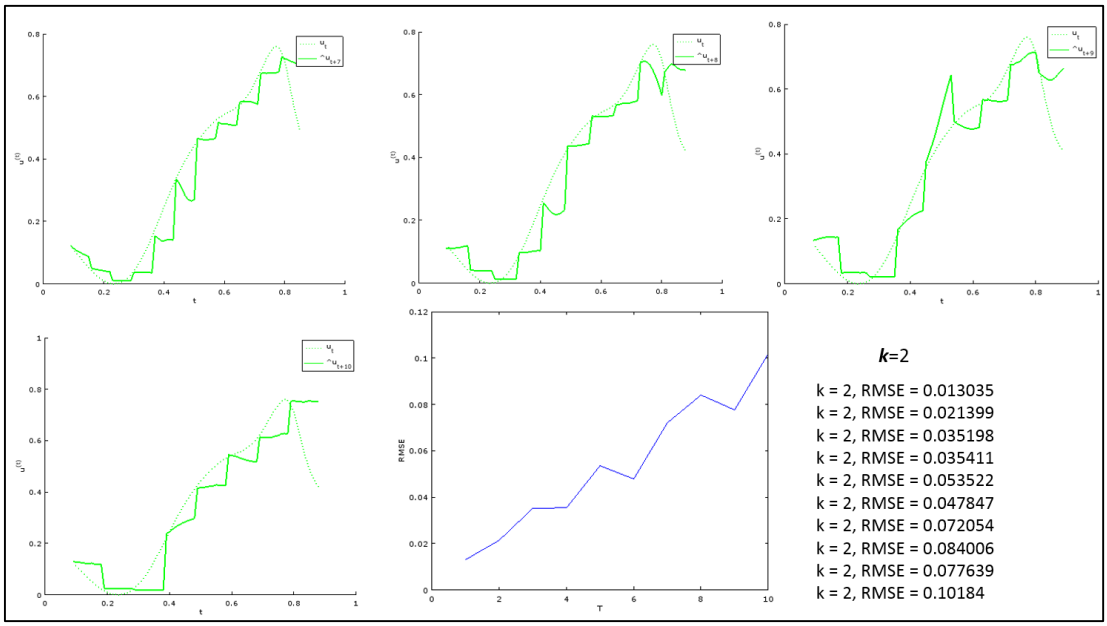


Figure 6.17: Software utilisation prediction for k=2 (from *t*+7 to *t*+10) with RMSE values

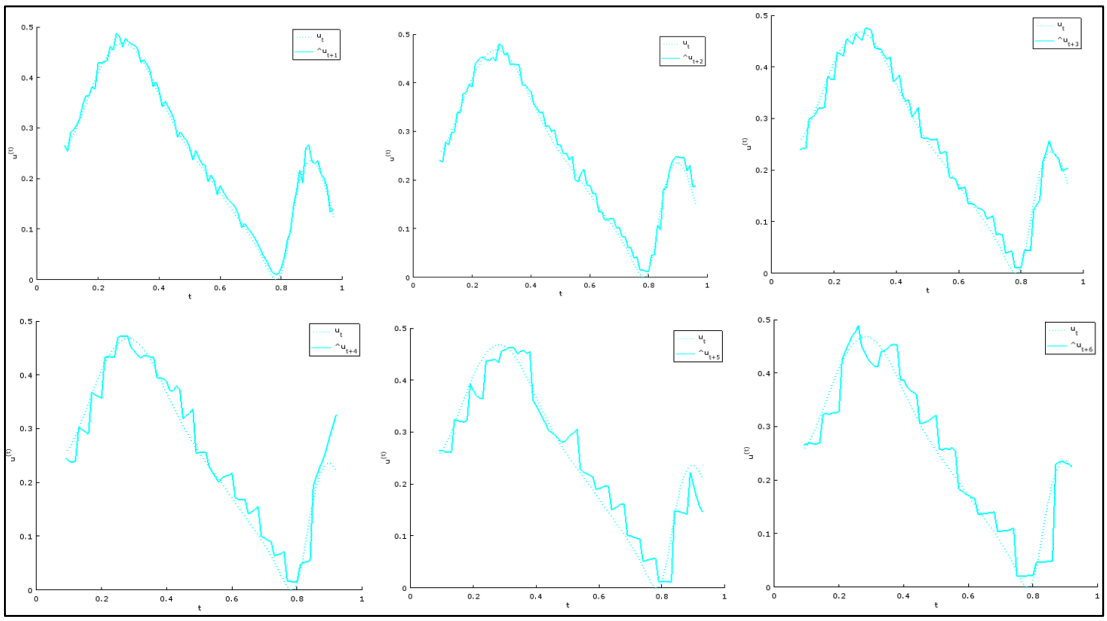


Figure 6.18: Software utilisation prediction for k=3 (from *t*+1 to *t*+6)

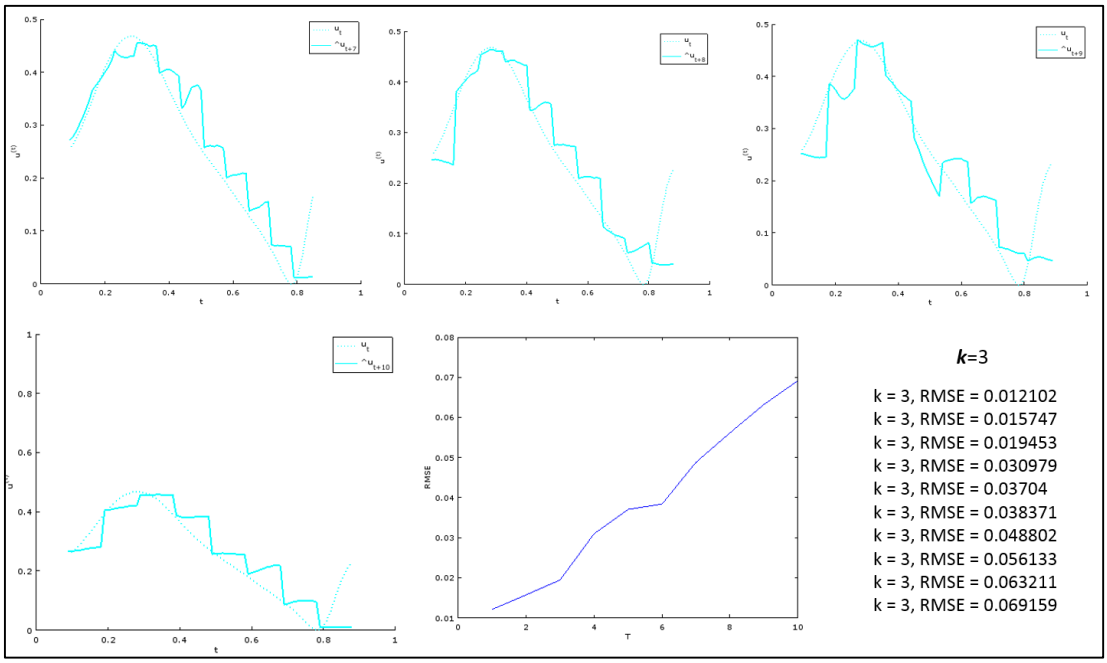


Figure 6.19: Software utilisation prediction for k=3 (from *t*+7 to *t*+10) with RMSE values

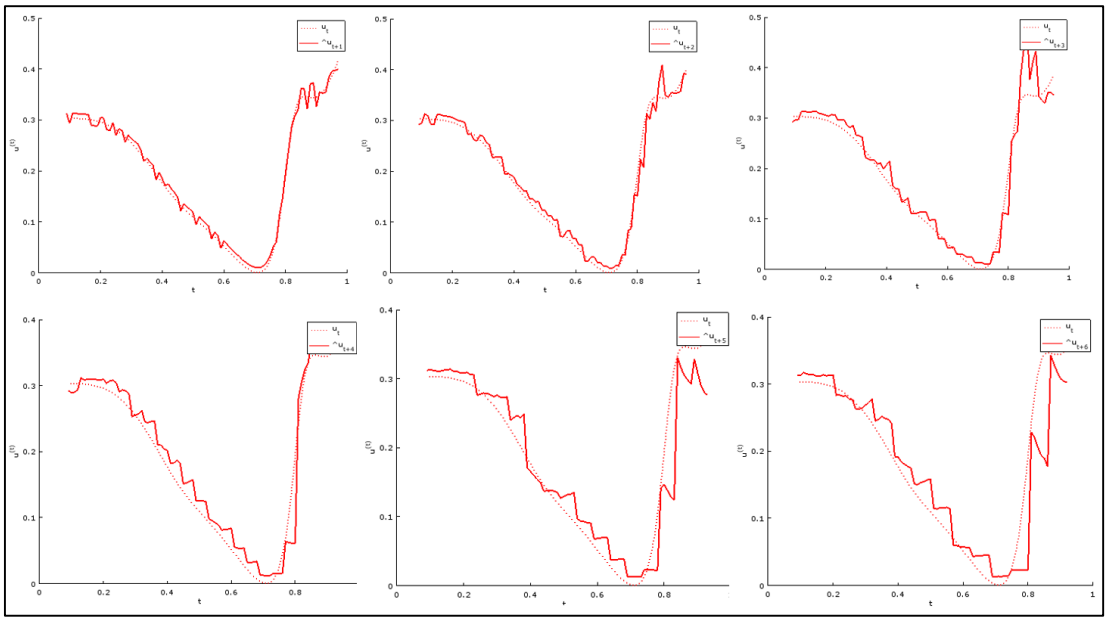


Figure 6.20: Software utilisation prediction for k=4 (from *t*+1 to *t*+6)

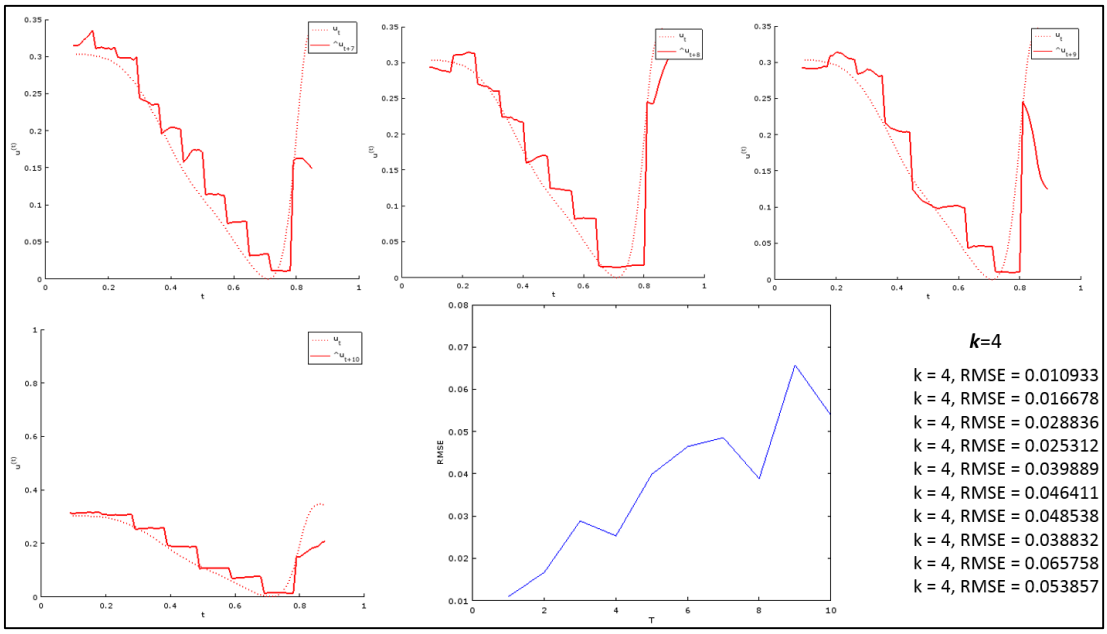


Figure 6.21: Software utilisation prediction for k=4 (from *t*+7 to *t*+10) with RMSE values

After performing individual software utilisation prediction, the following figures (Figure 6.22, Figure 6.23, Figure 6.24 and Figure 6.25) show combined prediction results for selected future software prediction lags, i.e. *t*+1, *t*+3, *t*+5 and *t*+10 where dotted lines show actual software utilisation and straight lines show predicted software utilisation.

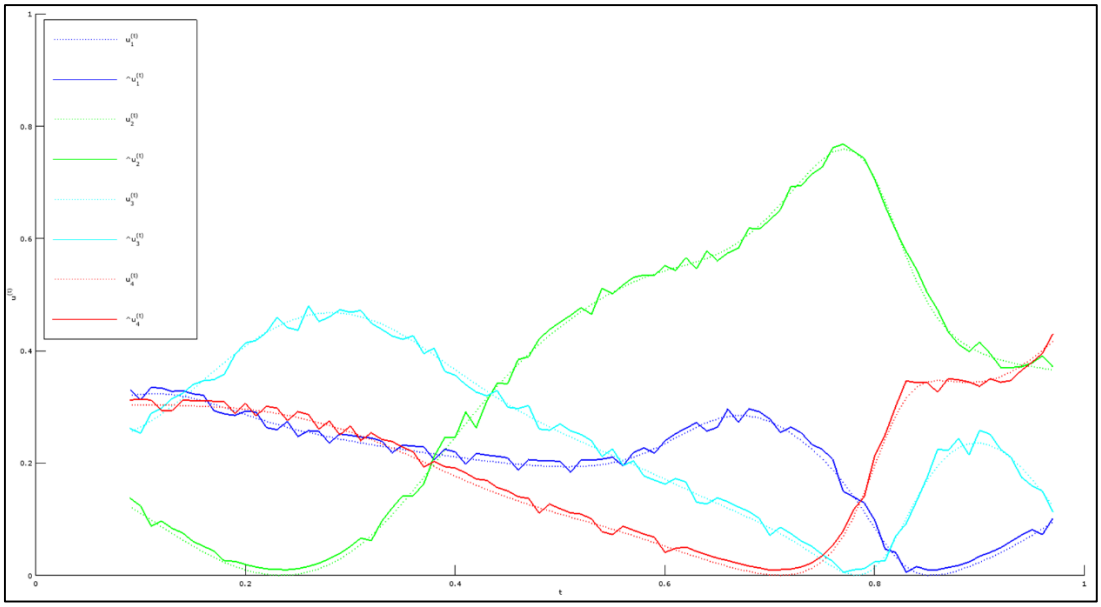


Figure 6.22: Software utilisation prediction for *t*+1

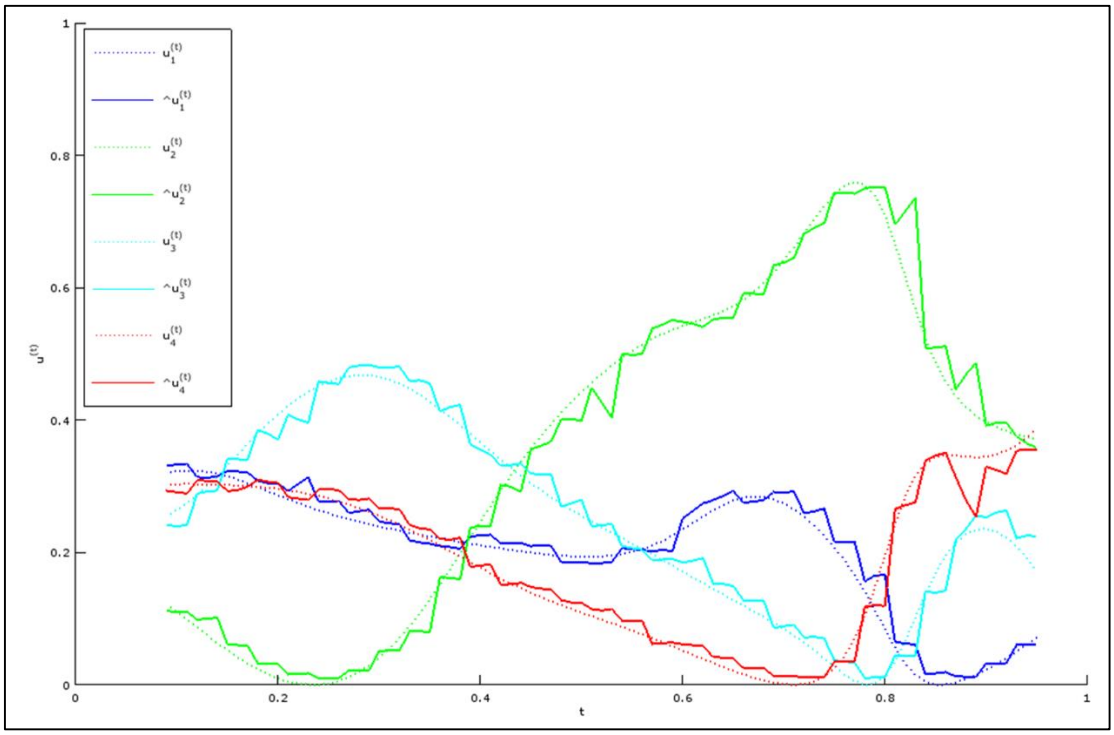


Figure 6.23: Software utilisation prediction for *t*+3

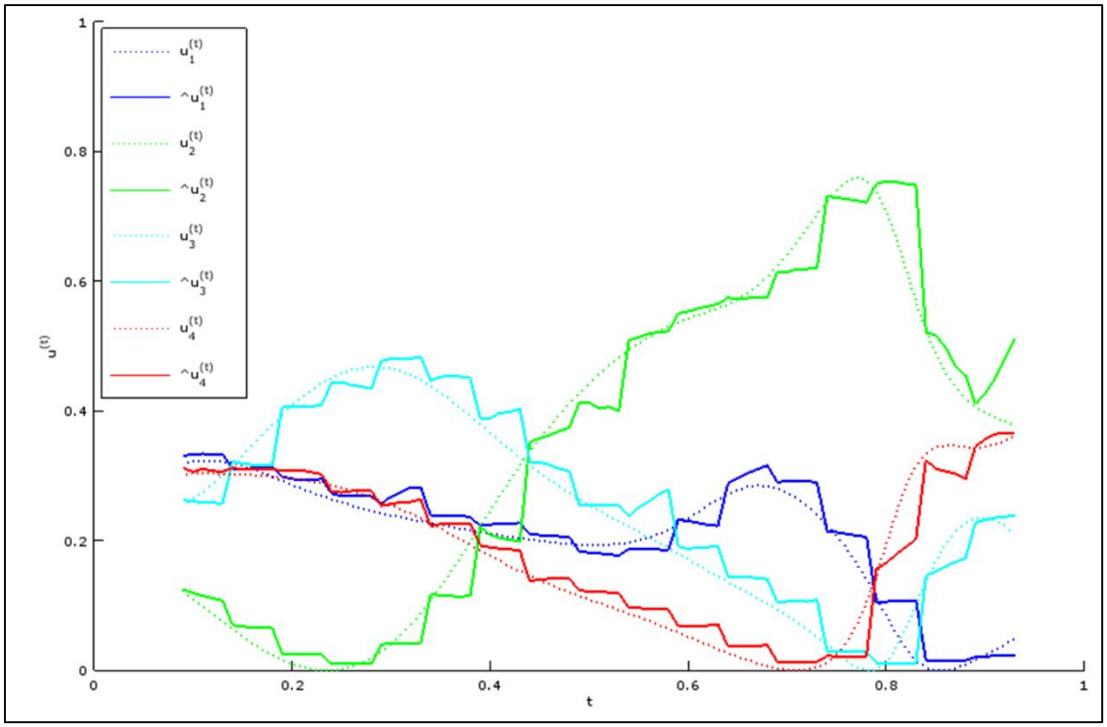


Figure 6.24: Software utilisation prediction for *t*+5

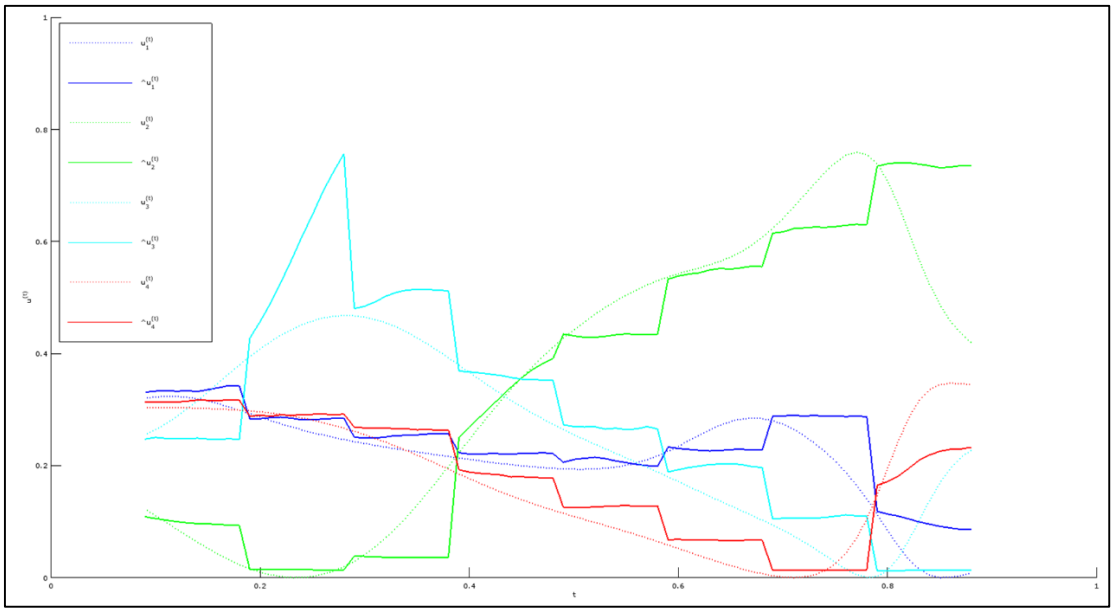


Figure 6.25: Software utilisation prediction for *t*+10

6.7 Discussion of the results

After developing a framework to conduct time-series analysis, a number of figures were generated to predict the future values of mine planning software utilisation. The results presented above show that the prediction made in this thesis is reliable and comparable to any successful time-series analysis outcome. It was noted that the prediction quality was the highest level in $\hat{u}_k^{(t+1)}$ (99% accuracy), and it was the lowest level in $\hat{u}_k^{(t+10)}$ (94% accuracy), which was an expected outcome, and expected to decrease with an increase in T_f.

6.8 Summary

In this Chapter, time-series analysis was performed by developing a framework to estimate future values of mine planning software utilisation in South Africa. In this framework, because of lack of available data, a simulated dataset was generated. Using a number of developed scripts and changing a number of variables within the scripts, a model was developed and prediction on that model was performed. The prediction results indicated that, the time-series analysis that was performed on the

mine planning software utilisation was successful in predicting the future values of mine planning software utilisation with up to 99% confidence level.

7 CONCLUSIONS AND RECOMMENDATIONS

7.1 Introduction

The South African mining industry makes use of mine planning software in every stage of the mining operations. To gain maximum benefit from these operations, efficiency within the mining sector needs to be improved. To improve efficiency, a methodology to evaluate mine planning software utilisation for enhanced decision-making strategies was developed by calculating mine planning software utilisation with two time-stamps. Through these informed decision-making strategies, the South African mining industry is now in a position to choose the leading software solutions, which will lead to improved production efficiencies. Furthermore, after calculating the software utilisation index, this research predicted future utilisation of mine planning software by conducting a time-series analysis. This chapter notes the conclusions drawn and recommendations made for further work from this research study.

7.2 Key findings

Key findings of this thesis can be summarised as follows:

- The methodology for the evaluation of utilisation of mine planning software in various commodity sectors was developed on the basis of three variables, namely, commodity, functionality, and time factor, as a key evaluation criteria. Even though calculations were done only on two commodities in Chapter 4, Table 7.1 extends the calculation results to four commodities namely coal, diamond, gold and PGM for the 2014 time-stamp and lists the top three most utilised software per category. Six functionalities namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation were applied on the four different commodities using two different time-stamps (September 2012 and April 2014);
- It was found that the market leaders in terms of mine planning software utilisation in South Africa differs, depending on the commodity that is being mined as well as the functionality that is being used;

- Now, the South African mining industry has a framework that can be utilised in order to ultimately enhance decision-making strategies for utilisation of the mine planning software. This information can then be used further too, to better inform decision-making strategies on software utilisation by software developers, educational institutions and consulting companies from different perspectives. South African mining companies could use this framework to choose the leading software combinations to improve productivity, the educational institutions could make use of the framework to choose which software solutions they can expose their students to, consulting companies can advise their clients more effectively by using the framework and more importantly, software providers can compare their positions and develop more informed decision-making strategies accordingly;
- To successfully predict future time-series values, the ANN's training/learning section was divided into two main categories namely; training and testing, as the dataset is not validated due to the temporal nature of the mine planning software database, the dataset used in time-series analysis is simulated data, hence the database does not require to be validated. Supervised learning, which is common in time-series analysis was used to train the mine planning software dataset;
- During prediction, the learned weights and obtained software utilisation is used to predict the target value, and every obtained new value subsequently used in a succeeding step; and
- After developing a framework to conduct time-series analysis, a number of figures were generated to predict the future values of mine planning software utilisation. A new framework to predict future utilisation of mine planning software was developed with an almost 99% confidence level.

Table 7.1: Commodity/Functionality table showing most utilised mine planning software in 2014

	Geological Data Management	Geological Modelling and Resource Estimation	Design and Layout	Scheduling	Financial Valuation	Optimisation
Coal	Minex	Minex	Minex	Xpac	Xeras	Xeras
	Sable Data Warehouse	Surpac	Talpac	Dragsim	Carbon Economics	Dragsim
	Surpac	Vulcan	Surpac	MineSched		Carbon Economics
Diamond	GEMS	GEMS	GEMS	Enhanced Production Scheduler (CAE)	Whittle	Enhanced Production Scheduler (CAE)
	Maptek I-Site	Studio 3 - Geology	Talpac	EPS (MineRP)	MRM	Studio 5D Planner
	Sirovision	Sirovision	Studio 5D Planner	Studio 5D Planner		Whittle
Gold	Geological Data Management Solution	Studio 3 - Geology	Studio 3 - Geology	MRM	MRM	Studio 3 - Geology
	Pegs Lite	CADSMine	MRM	CADSMine	Carbon Economics	Enhanced Production Scheduler (CAE)
	MRM	MRM	CADSMine	Enhanced Production Scheduler (CAE)	Maxipit	Carbon Economics
PGM	Geological Data Management Solution	Studio 3 - Geology	MRM	MRM	MRM	Studio 3 - Geology
	Pegs Lite	MRM	CADSMine	CADSMine	Carbon Economics	MRM
	MRM	CADSMine	Studio 3 - Engineering	Bentley Scheduler	Bentley Evaluation	Carbon Economics

7.3 Research contributions: Methodology and decision-making strategies

This research provided a tool, a framework, a methodology to evaluate mine planning software utilisation for enhanced decision-making strategies, which was the main objective of this thesis. By making use of this tool, the South African mining industry, mine planning software providers, institutions that offer mining related qualifications as well as mining consultancy companies are now in a position to develop enhanced decision-making strategies broadly outlined as follows:

- These enhanced decision-making strategies can enable the South African mining industry to choose the optimal mine planning software combinations, which can ultimately lead to improved production efficiencies;
- Using this methodology, educational institutions in South Africa can easily make decisions regarding which strategically important mine planning software solutions they are going to expose their students to;
- As mergers and acquisitions will continue re-shaping the mine planning software field in this very competitive market, the mine planning software

providers are now also in a position to understand their market impact so that they can develop appropriate decision-making strategies for future growth; and

- Mining consultancy companies can advise their clients more effectively as the methodology provided through this thesis will help them to make the right choices of software solutions so that they can better serve their clientele.

This research also provided a new framework to predict future utilisation of mine planning software. The prediction results indicated that, the time-series analysis that were performed on the mine planning software utilisation was successful in predicting the future values of mine planning software utilisation with up to a 99% confidence level.

7.4 Research limitations

There were a number of limitations to this research. The major problem was obtaining the required information, i.e. it took more than a year to get up to date information, including data from CAE Mining. However, after presenting some of the outcomes from this research during SAIMM's "Mine Planning School – 2014", a few more software providers showed interest in participating and becoming part of the database. Furthermore, it was impossible to check and confirm if the information provided by the software producers in fact was accurate and the database heavily relied on their good faith. Finally, this research made use of scripting language intensively which was a time consuming exercise and it took a lot of time to implement.

7.5 Recommendations for future research work

As mine planning software providers have established offices around the globe, it would be ideal to extend the current database to include datasets from other mining countries such as Australia, Canada and USA. This would bring an added dimension to comparison when it comes to multi-national software utilisation.

One of the outcomes of this research was to predict future software utilisation but due to the lack of available data, a simulated dataset had to be used to conduct the time-

series analysis. For future improvements, software producers need to be approached on a regular basis to obtain the latest available dataset, initially at two-year intervals and continue testing the ANN time-series analysis.

8 REFERENCES

acQuire, 2013. *Geoscientific Information Management System.* [Online]
Available at: http://www.acquire.com.au
[Accessed 18 December 2013].

Aha, D. & Bankert, R., 1997. Cloud classification using error-correcting output codes. *Journal of Artificial Intelligence Application for Natural Resources in Agriculture Environtental Science,* 11(1), pp. 13-28.

Apte, C., Damerau, F. & Weiss, S. M., 1994. Automated learning of decision rules for text categorization. *Information Systems,* 12(3), p. 233–251.

ArcGIS, 2013. *The Mapping Platform for Your Organization.* [Online]
Available at: http://www.arcgis.com
[Accessed 18 December 2013].

Baldi, P. & Pollastri, B., 2002. A machine learning strategy for protein analysis. *IEEE Intelligent Systems,* 17(2), pp. 28-35.

Bascetin, A. & Sertabipoglu, Z., 2010. *Strategic Mine Planning and Design for Quarrying.* Phoenix, Society for Mining, Metallurgy & Exploration.

Bishop, C. M., 1996. *Neural networks for pattem recognition.* New York: Oxford University Press.

Bogotobogo, 2014. *PHP & MySQL Tutorial - Request 2014.* [Online]
Available at: http://www.bogotobogo.com/php/php1.php
[Accessed 29 July 2014].

Brill, E., 1994. *Some advances in transformation-based part of speech tagging.* s.l., AAAI Conference.

CADSmine, 2013. *Specialised mine planning and design for tabular orebodies.* [Online]
Available at: http://www.minerpsolutions.com/en/software-products/planningsched/CADSmine.aspx
[Accessed 20 December 2013].

CAE NPV Scheduler, 2013. *The complete strategic mine planning system.* [Online]
Available at: http://www.cae.com/mining/software/open-pit-planning-solution/
[Accessed 20 December 2013].

CAE Studio 5D Planner, 2014. *CAE's successor to Mine2-4D.* [Online]
Available at: http://www.cae.com/mining/software/underground-planning-solution/
[Accessed 12 May 2014].

CAE Studio3, 2013. *The Complete Solution for Geology and Mine Design..* [Online]
Available at: http://support.caemining.com/studio3-brochure
[Accessed 20 December 2013].

Chamber of Mines, 2014. *Natural Minerals | Mining Industry | Chamber of Mines SA.* [Online]
Available at: http://chamberofmines.org.za/mining-industry/comsa-members
[Accessed 12 June 2014].

Chaney, A., 2014. *Server side scripting.* [Online]
Available at: http://www.princeton.edu/~achaney/tmve/wiki100k/docs/Server-side_scripting.html
[Accessed 1 January 2014].

Clare, A. & King, R. D., 2001. *Knowledge discovery in multi-label phenotype data.* Berlin, Springer.

Dagdelen, K. & Kawahata, K., 2007. *Value Creation Through Strategic Mine Planning and Cutoff Grade Optimization.* Denver, SME: Mining Metallurgy and Exploration.

Department of Mineral Resources, 2012. *South African Minerals Industry (SAMI).* [Online]
Available at: http://www.dmr.gov.za/publications/south-africas-mineral-industry-sami/viewdownload/217/1021.html
[Accessed 12 March 2014].

Deswik, 2013. *Leaders in innovative mining software & consultancy.* [Online]
Available at: http://www.deswik.com
[Accessed 20 December 2013].

El-Ramly, M. & Stroulia, E., 2004. *Mining software usage data.* Scotland, s.n., pp. 64-68.

Esri, 2014. *Exploration.* [Online]
Available at: http://www.esri.com/industries/mining/business/exploration
[Accessed 28 July 2014].

Even-Zohar, Y. & Roth, D., 2001. *A sequential model for multiclass classification.* s.l., SIGDAT Conference on Empirical Methods in Natural Language Processing.

GEOVIA GEMS, 2013. *Integrated Geology, Resource Modelling, Mine Planning and Production.* [Online]
Available at: http://www.gemcomsoftware.com/products/GEMS
[Accessed 22 December 2013].

GEOVIA MineSched, 2013. *Great Results with Proven Technology.* [Online]
Available at: http://www.gemcomsoftware.com/products/MineSched
[Accessed 23 December 2013].

GEOVIA Minex, 2013. *The Only Integrated, End-to-end Geology and Mine Planning Solution for Coal and other Stratified Deposits.* [Online]
Available at: http://www.gemcomsoftware.com/products/Minex
[Accessed 24 December 2013].

GEOVIA Surpac, 2013. *Integrated Geology, Resource Modelling, Mine Planning and Production.* [Online]
Available at: http://www.gemcomsoftware.com/products/Surpac
[Accessed 20 December 2013].

GEOVIA Whittle, 2013. *Strategic Mine Planning - Optimise Mine Profits.* [Online]
Available at: http://www.gemcomsoftware.com/products/Whittle
[Accessed 21 December 2013].

Gibbs Associates, 2004. *Mining Software Summaries.* [Online]
Available at: http://www.miningsoftware.com/swmine/software.html
[Accessed 19 May 2014].

Gibbs, B., 1994. Computers – The Catalyst for Information Accessibility. *SME.*

Hertz, J., Krogh, A. & Palmer, R. G., 1991. *Introduction to the Theory of Neural Computation.* Redwood City: Addison-Wesley.

Immersive Technologies, 2013. *Advanced Equipment Simulators for Mining and Earthmoving Equipment Training.* [Online]
Available at: http://www.immersivetechnologies.com
[Accessed 29 December 2013].

Jacobson, L., 2013. *Introduction to Artificial Neural Networks - Part 1.* [Online]
Available at: http://www.theprojectspot.com/tutorial-post/introduction-to-artificial-neural-networks-part-1/7
[Accessed 10 June 2014].

Jain, A. K., Mao, J. & Mohiuddin, K., 1996. Artificial Neural Networks: A Tutorial. *IEEE,* 29(3).

Johnson, K., 2011. *Visual Computing and Mine Design.* Wollongong, 35th APCOM Symposium.

Johnstone, W., 2003. Effective multi-dimensional orebody modelling: A key component of the mining value chain. *Journal of the Southern African Institute of Mining and Metallurgy.*

Kapageridis, I. K., 2005. *The Future of Mine Planning Software – New Tools and Innovations.* Izmir, 19th International Mining Congress and Fair of Turkey.

Karimi, B., Menhaj, M. & Saboori, I., 2010. Multilayer feed forward neural networks for controlling de-centralized large-scale non-affine nonlinear systems with guaranteed stability. *International Journal of Innovative Computing, Information and Control,,* 6(11), pp. 4825-4841.

Karlik, B. & Olgac, V., 2010. Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks. *International Journal of Artificial Intelligence And Expert Systems (IJAE),* 1(4), pp. 111-122.

Katakwa, T., Musingwini, C. & Genc, B., 2013. Online database of mine planning and peripheral software used in the South African mining industry. *Journal of the Southern African Institute of Mining and Metallurgy,* June, Volume 113, pp. 497-504.

Kimberling, E., 2009. *Not all ERP software is created equal: Comparing, SAP, Oracle EBS and MS Dynamics.* [Online]
Available at: http://www.youtube.com/watch?v=dl3ZxTTlrF8
[Accessed 19 June 2014].

Latimer, C., 2013. *Mapping: Planning to production.* [Online]
Available at: http://www.miningaustralia.com.au/features/mapping-planningtoproduction
[Accessed 16 July 2014].

LeCun, Y. et al., 1989. Back-propagation applied to handwritten zip code recognition. *Neural Computation,* 1(4), p. 541–551.

MathWorks, 2015. *Improve Neural Network Generalization and Avoid Overfitting.* [Online]
Available at: http://www.mathworks.com/help/nnet/ug/improve-neural-network-generalization-and-avoid-overfitting.html
[Accessed 8 February 2015].

McCulloch, W. S. & Pitts, W., 1943. *A logical calculus of the ideas immanent innervous activity.* s.l.:Bulletin of Mathematical Biophysics.

Mine2-4D, 2013. *The best underground mine planning tool by far!.* [Online]
Available at:
http://www.http://www.minerpsolutions.com/en/software/planningsched/Mine2-4D.aspx.com/en/software/planningsched/Mine2-4D.aspx
[Accessed 19 December 2013].

Minemax, 2013. *Mine Planning and Scheduling Solutions.* [Online]
Available at: http://www.minemax.com
[Accessed 18 December 2013].

MineSight, 2013. *Mining is essential.* [Online]
Available at: http://www.minesight.com/en-us/packages.aspx
[Accessed 18 December 2013].

Minsky, M. & Papert, S., 1969. *Perceptrons.* Cambridge: MIT Press.

Modular's DISPATCH, 2013. *The DISPATCH fleet management system.* [Online]
Available at: http://modularmining.com/product/dispatch
[Accessed 29 December 2013].

Murphey, Y. L. & Luo, Y., 2002. *Feature extraction for a multiple pattern classification neural network system.* s.l., IEEE International Conference on Pattern Recognition.

Murphey, Y. L. & Luo, Y., 2002. *Feature extraction for a multiple pattern classification neural network system.* s.l., IEEE International Conference on Pattern Recognition.

Nakajima, S., 1998. *Introduction to Total Productive Maintenance..* Cambridge: MA.

Opoku, S. & Musingwini, C., 2012. *Modelling Geological Uncertainty For Open-Pit To Underground Transition In Gold Mines.* New Delhi, Reading Matrix Inc..

Owram, K., 2014. *Flight simulator firm CAE puts mining unit up for sale as civil aviation profit soars.* [Online]
Available at: http://business.financialpost.com/2014/08/13/cae-earnings/
[Accessed 1 September 2014].

Oxford English Dictionaries, 2014. *Utilize.* [Online]
Available at: http://www.oxforddictionaries.com/definition/english/utilize?q=utilize

Prinsloo, L. & Marais, J., 2014. *Outlook gloomy for SA's mining sector on eve of indaba.* [Online]
Available at: http://www.bdlive.co.za/business/mining/2014/02/02/outlook-gloomy-for-sas-mining-sector-on-eve-of-indaba
[Accessed 19 July 2014].

Rakesh, B., 2014. *Why choose PHP for enterprise web development?.* [Online]
Available at: http://php.osscube.com/blog/why-choose-php-enterprise-web-development
[Accessed 2 January 2014].

RocScience Software, 2013. *Virtual Reality at MIRARCO - Mining Innovation.* [Online]
Available at: http://www.rocscience.com/library/rocnews/fall2002/Miraco.pdf
[Accessed 29 December 2013].

RocScience, 2013. *Software tools for rock and soil.* [Online]
Available at: http://www.rocscience.com
[Accessed 16 December 2013].

Rosenblatt, F., 1958. *he perception: A probabilistic model for information storage and organization in the brain.* s.l.:Psychological Reviews.

Rumelhart, D. E., McClelland, J. L. & Group, P. R., 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Cambridge: MIT Press.

Sable, 2013. *Data that works.* [Online]
Available at: http://www.sable.co.za
[Accessed 18 December 2013].

Sable, 2014. *Data that works.* [Online]
Available at: http://www.sable.co.za
[Accessed 16 June 2014].

Shirose, K., 2013. *Equipment Utilization Metrics.* [Online]
Available at:
http://www.ombuenterprises.com/LibraryPDFs/Equipment_Utilization_Metrics.pdf
[Accessed 15 June 2014].

Smit, A. & Lane, G., 2010. *Mine Optimization and its Application Using the Anglo Platinum Mine Optimisation Tool (APMOT).* Sun City, The Southern African Institute of Mining and Metallurgy.

Squelch, A. P., 2001. Virtual reality for mine safety training in South Africa. *Journal of the Southern African Institute of Mining and Metallurgy,* pp. 209-217.

Stergiou, C. & Siganos, D., 1996. *Neural Networks.* [Online]
Available at: http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html
[Accessed 14 June 2014].

Surfer, 2013. *Powerful Contouring, Gridding, and 3D Surface Mapping Software for Scientists and Engineers.* [Online]
Available at: http://www.goldensoftware.com/products/surfer
[Accessed 20 December 2013].

The Open Group, 2010. *The exploration and mining business reference model.* [Online]
Available at: https://collaboration.opengroup.org/emmmv/documents/22706/Getting_started_with_the_EM_Business_Model_v_01.00.pdf
[Accessed 23 July 2014].

Trimble, 2014. *Trimble Open Pit Design.* [Online]
Available at: http://pinpoint.microsoft.com/en-us/applications/trimble-open-pit-design-12884966681
[Accessed 26 July 2014].

ViMINE, 2013. *ViMINE: UNSW's innovative mine planning tool.* [Online]
Available at: http://www.engineering.unsw.edu.au/mining-engineering/vimine-unsws-innovative-mine-planning-tool
[Accessed 29 December 2013].

Vorne Industries, 2008. *The Fast Guide to OEE.* [Online]
Available at: http://www.vorne.com/pdf/fast-guide-to-oee.pdf
[Accessed 15 June 2014].

Vulcan, 2013. *3D Modelling and Mine Planning Software.* [Online]
Available at: http://www.maptek.com/products/vulcan
[Accessed 29 December 2013].

Whittle, G. & Burks, S., 2010. *Simultaneous Mining and Mineral Processing Enterprise Optimization for the Platinum Industry.* Sun City, The Southern African Institute of Mining and Metallurgy.

Wolf, T., 2013. *What is ERP Software?.* [Online]
Available at: http://www.tmcnet.com/channels/erp-software/articles/173165-what-erp-software.htm
[Accessed 28 December 2013].

XERAS, 2014. *XERAS Strategic Financial Evaluation.* [Online]
Available at: http://www.rpmglobal.com/mining-software/financial-modelling-xeras/xeras-strategic-financial-evaluation
[Accessed 14 May 2014].

XPAC, 2014. *XPAC mine scheduling software.* [Online]
Available at: http://www.rpmglobal.com/mining-software/mine-scheduling-xpac
[Accessed 19 May 2014].

Zainuddin, Z. & Ong, P., 2007. Function Approximation Using Artificial Neural Networks. *International Journal of Systems Applications, Engineering & Development,* 1(4), pp. 173-178.

Zhang, B., Chen, Z. & Murphey, Y. L., 2005. *Protein secondary structure prediction using machine learning.* s.l., IEEE International Joint Conference on Neural Networks.

9 APPENDICES

9.1 Summary of paper abstracts on the research

The publications listed below have originated from this thesis so far:

9.1.1 Paper 1

Parts of Chapters 3 were compiled and published as the paper: Katakwa, TP, Musingwini, C and **Genc, B (2013).** Online database of mine planning and peripheral software used in the South African mining industry. *Journal of The Southern African Institute of Mining and Metallurgy*, Volume 113, No. 6, June 2013, pp. 497-504. An electronic copy of the paper is available on SAIMM website at INTERNET: http://www.saimm.co.za/Journal/v113n06p497.pdf. The following is an extract of the paper abstract:

"Abstract

The utilization of software is now inherent to virtually every activity along the mining value chain. However, apart from the software survey done by Gibbs in the 1990s and work on the extent of diffusion of information and communication technology (ICT) in the South African platinum sector by Mugodi and Fleming in 2003, the nature and extent of software utilization in the South African mining industry has never been evaluated. The Mine Planning, Optimisation and Valuation (MPOV) Research Group in the School of Mining Engineering, University of Witwatersrand, therefore initiated a project to collate and analyse the current utilization of software in the South African mining industry. This was done through the development of a web-based database of the relevant software. Snowball sampling was used to collect the data because the South African mining industry is diverse and software utilization is fragmented across and within the sectors of the industry. The data was then organized into distinct categories so that the information from a variety of sources could be evaluated on the same basis. A beta version of the database can be accessed online through a user-friendly front-end platform at http://db.mining.wits.ac.za.

The database is expected to help at least 13 educational institutions with decisions on facilities and training that are vital to the education of mining and mining-related professionals. Exploration, mining, and consulting companies will also benefit from information in the database relating to availability and useful combinations of software solutions. The database is also strategic to software providers by providing a better understanding of their respective relative market share along the mining value chain. An analysis of the data collated in this research shows that about 77% of the software users are mining companies, 17% are consulting companies, 3% are mineral exploration companies, and the rest are software providers and educational institutions. The software used in the South African mining industry is largely provided by Gemcom Software International, MineRP Solutions, and MRM Mining Services. CAE Mining, who in 2010 acquired the Datamine Group providing Datamine software, is also widely acknowledged as a major software supplier in South Africa, but data from them had not been obtained at the time of producing the beta version of the database due to proprietary constraints. The work reported in this paper is part of an MSc research study in the School of Mining Engineering at the University of the Witwatersrand.

Keywords: *Mine planning, mining, value chain, software, database, South African mining industry, utilization"*

9.1.2 Paper 2

Parts of Chapters 3 and 4 were compiled and will be published as the paper in October: **Genc, B.**, Musingwini, C. and Katakwa, T.P. **(2014)** Estimating mine planning software utilisation for decision-making strategies in the South African platinum group metals mining sector, in *Proceedings of the 6th International Platinum Conference – "Platinum - Metal for the Future"*, Sun City, South Africa, 20-22 October 2014, Southern African Institute of Mining and Metallurgy:

"Abstract

This paper discuses a new methodology to define and measure mine planning software utilisation in the South African platinum group metals (PGM) mining sector within an evolving dataset framework. An initial dataset showing the mine planning

software providers, their corresponding software solutions as well as the software capabilities and information on the number of licences was collected and compiled in 2012 in an online database for software utilised in the South African mining industry. The database development and implementation was published in the Journal of the Southern African Institute of Mining and Metallurgy in 2013. In 2014 the dataset was updated with additional and new information.

Using the 2012 and 2014 time-stamps, a methodology for estimating the software utilisation was developed. In this methodology, three variables namely commodity (i), functionality (l) and time factor (t) were used to define and measure the software utilisation in order to ultimately inform decision-making strategies for software utilisation by various stakeholders. Using six different functionalities namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation, utilisation for PGM was measured. This paper presents the methodology employed for measuring the mine planning software utilisation. The work presented in this paper is part of a Ph.D. research study in the School of Mining Engineering at the University of the Witwatersrand.

Keywords: *PGM sector mine planning software utilisation, database, South African mining industry"*

9.1.3 Paper 3

Parts of Chapters 3 and 4 were compiled and submitted as the paper in July, 2014: **Genc, B.**, Musingwini, C. and Celik, T. **(2015)** Estimating mine planning software utilisation for decision-making strategies in the South African gold mining sector. *Journal of the Southern African Institute of Mining and Metallurgy*. The following is an extract of the paper abstract:

"Abstract

This paper discuses a new methodology to define and measure mine planning software utilisation in the South African gold mining sector within an evolving dataset framework. An initial dataset showing the mine planning software providers, their

corresponding software solutions as well as the software capabilities and information on the number of licences was collected and compiled in 2012 in an online database for software utilised in the South African mining industry. The database development and implementation was published in the Journal of the Southern African Institute of Mining and Metallurgy in 2013. In 2014 the dataset was updated with additional and new information.

Using the 2012 and 2014 time-stamps, a methodology for estimating the software utilisation was developed. In this methodology, three variables namely commodity (i), functionality (l) and time factor (t) were used to define and measure the software utilisation in order to ultimately inform decision-making strategies for optimal software utilisation. Using six different functionalities namely Geological Data Management, Geological Modelling and Resource Estimation, Design and Layout, Scheduling, Financial Valuation and Optimisation, utilisation for gold was measured. This paper presents the methodology employed for measuring the mine planning software utilisation. The work presented in this paper is part of a Ph.D. research study in the School of Mining Engineering at the University of the Witwatersrand.

Keywords: *Gold sector software utilisation, database, South African mining industry"*

9.1.4 Paper 4

Parts of Chapters 3 and 4 were compiled and submitted as the paper in October, 2014: **Genc, B.**, Musingwini, C. and Celik, T. **(2015)** Measuring "optimisation" software utilisation in the South African mining industry. *International Journal of Mining, Reclamation and Environment*. The following is an extract of the paper abstract:

"Abstract

This paper discusses a new methodology to define and measure utilisation of mine planning software. An initial dataset compiled in 2012 in an online database was updated in 2014 with additional and new information for software utilised in the South African mining industry. However, the online database only serves as a data repository and presentation platform. A methodology was required to analyse the data

in the database to infer extent of utilisation of the software to enable better decision-making strategies by various stakeholders that include mining companies, software providers, consulting companies and educational institutions. The new methodology uses three variables, namely commodity, functionality and time-stamp. In this paper, only one of the six mine planning software functionalities called the "Optimisation" functionality is applied on five selected commodities, using the 2012 and 2014 time-stamps. The work presented in this paper is part of a PhD research study in the School of Mining Engineering at the University of the Witwatersrand.

Keywords: *Mine planning; software utilisation; optimisation software; database; South African mining industry."*

9.2 The script details used to calculate mine planning software utilisation

Two Octave scripts are written to calculate software utilisation by using all three variables, namely, commodity, functionality and the time-stamp. These include scripts for utilisation and validation, respectively.

9.2.1 Utilisation

```
data;
commodityid2names;
softwareid2names;
functionalityid2names;
%id     site_id software_id     provider_id     number_licenses_installed
        timestamp_id
%siteSoftware
%commodity_id         site_id
%siteCommodity
%functionality_id     software_id
%functionalitySoftware

functionality = 3;
commodity = 3;
timeStampID = [1];
```

```
timeStamp{1} = "Sep 2012";
timeStamp{2} = "Apr 2014";

%find functionality vs software_id
software_id_index = find(functionalitySoftware(:,1)==functionality);
temp = functionalitySoftware(:,2);
software_id = temp(software_id_index);
%disp("software_id");
%disp(software_id);



%find commodity vs site_id
site_id_index = find(siteCommodity(:,1)==commodity);
temp = siteCommodity(:,2);
site_id = temp(site_id_index);
%disp("site_id");
%disp(site_id);

%creating table
f = zeros(length(software_id),1);
tempSiteSoftware = [];
for i = 1:size(siteSoftware,1)
        if(~isempty(find(site_id == siteSoftware(i,2))) && ~isempty(find(timeStampID
== siteSoftware(i,6))))
                tempSiteSoftware  = [tempSiteSoftware ; siteSoftware(i,:)];
        end
end
for i = 1:size(tempSiteSoftware,1)
        index = find(software_id == tempSiteSoftware(i,3));
        f(index) = f(index) + tempSiteSoftware(i,5);
end
commodityName = [];
for i = 1:length(commodityID2Names)
        if(commodityID2Names{i}{1} == commodity)
                commodityName = commodityID2Names{i}{2};
```

```
                break;
        end
end

disp(sprintf("%-16s%-32s","Commodity:",commodityName));

functionalityName = [];
for i = 1:length(functionalityID2Names)
        if(functionalityID2Names{i}{1} == functionality)
                functionalityName = functionalityID2Names{i}{2};
                break;
        end
end

disp(sprintf("%-16s%-32s","Functionality:",functionalityName));
disp(sprintf("%-16s%-32s","Time stamp:",timeStamp{timeStampID}));

disp("**********TABLE**********");
%displaying table
disp(sprintf("%-16s%-16s%-16s%-16s", "m", "software_id", "# of sites", "f(m)"));
for m = 1:length(f)
        disp(sprintf("%-16d%-16d%-16d%-16.4f", m, software_id(m), f(m),
f(m)/sum(f)));
end
disp("**********LEGEND**********");
for i = 1:length(software_id)
        for j = 1:length(softwareID2Names)
                if( softwareID2Names{j}{1}== software_id(i))
                        disp(sprintf("%-16d-
>%s",software_id(i),softwareID2Names{j}{2}));
                        break;
                end
        end
end
```

9.2.2 Validation

data;

siteIDGold = [
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144

2180
2182
2183
2184
2185
2186
2187
2188
2190
2191
2192
2193
2194
2195
2196
2197
2200
2202
2203
2204
2207
2208
2209
2210
2212
2351
2352
2354
2355
2358
2363
2648
2734
3375
3480

```
3486
3520
3532
3533
3536
3546
3551
3555
3564
3575
];

siteSoftwareTemp = [];
for i = 1:size(siteSoftware,1)
        index = find(siteIDGold  == siteSoftware(i,2));
        if(~isempty(index) && (siteSoftware(i,6)==1))
                siteSoftwareTemp = [siteSoftwareTemp; siteSoftware(i,:)];
        end
end
fid = fopen ("myfile.txt", "w");
for i = 1:size(siteSoftwareTemp,1)
        fdisp(fid, sprintf("%-8d,%-8d,%-8d,%-8d,%-8d,%-
8d\n",siteSoftwareTemp(i,1),siteSoftwareTemp(i,2),siteSoftwareTemp(i,3),siteSoftw
areTemp(i,4),siteSoftwareTemp(i,5),siteSoftwareTemp(i,6)));
end
fclose(fid);

z = [
2
2
2
2
2
2
2
```

2
4
4
4
4
4
4
4
4
4
5
5
5
9
9
9
9
10
10
10
10
10
10
10
10
10
10
10
10
10
12
12
12
12
12
12

13
13
14
14
15
15
15
15
16
16
16
16
16
16
16
16
19
19
19
20
22
22
22
23
23
23
25
25
26
28
28
33
34
35
36

38
39
40
41
66
68
69
70
70
71
71
71
76
76
76
76
76
76
76
76
76
77
77
77
77
77
77
77
77
77
78
80
81
81
83
83

83
83
83
83
83
83
83
97
97
97

97
97
97
97
98
98
98

```
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
99
101
103
];
```

9.3 The script details used to predict time-series analyses

Number of Octave scripts are written to model the dataset. These include scripts for Backpropagation algorithm, Feedforward ANN, forecasting, simulations, testing and training, respectively.

9.3.1 Backpropagation

```
function [Whi,Woh] = backpropagation(x,h,e,alpha,Whi,Woh)
%update Woh
Woh = Woh - alpha * e * h;
%update Whi
Whi = Whi - alpha * e * (Woh .* (h .* (1-h))) * x';
end
```

9.3.2 Feedforward ANN

```
function [o,h] = feedforward(x,Whi,Woh)
Whi(1,:) = 0;
h = activation(Whi*x);
h(1) = 1;
o = activation(h'*Woh);
end

function [g] = activation(x)
g = 1./(1+exp(-x));
end
```

9.3.3 Forecasting

```
close all;
clear all;
clc;
```

```
randn('seed',7);
rand('seed',7);
M = 100; %defines the number of Gaussian with random parameters used in
simulation
K = 4; %defines the number of software operating on a specific commodity and
functionality
tSampling = 0.01; %sampling rate
color = [
    0 0 1;
    0 1 0;
    0 1 1;
    1 0 0;
    1 0 1;
    1 1 0];
color = [color; color/2];
color = [color; color/2];
[f, t] = SimulateData(tSampling, M, K);
figure, hold on;
linewidth = 2.0;
labels = {};
for k = 1:K
    c = color(k,:);
    plot(t,f(k,:),'color',c,'linewidth',linewidth);
    labels{k} = ['u_{' num2str(k) '}^{(t)}'];
end
xlabel('t');
ylabel('u^{(t)}');
legend(labels,'location','best')
box;
ylim([0 1.0]);
hold off;

figure,hold on;
to = t;
ft = [];
```

```
fp = [];
for k = 1:K
    y = f(k,:);
    ut = [];
    up = [];
    tt = [];
    train_window = 10;
    test_window = 1;
    window_width = train_window + test_window;
    ni = train_window; %number of nodes in the input layer
    nh = 2*train_window; %number of nodes in the hidden layer
    learning_rate = 0.01;
    for t = 1:test_window:length(y)-test_window-window_width
        Xwindow = y(t:t+window_width-1)';
        twindow = to(t:t+window_width-1);
        X = [1; Xwindow(1:train_window-1)];
        Y = Xwindow(train_window);
        [Whi,Woh] = train(X,Y,ni,nh,learning_rate);
        XXwindow = Xwindow(1:train_window);
        for kk = 1:test_window
            X = [1; XXwindow(1+kk:1+kk+train_window-1-1)];
            [u] = feedforward(X,Whi,Woh);
            XXwindow = [XXwindow;u];
            ut = [ut Xwindow(1+kk+train_window-1-1)];
            tt = [tt twindow(1+kk+train_window-1-1)];
            up = [up u];
        end
        %disp(t);
    end
    c = color(k,:);
    plot(tt,ut,':','color',c,'linewidth',1.5);
    plot(tt,up,'color',c,'linewidth',1.5);
    labels{2*k-1} = ['u_{' num2str(k) '}^{(t)}'];
    labels{2*k} = ['\^{u}_{' num2str(k) '}^{(t)}'];
    disp(['k = ' num2str(k) ', RMSE = ' num2str(sqrt(mean((ut-up).^2)))]);
```

```
    disp([tt' ut' up']);
    ft(k,:) = ut;
    fp(k,:) = up;
end
xlabel('t');
ylabel('u^{(t)}');
legend(labels,'location','northwest')
box;
ylim([0 1.0]);
hold off;
```

9.3.4 Simulations

```
%close all;
clear all;
randn('seed',17);
rand('seed',17);
M = 100;
K = 4;
tSampling = 0.01;
color = [
    0 0 1;
    0 1 0;
    0 1 1;
    1 0 0;
    1 0 1;
    1 1 0];
color = [color; color/2];
color = [color; color/2];
[f, t] = SimulateData(tSampling, M, K);
figure, hold on;
linewidth = 2.0;
labels = {};
for k = 1:K
    c = color(k,:);
```

```
    plot(t,f(k,:),'color',c,'linewidth',linewidth);
    labels{k} = ['u_{' num2str(k) '}^{(t)}'];
end
xlabel('t');
ylabel('u^{(t)}');
legend(labels,'location','best')
box;
ylim([0 1.0]);
```

9.3.5 Testing

```
f = 2;
ts = 0.01;
t = 0:ts:2;
to = t;
y = (cos(2*pi*f*t)+1)/2;
ut = [];
up = [];
tt = [];
train_window = 7;
test_window = 3;
window_width = train_window + test_window;
for t = 1:test_window:length(y)-test_window-window_width
    Xwindow = y(t:t+window_width-1)';
    twindow = to(t:t+window_width-1);
    X = [1; Xwindow(1:train_window-1)];
    Y = Xwindow(train_window);
    [Whi,Woh] = train(X,Y,train_window,2*train_window,0.01);
    XXwindow = Xwindow(1:train_window);
    for k = 1:test_window
        X = [1; XXwindow(1+k:1+k+train_window-1-1)];
        [u] = feedforward(X,Whi,Woh);
        XXwindow = [XXwindow;u];
        ut = [ut Xwindow(1+k+train_window-1-1)];
        tt = [tt twindow(1+k+train_window-1-1)];
```

```
        up = [up u];
    end
    disp(t);
end
figure,plot(tt,ut,'b');
hold on
plot(tt,up,'r');
```

9.3.6 Training

```
function [Whi,Woh] = train(X,Y,ni,nh,alpha)
Whi = randn(nh,ni);
Woh = randn(nh,1);
converged = 0;
while(converged == 0)
    e_max = -inf;
    Whip = Whi;
    Wohp = Woh;
    for n = 1:size(X,2)
        x = X(:,n);
        [o,h] = feedforward(x,Whi,Woh);
        e = (o-Y(n));
        [Whi,Woh] = backpropagation(x,h,e,alpha,Whi,Woh);
        if(abs(e)>e_max)
            e_max = abs(e);
        end
    end
    e_max = max(max(max(abs(Whip-Whi))),max(max(abs(Wohp-Woh))));
    %disp(e_max);
    if(e_max < 1e-4)
        converged = 1;
    end
end
```