Competitiveness and performance prediction of surface coal mining

by

Mussa Daniel Budeba

Presented as partial fulfilment for the degree

Philosophiae Doctor (Mining)

in the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT, AND INFORMATION TECHNOLOGY
DEPARTMENT OF MINING ENGINEERING
UNIVERSITY OF PRETORIA

July 2016

DECLARATION

I hereby declare that this thesis is my own unaided work. It is being submitted for the degree Philosophiae Doctor (Mining) at the University of Pretoria, Pretoria. It has not been submitted before for any degree or examination in any other University. This document represents my own opinion and interpretation of information obtained from literature survey.

________________________________(signature on copy of final document)

MUSSA DANIEL BUDEBA

Dated this____________________day of ____________________

UNIVERSITY OF PRETORIA

FACULTY OF ENGINEERING, BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

DEPARTMENT OF MINING ENGINEERING

The Department of Mining Engineering places great emphasis upon integrity and ethical conduct in the preparation of all written work submitted for academic evaluation. While academic staff teaches you about systems of referring and how to avoid plagiarism, you too have a responsibility in this regard. If you are at any stage uncertain as to what is required, you should speak to your lecturer before any written work is submitted.

You are guilty of plagiarism if you copy something from a book, article or website without acknowledging the source and pass it off as your own. In effect you are stealing something that belongs to someone else. This is not only the case when you copy work word-by-word (verbatim), but also when you submit someone else's work in a slightly altered form (paraphrase) or use a line of argument without acknowledging it. You are not allowed to use another student's past written work. You are also not allowed to let anybody copy your work with the intention of passing if of as his/her work.

Students who commit plagiarism will lose all credits obtained in the plagiarised work. The matter may also be referred to the Disciplinary Committee (Students) for a ruling. Plagiarism is regarded as a serious contravention of the University's rules and can lead to expulsion from the University. The declaration which follows must be appended to all written work submitted while you are a student of the Department of Mining Engineering. No written work will be accepted unless the declaration has been completed and attached.

I (full names) ______________________________

Student number ______________________________

Topic of work ______________________________

Declaration

1. I understand what plagiarism is and am aware of the University's policy in this regard.
2. I declare that this ________(e.g. essay, report, project, assignment, dissertation, thesis etc) is my own original work. Where other people's work has been used (either from a printed source, internet or any other source), this has been properly acknowledged and referenced in accordance with departmental requirements.
3. I have not used another student's past written work to hand in as my own.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature ______________________________

Acknowledgments

I am grateful to the Lord God Almighty for taking care of me during my research period at the University of Pretoria. I would also like to take the opportunity to acknowledge the following individuals and institutions who in one way or another contributed to the completion of my research:

1. The University of Dar es Salaam management for financial support of my PhD research through its staff development programme under the support of world bank project.
2. My supervisor Prof. R.C.W. Webber-Youngman, Head of the Department of Mining Engineering for financial support and advice, and my co-supervisor Prof. J.W. Joubert, for the financial support , constructive ideas, advice, time for reading report drafts and guidance of my research at the University of Pretoria.
3. All those people who have published scholarly works that I have cited and referenced in my research - their work contributed to the concepts and ideas that I used in my own research.
4. Mr Johan Hagger, for his advice and technical help regarding the feasibility studies of large surface mines.
5. Staff members of the Department of Mining Engineering, including Professor Jan du Plessis, Dr Bertie Meyer, Mr Clive Knobs, Mr Wolter de Graaf, Mr Jannie Maritz, Mrs Gudmanz, Mr Johan Lou, Mr Eugen Preis and Mr Elphus Ramashau. Their encouragement highly motivated me to work hard at my research work.
6. My wife Christina, for her patience and tolerance during the difficult time of raising three sons alone in my absence for studies. My mother, Ester Daniel for her consistent support and encouragement and her prayers for my research work. Thanks also to others in the Daniel family: Mr Charles K. Budeba, Mr Paul Daniel, Mr Tito Daniel and Miss Teleza Daniel for their encouragement.
7. Dr Afrael Sarakyikya, Dr Jestina Katandukila, Dr Silipius Mbawala, Dr George Fasha, Eng. Jean Rene, Eng. Julius Komba, Mr Abdul Kenan and Mr Linus Kinyondo for their encouragement of my research.

Contents

List of Figures

List of Tables

List of Algorithms

Acronyms

AIC Akaike Information Criterion

ANN Artificial Neural Network

ATCOM Arthur Taylor Colliery Opencast Mine

BCC Banker–Cooper–Charnes

BESR Break Even Stripping Ratio

BLUE Best Linear Unbiase Estimator

BP British Petroleum

CAGR Cumulative Annual Growth Rate

CANMET Canada Center for Mineral and Energy Technology

CAPCOSTS Capital Cost

CAPEX Capital Expenditure

CCR Charnes–Cooper–Rhodes

CES Cost Estimation System

COLS Corrected Ordinary least Squares

CRS Constant Return to Scale

CSLE Combined System for Local and Export

CV Calorific Value

DEA Data Envelopment Analysis

DMU Decision Making Unit

ECMS Export Coal Mine Supply

EP	Emission Penalty
FOB	Free on Board
FP	Fractional Program
GDP	Gross Domestic Product
GLM	Generalised Linear Model
IEA	International Energy Agency
IMF	International Monetary Fund
LP	Linear Program
LCMS	Local Coal Mine Supply
LG	Lerchs–Grossmann
MOLS	Modified Ordinary least Squares
ND	Non-Discretionary
NPV	Net Present Value
NR	Net Revenue
OECD	Organisation for Economic Cooperation and Development
OEM	Original Equipment Manufacturer
OLS	Ordinary Least Squares
PFP	Partial Factor Productivity
RMG	Raw Material Group
RMSE	Root Mean Square Error
RMSR	Root Mean Square Residuals
ROM	Run-of-Mine
RSS	Residual Sum Square
SE	Scale Efficiency
SFA	Stochastic Frontier Analysis

SI	Sample and Iterate
SNL	Savings and Loan
SR	Stripping Ratio
TFP	Total Factor Productivity
USA	United States of America
USGS	United States Geological Survey
USBM	United States Bureau of Mines
VRS	Variable Return to Scale
WCA	World Coal Association
WEC	World Energy Council

Abstract

Competitiveness and performance prediction of surface coal mining

by

Mussa Daniel Budeba

Supervisor	:	Prof. R.C.W. Webber-Youngman
Co-supervisor	:	Prof. J.W. Joubert
Department	:	Mining Engineering
University	:	University of Pretoria
Degree	:	Philosophiae Doctor (Mining)

The purpose of this research is to formulate mathematical models for assisting the management of either a new or operating surface coal mine to assess its competitiveness relative to other coal producers for a given market of thermal coal.

As an alternative of efficiency measurement to provide a new way to assess better the competitiveness of surface mining, Data Envelopment Analysis (DEA) method is proposed. DEA uses linear programming to determine the relative efficiencies of (competing) mines, each referred to as a Decision Making Unit (DMU). In this research, the methodology applied involves three stages: First, applying DEA to formulate the mathematical models basing on the structure of coal extraction, processing and supply to the markets. Second, evaluate the models performance and illustrate the use case, and thirdly develop predictive model for the efficiency and performance of a new mine.

Three DEA models were developed, each representing a specific configuration of extraction, processing and sale of coal to the markets. The main model, referred to as Combined System for Local and Export (CSLE), supplied both the local and export markets. Two special cases, referred to as Local Coal Mine Supply (LCMS) and Export Coal Mine Supply (ECMS) respectively, looked at the individual markets in isolation.

The results from the numerical illustrations of the application of the DEA models showed that the models were able to discriminate between the efficient (best practice) and inefficient mines. This provides a quantitative measure that mining companies can use to benchmark themselves against other competitors in a multi-dimensional manner. Also, the proposed

method allows for generating realistic, quantitative targets for those DMUs that are considered *inefficient*. After formulating the three DEA models, use cases are presented for the CSLE model to demonstrate the significance of the proposed model for decisions making.

Predictive models for technical efficiency and mine performance developed in this thesis, target new mining operations wanting to enter the market. A statistical method known as *supervised learning* was employed in this case. It was found that the predictor variables in the model can only explain 54.5% of the variation in technical efficiency. To test the prediction accuracy, the mining entities were separated into training and test sets. On the test set, the model predicted efficiency scores within ±20% of the actual (known) values. To improve the performance of this model, this thesis suggests investigating the influence of qualitative variables on mining efficiency. Such qualitative variables may include worker morale, work satisfaction and salary disputes.

Mine planning is non-trivial as it requires various perspectives and involves the interdependence of many variables with different units of measure. This research is significant as it provides mining management with a sound and rigorous model to handle the multiplexity of the decision variables. The quantitative approach provides for evidence-based decision support where large capital amounts are at risk. Mine planning parameters can be evaluated taking the mine's particularities into account before proceeding to the production stage. The DEA approach is useful both for current mining operations to evaluate its competitiveness in given markets, as well as new mining operations who need to anticipate the type and quantity of capital to invest given their project characteristics.

Therefore, the mine management can use the models to determine the optimal technical inputs such as capital, labour and the stripping ratio while considering mine-specific challenges that influence the competitiveness of the project, such as the location of the mine from the market and coal seam thickness that can not be controlled.

Keywords: DEA, technical efficiency, coal mine competitiveness, predictive models

Chapter 1

Introduction

Coal is one of the major sources of energy, contributing approximately 29% of the total primary energy. This is only after fuel oil, which accounts for 31% (International Energy Agency 2016). Coal is used to generate 41% of the electricity in the world, and this is predicted to increase to 46% by 2030 (Schernikau 2010). Meanwhile, the demand for energy is expected to increase at an average Cumulative Annual Growth Rate (CAGR) of 1.7% per annum in the period between 1990 and 2030 (Schernikau 2010). This provides incentive to increase coal production and start new coal mines. However, new mines cannot be started unless they are profitable.

To start a profitable new coal mine is challenging as it requires that once the mine begins operation, it produces tonnage for selling in a competitive market with many external factors that can hamper its profitability. Suppose a group of surface coal mines located in Queensland, Australia and Mpumalanga, South Africa produce and supply coal to the international market. A fixed price for coal of a specified energy content is offered. Each mine has its own production rate based on its coal characteristics, such as the varying thickness in the ground, the calorific value, which is the amount of energy present when the coal is burnt for electricity and price of coal offered by the market. The mines apply different technical variables in producing coal, such as the type and quantity of capital and the stripping ratio, which is the quantity of overburden to be removed per tonne of coal extracted. Some of these mines achieve their production target using minimum technical variables [1], while operating with good safety records and managing the impact of the operations to the environment such as that of water pollution. These mines are efficient and cost-effective. They considered competitive and to exhibit best practices.

On the other hand, some mines are inefficient. They use excess inputs to achieve their production target, and thus they cannot survive as a competitive business. Assume that a

[1] This refers to the optimum inputs used to produce target coal tonnages. The operation achieve its target by using less quantity of controllable inputs such as capital spending

new mine located in South Africa, say Mine **A**, is about to start production. It will face the challenge of determining whether it is competitive locally and internationally given its unique characteristics, such as the thickness of the coal in the ground. If mine production starts and is inefficient, it will fail to generate a return on investment and will not survive in the business. It remains challenging for a new mine to predict its competitiveness among the existing producers or to identify the best practices and position itself competitively.

Examples of critical problems such as cost overruns in both existing and new mines are known in the mining industry. That is, the actual cost of the project relative to the estimated costs. For example, a global study conducted by Ernst & Young Global Limited (2015) found cost overruns on average of 62% in the mining industry. A total of 108 mega projects at different stages were investigated in October 2014, some at the initial stage of operation and others at the delivery stage. The study included both coal and metal mines, such as copper, iron ore, gold, and nickel. Some reasons for these overruns included project management factors, stakeholder conflicts, resource constraints, and regulatory and policy-related challenges, together with an unfavourable external environment such as commodity price movements.

A new mine will not operate in isolation. It has to consider the influence of the other producers supplying coal to the same market. In addition, management of a mine project should be conducted properly, from the initial stage of prospecting to production and delivery, by making decisions regarding the technical inputs needed, such as capital and the number of employees, to make the mine competitive relative to other producers and to generate returns for the investments made.

1.1 The process for starting a new mine

Before producing coal and selling it, mines go through a staged process: prospecting, exploration, development, exploitation, and reclamation. These are known as stages in the life of a mine (Hartman and Mutmansky 2002). The first four stages ensure the delivery of the product, whereas reclamation ensures proper closure of the mine. The first four stages are shown in Figure 1.1.

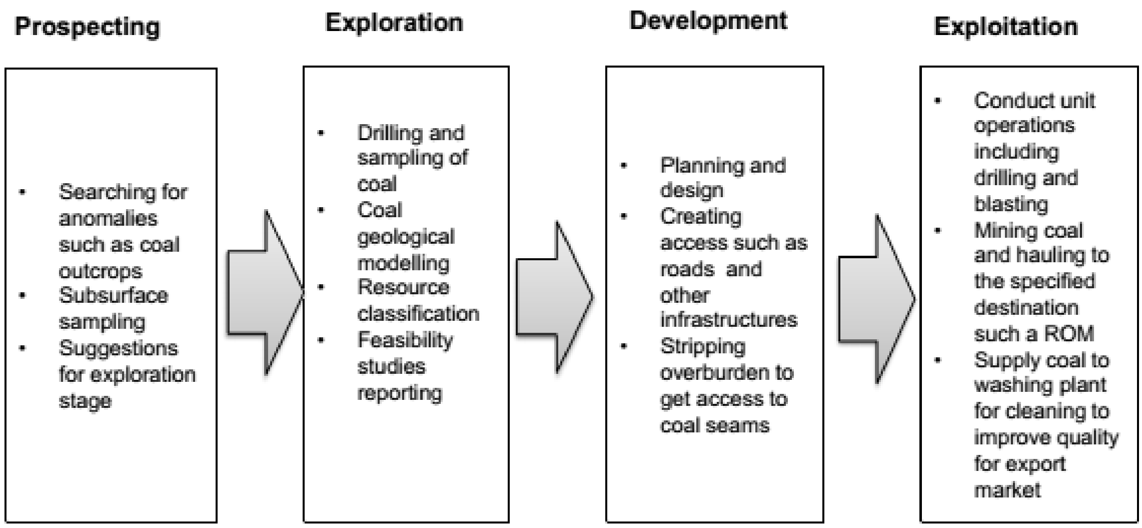

Figure 1.1: Stages in the life of a mine (Author‘s construction).

Prospecting involves searching for the occurrence of coal. It involves looking for surface-appearing coal known as *outcrops*. Samples of coal are taken from the surface and suggestions for the exploration stage are made. *Exploration* involves drilling thin deep holes into the coal in the ground. Drill hole data such as location of holes, Calorific Value (MJ/kg), Ash(%) and lithology are compiled in order to estimate the tonnage and quality of the coal seam available on the property. This is also known as *resource modelling*. After the exploration stage, *feasibility studies* are carried out to evaluate whether the project is viable. This stage includes estimating the resource and its classification, determining the reserve, estimating the production rate and costs, and conducting environmental studies on the effect of the project on the environment during and after production.

There are three major stages of feasibility studies. The first stage is a conceptual study, in which little information is known about the deposit of interest. Cost estimates are initially carried at an accuracy of ±30% (Hustrulid and Kuchta 2006; McCarthy 2013). For example, if the actual cost of the project according to the conceptual study is US$50 million, then the cost estimate may vary between US$35 million and US$65 million. The second stage is a pre-feasibility study, which involves further data gathering as compared to the conceptual stage. During this stage, a better understanding of the mine project is obtained, more samples are collected, and the cost estimates are done at an accuracy of ±20% (Hustrulid and Kuchta 2006). The third stage is a detailed feasibility study, which requires more details about the project, even more samples of the deposit of interest, mine plans, mine layout, and a process plan about the project. By this stage, most of the engineering analyses have been completed and the accuracy of the project estimates will be approximately ±10% (Hustrulid and Kuchta 2006).

The development stage brings the mine into full production. It includes planning, design, and construction. During this stage, the feasibility study is reviewed and implemented. For example, the plan at this stage may change depending on whether new information becomes available and changes in the economics of the project. Activities at this stage include acquiring funds for the project, creating access to the coal for extraction, and installing infrastructures such as washing plants, to name a few.

Exploitation is the actual production, where the coal is extracted and shipped or delivered to the washing facility for cleaning to sell for export and local power plants. Exploitation is the production cycle, which includes drilling, blasting, loading, and hauling. During drilling, holes are created in the ground as designed by the engineer. The holes are charged with explosives and then blasted to fragment coal. Loading follows blasting, using machines such as shovels to excavate the fragmented coal, which is then loaded into trucks. The loaded trucks transport the coal from the pit to either a destination prepared at the mine to stockpile the coal for local sales or to the washing plant, where the coal is cleaned to remove wastes such as ash to improve its quality for export. In some other mines blasting such that of overburden is not done due to the ground being soft such that it does not require blasting, dragline excavate the overburden and dump in the mine out area and shovel loads coal into trucks hauling to the specified destination.

Despite the clear process indicated in Figure 1.1, Mohnot et al. (2001) claims that new mines are faced with the problem of combining technical design and economic parameters to generate value for their stakeholders. In addition to that, safety and environmental issues affect the realization of value from new and operating mines. Mines operation practices are required to minimize fatalities and manage impact to the environments generated by the effect of mining activities.

1.2 Mining as a turbulent operation

The mining business is a turbulent operation in which some mines are successful in achieving their production target using minimum technical inputs, whereas others do not perform well. Most mines face challenges that cause uncertainty in production delivery. These challenges includes those that mine management can control, called discretionary variables, and those that it cannot control, known as non-discretionary variables. Both types of variables affect the mining revenue. They can be summarized as shown in Figure 1.2.

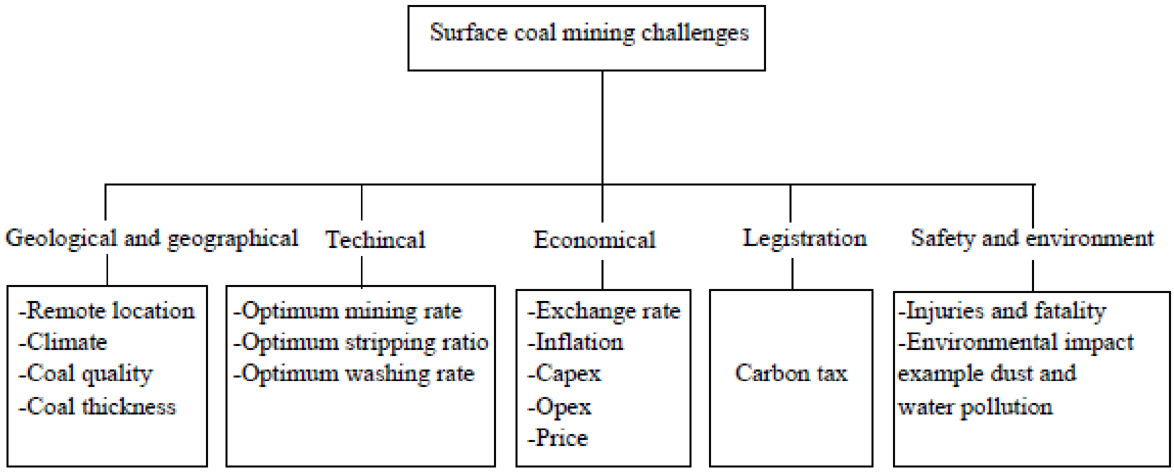

Figure 1.2: Supply challenges for surface coal mines.

1.2.1 Controllable variables of a new mine

Controllable variables are technical variables that the mine management has control over. For example, mine management can determine the production rate, the washing rate of coal, the design of the pit to generate maximum profit, the number of employees needed, and the amount of capital needed for the initial investment. These can be changed at the discretion of the mine management, and thus they are also known as discretionary variables. For example, a mine can increase capital or reduce the number of employees at its discretion.

1.2.2 Non-controllable variables of a new mine

Each coal project has unique characteristics that mine management cannot control, such as geographical and geological, economic, and legislative variables. These differentiate one mine project from another, and they are known as non-discretionary variables.

Geographical variables include the location of the mine. Some coal mines are close to their markets, whereas others are located in regions very far from their markets. For a mine that is far from a port, the distance for transporting coal affects the tonnage that can be transported, and the transportation costs for a mine located in a remote area with poor infrastructure are higher than those for a mine close to its market. Climate is another variable that influences productivity. Some mines are located in regions with high rainfall, which interferes with production and requires frequent pumping of water. Geological variables include the calorific value of the coal and the presence of ash, which requires cleaning the coal to improve the energy content. In addition, some coal deposits are thicker, which makes coal extraction easy, whereas others are thinner, which requires special equipment for selectivity (Shafiee and Topal 2012; Shafiee et al. 2009; Chan 2008; Haftendorn et al. 2012; Höök et al. 2010).

Economic variables include price, exchange rate, and inflation. These change from time to time in a given country, and they affect mine projects within the respective country. The other factor is legislation, which has to do with environmental issues such as management of dust, water quality and control of carbon emissions from mine operations. For example, mines are required to pay a carbon tax if they exceed the emission limits established by the countries having the policy for carbon emission control in which the mines operates (Gordon 1976; Schneid and Torries 1991). This affects the revenue stream of these coal mines.

1.2.3 Measuring the competitiveness of coal mines

Measuring the competitiveness of a new coal mine is a challenge that requires investigation. The existing method measures the competitiveness of operating mines based on a cost curve. The cost curve refers to the curve developed by collecting cash costs for each mine involved in production and plotting those costs against the cumulative production rates. Mines that are in the lower part of the curve are considered competitive, whereas those in the higher part of the curve above a given price are operating at a loss. To illustrate the use of the cost curve, consider Figure 1.3.

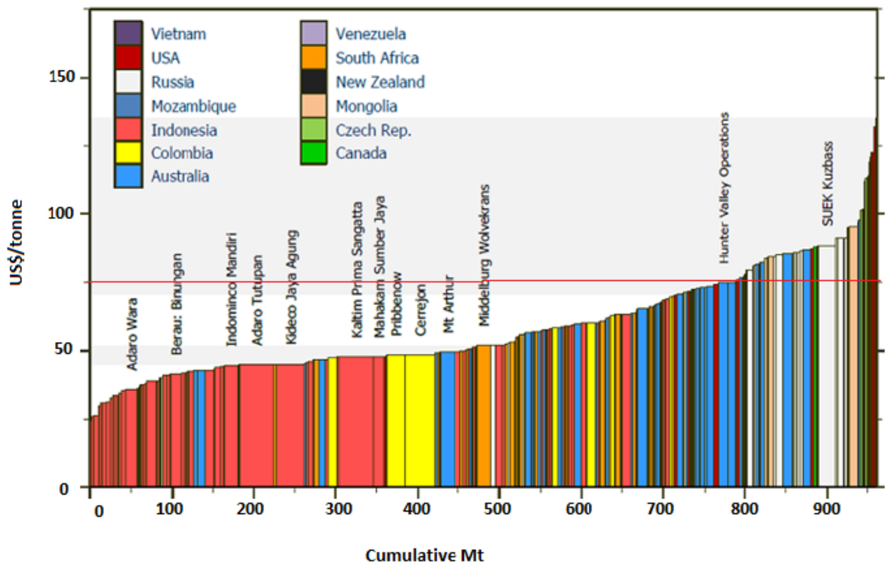

Figure 1.3: Example of thermal coal mine cost curve. Adapted from Morgan (2014).

For example, assume that the price of thermal coal is US\$75/tonne. All mines with costs above this price are operating at a loss, whereas those with costs below this price are generating a profit. Taking the example of SUEK Kuzbass, this company cannot survive in the business because its cost of approximately US\$125/tonne is above US\$75/tonne, and thus is operating at a loss. On the other hand, Hunter Valley Operations is operating at a cost of US$75/tonne, which is equal to the market price of coal, and hence it is not generating a profit; this company will also not survive in the business.

The use of the cost curve for a new or planned mine is for cost-comparison purposes. It cannot locate the position of a new mine competitively. Mine management must always compare their projects against those of lower-cost producers, which is often uncertain. It also does not make the mines efficient because each mine has unique non-controllable variables that require optimal choice of the controllable variables when commencing a new mine. The mining industry needs a method that can help new mining companies assess their plan and make informed decisions about the project before commencing. In addition, the cost curve cannot clearly indicate whether a mine is using excessive input of its resources to achieve the same tonnage as other mines of the same scope.

In addition, mine projects still have cost and production issues. For example, Bullock (2011) highlighted some mine projects with problems. Studies of mines operating between 1965 and 2002 found that the lowest cost overrun was 22%, whereas the highest was 35%. Another study of 60 projects found that 58% had overruns of between 15% and 100% of the capital cost. A due-diligence investigation of one mine project showed that the mine's operating costs were set 17% too low and not estimated, working capital was underestimated by at least 75%, and contingency was underestimated by 10%.

Other examples have been identified in reviews of megaprojects. In 2011, it was found that over a period of two years, approximately 70% of megaprojects, including mines, had greater than 25% cost competitiveness, apart from cost overruns, schedule overruns, and operational problems (PwC 2012; Merrow 2011). When the actual outcome of the real operation was compared to the estimates, among the causes of these deviations was a poor understanding of the project risks and poor estimation techniques.

Van Aswegen and Koster (2008) also conducted qualitative research into the South African mining industry. The study involved the use of 144 questionnaires sent to middle and senior management and the directorate of companies such as Anglo American, Anglo Platinum, Murray and Roberts, Impala Platinum, and other consulting companies. Of these, 49% responded, and from the responses it was found that there was a gap between the feasibility study and the actual outcome during project execution, which was attributed to cost deviations, schedule deviations, a number of project scope changes, and inaccurate prediction of operational performance.

In addition, cost overruns have been identified in mining for decades when as-built capital costs are compared to the feasibility study capital cost estimates. For example, a study of 18 mining projects operating from 1965 to 1981 found average cost overruns of 33% when compared to the respective feasibility studies. Another study of 60 mining projects operating from 1980 to 2001 found average cost overruns of 22% and half the projects had overruns of more than 20%. The reason was established as optimistic and poor cost estimation (Noort and Adams 2006). Other examples of mining cost overruns are shown in Table 1.1. Ranging from 10% to 100%, cost overruns affect the expected net value of the project because of the extra costs incurred.

Table 1.1: Cost overruns (Noort and Adams 2006)

Project	**Company**	**Feasibility**	**Actual forecast overrun**
Ravensthorpe/ YabiluExpansion	BHP Billiton	A$1.4 billion	30%
Spence (Chile)	BHP Billiton	US$990 million	10%
Telfer Mine	Newcrest	A$1.19 billion	17.5%
Stanwell Magnesium	AMC	A$1.3 billion	30%
Boddington	Newmont	A$866 million	100%
Goro Project (Indonesia)	Inco	US$1.45 billion	15%
Prominent Hill Project	Oxiana	A$350 million	51%

1.3 Research question

There are many surface mines producing thermal coal for sale both locally and internationally. Each mine produces equal or different quantities of coal as compared to other mines while using similar inputs that can also be varied. The mines supply coal to the same markets, which offer the same price for a specific standard energy content in the coal. These mines realize their level of competitiveness once they are in production. On the other hand, new mines before starting production cannot predetermine their competitiveness relative to existing producers. Instead, new mines apply their estimated technical variables to start operating while being uncertain whether they can achieve best practices and competitiveness. Cost curves can only tell the relative cash cost of production of an operating mine; however, it fails to position a new mine competitively or to assess its use of inputs to achieve the target coal tonnage. There is no method to help mine management apply an end-to-end analysis to see if the mine supply system is operating according to best practices nor to help them know how to improve the mine's processes to make it more competitive. This leads to the following main research question:

How can a new surface mine producing thermal coal evaluate its competitiveness *relative to other operating coal mines considering each mine's specific variables those that mine management can control and those it cannot, given the market of thermal coal?*

The main research question is divided into the following secondary questions:

- *What model representing the structure of the mine coal supply can be used to measure the relative technical efficiency of a mine considering variables that mine management can control and those it cannot?*
- *How can it be determined which are best practice surface mines producing thermal coal given the unique mine variables?*
- *What models can be used to predict the technical efficiency and estimate production for competitiveness of a new mine?*

1.4 Research design

The current research is quantitative in nature. It will generate two models, one for measuring the technical efficiency of a thermal coal mine and the second for predicting the efficiency of a surface coal mine. The process considered in developing the models is presented in Figure 1.4.

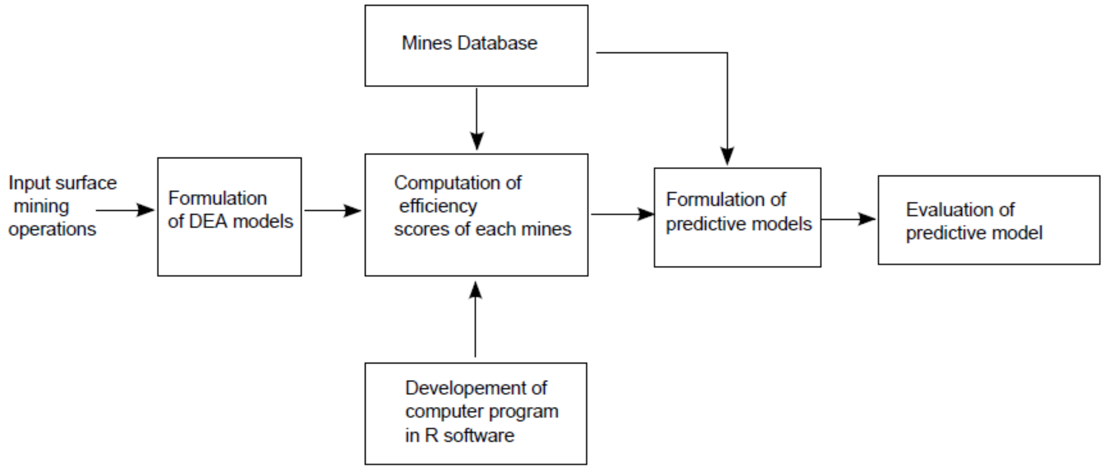

Figure 1.4: Approach for model development.

1.4.1 The model for technical efficiency

The model for technical efficiency will be used to estimate the efficiency of mines supplying coal for export and to local markets. In this research, the model is denoted as Combined System for Local and Export (CSLE). The model will be formulated using Data Envelopment Analysis (DEA), which is a method that uses linear programming to evaluate the relative efficiency of a group of entities considering their inputs and outputs. The reasons for using DEA for this research are, first, it can be used to assess the overall efficiency of each subsystem, unlike the parametric method based on regression, which considers inputs and a single output. Second, it does not require the assumption of a specific production function. Instead, it measures multiple inputs and outputs of either the whole system or a subsystem, such as those employed for mines. Third, it can be applied even if there is insufficient data. Fourth, the mine is involved in evaluating the efficiency of each unit, and hence its influence is significantly realized.

The CSLE model is composed of interlinked functioning mining operations considered as subsystems, each of which has inputs and outputs. The subsystems are extraction, washing, and port. The washing plant is an intermediate subsystem that has two forms of inputs: one is the output from extraction and the other is the input at the beginning of the washing plant and generate the output. The port has two forms of input, one at the beginning of the port and the other is the output from the washing plant. The input and output of each subsystem are used in the DEA method to formulate the efficiency of the overall mining system and of each subsystem. Special-case DEA models will be generated from the CSLE model, one for the supply of coal for export only and the other for the local market only.

Computer code will be developed in R-software to solve the CSLE and special-case models. The outputs from solving the models will be the overall system and subsystem efficiency scores for each mine assessed for its coal production to generate revenues.

1.4.2 The predictive models for a new mine

Two models will be formulated using the regression method of supervised learning type. The technical efficiency predictive model will use the efficiency scores obtained from the CSLE model, and a relationship will be developed against all the input variables in all mining subsystems. The predictive model for production will be developed based on the relationship between all the input variables in mining operation; the production rate will be the dependent variable.

1.5 Research methodology

The methodology used for this research consists of two parts: modelling and evaluation of the models.

1.5.1 Modelling

- Formulating the mathematical equations of the overall efficiency of a mine supply system using the DEA method considering the flow of inputs and outputs from interlinked extraction, washing, and port (market) subsystem structures.
- Formulating the predictive models using the number of inputs and the output variables of the mine supply system, and applying the supervised learning regression method to formulate the models.

1.5.2 Evaluating the models

- Collecting both controllable and non-controllable variables from the Raw Material Group (RMG) database and from annual mining reports for coal commodities.
- Creating a correlation matrix of relationships among the variables of the mines collected.
- Using the multivariate simulation technique to generate mines mimicking existing mine projects by reproducing the correlation matrix of the original data collected.
- Developing code using software developed by R Core Team (2015), applying it to solve the DEA models using the simulated data, and interpreting the results.
- Using statistical packages for regression-supervised learning in R-software to specify the parameters of the predictive models using training datasets and evaluating the performance of the models using the test datasets

1.6 Structure of the thesis

The literature review pertaining to the estimation of the performance and methods of evaluating the technical efficiency of a new surface coal mine is covered in Chapter 2. The chapter explains the available literature regarding estimation of the production rate and costs of a new mine project, identifying the shortcomings, discussing the available methods for evaluating the efficiency of the new mine, and justifying the choice of the DEA method for this research. Chapter 3 examines the source of the data and the method of simulating the data to generate mining supply systems for use in answering the research questions stated in this study. Chapter 4 details the formulation of models for measuring the technical efficiency of surface coal mining supply systems using the DEA method. The chapter concludes with one main model, referred to as CSLE, and two special-case models. The evaluation of the application of the CSLE and two special-case models is detailed in Chapter 5. Chapter 6 discusses the formulation of the predictive models for the technical efficiency and performance of a mine. The chapter illustrates specification of the parameters of the model and evaluates the capability and performance of the predicative models considering their application to new mines. Chapter 7 concludes the research work covered from Chapters 1 to 6. It reviews the research questions and describes the findings with regard to the questions stated. It highlights the answers to each question by recapitulating the specific chapter answering the question. Moreover, the chapter discusses the research contribution and limitations of the current research, and it suggests areas requiring further investigation that the candidate believes will advance the knowledge contributed by this research work. Finally, it gives recommendations for the study conducted.

Chapter 2

Performance and efficiency of a new mine project

Mining operations, like other production industries, measure the achievement of their objectives for the business by means of performance measurements. Every industry defines performance based on the nature of its operations and uses some type of metric to measure these achievements. Examples of metrics used include financial metrics, such as profits, and sales per tonne produced by the operation. However, Bourne et al. (2007) claim that there is no universally agreed upon definition of performance. Markovits-Somogyi (2012) defines performance as a quantifiable, data-like result that can be reached by someone or something in the course of work or other professional activity in a given time frame. So, quantification of the achievement of objectives reflects the performance of the business.

In the mining context, Hustrulid and Kuchta (2006) state that performance can be reduced to throughput and recovery. Throughput refers to the tonnage of rock containing minerals that are produced and fed into the plant, whereas recovery is the percentage of the mineral content in rocks that are extracted as final products. Mining companies include cost effectiveness as a measure of performance to attain corporate objectives such as profitability.

Different operating mines have variable throughputs, recovery, and costs that depend on the methods of mining and washing the coal. For a given mining method, the use of equipment such as trucks and shovels to mine coal and overburden in a specific time helps determine the mining performance. Likewise, for a new mine, understanding the available surface mining methods and selecting the appropriate one is essential for ensuring that the estimated production rate is achieved during operation.

2.1 Selecting a surface coal mine method for a new project

A new mine requires a method of mining coal given the estimated production rate. There are three major mechanized surface coal mining methods: Area strip mining, open pit mining, and contour strip mining (Curley 2011). These methods apply to coal formations (seams) that are close to the surface, making it profitable to extract them. All three methods involve the removal of the overburden over the coal seam, followed by extraction of the coal from the seam. The choice of which method to use in mining the coal depends on the geology, including the thickness and depth of the overburden, topography, and economics, such as capital costs, to name a few.

Area strip mining applies where the ground is relatively horizontal, shallow, or with enough area to create strips. The mine starts by opening a long strip, called a *box cut*, which is created by removing waste to expose the coal seam for extraction. Once coal is mined in the first strip, the next strip is created and the mined-out overburden is dumped into the first mine cut. The process continues with the removal of overburden ahead of coal extraction. The major equipment used is a dragline, which removes the waste. The truck and shovel are used for extracting the coal. Examples of this type of mine include Arthur Taylor Colliery Opencast Mine (ATCOM) in South Africa. To illustrate this mining method, Figure 2.1 shows the dragline excavating waste while the shovel loads coal into the truck.

Figure 2.1: Illustration of area strip coal mine (Wikipedia, the free encyclopedia 2016).

The area strip mining method requires large capital for the purchase of the equipment, which includes a dragline costing more than US$100 million. It has a low operating unit cost as compared to the other methods because it allows for many tonnes of production. The limitation of area strip mining is the lack of flexibility as the operations increase in depth (Scott et al. 2010; Westcott 2004; Baafi and Mirabediny 1998; Mitra and Saydam 2012).

Open pit mining involves extracting coal in a system of single or multiple steps known as benches. The benches are created in both overburden and coal seams. The typical appearance of open pit mining is indicated in Figure 2.2. In this method, stripped overburden is hauled to designated dumps and coal is transported to the plant to be either stockpiled or tipped directly into a crusher. Coal tonnage hauled from the pit to the plant is known as Run-of-Mine (ROM). The major pieces of equipment for this method are trucks and shovels.

Figure 2.2: Illustration of open pit coal mine (wikimapia 2016).

Open pit mining applies to complex coal formations with varying thicknesses and overburden. The capital cost is smaller than that of the area strip mining method. An example of an open cut mine is Mangoola coal mine, which is located in the Wybong area in Australia (Glencore 2016). It has an annual production of 9.2 Mtpa of thermal coal suitable for both export and domestic sales.

Contour strip mining is commonly used for extracting coal seams formed in mountainous areas. The overburden covering the coal that outcrops (appears) on the side of the hill is removed and then coal is extracted. The excavated overburden is dumped in the mined out bench, then mining proceeds inside the hill and stops when the overburden becomes so large that it is unprofitable to extract the coal. An illustration of this method is shown in Figure 2.3

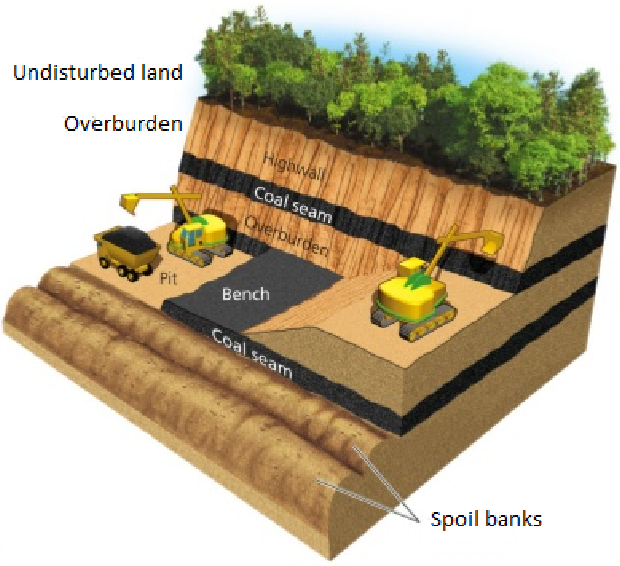

Figure 2.3: Illustration of contour mining (Myanmar 2016).

2.2 Estimating performance for a new mine project

The performance of a new mine involves estimating the production rate and costs based on the available information, such as the size of the deposit. The production rate is the first parameter to be estimated. It is the basis for the equipment selection and estimation of capital and operating costs for the mine.

2.2.1 Estimation of production rate for a new mine

There are different approaches for estimating the production rate of a new mine. Taylor in 1977 proposed a method for estimating this rate based on the expected tonnes to be mined. This is known as *Taylor's Rule* and is given in Equation (2.1), which also can be written in a general form as Equation (2.2) (Long 2009).

$$\text{Tonnes per day} = 0.014 \times (\text{expected tonnes})^{0.75} \tag{2.1}$$

$$C = bT^{a} \tag{2.2}$$

where C is the capacity in metric tons per day, also referred to as the production rate, a and b are coefficients to be estimated, and T is the reserve tonnage in metric tons.

Long (2009) reviewed *Taylor's Rule* and established that different model coefficients could be obtained. For example, a is less than the 0.75 initially suggested by Taylor. Two criticisms were raised during model evaluation: a) It is not a homogeneous model because it was developed from mines that were producing different commodities and they were not deposit-specific; b) the assumption of constant elasticity in the model is not valid because the relationships between the reserve and capacity are inelastic in nature. Technological changes at different times were not considered because the deposit might need a certain kind of technology for it to be extracted, and this could affect estimates of the production rate.

Investigation is thus needed into generating a production rate estimation model that could be used specifically for coal mines, taking into consideration the unique characteristics of the particular coal deposit.

A second method for determining the production rate is multiple economic analysis. This involves a series of production rates. The cost is computed for each production rate, and the production rate that gives the maximum Net Present Value (NPV) is chosen for application to the mine project under study. This method requires the use of software to test multiple scenarios of production rates and to determine the one that generates the maximum NPV (Leinart and Schumacher 2010). The production rate established using this approach does not guarantee that the mine will be efficient in practice relative to other producers or make the mine efficient and cost effective for competitiveness.

2.2.2 Estimation of costs for a new mine

In mining projects, cost estimates are done once the production rate has been estimated. Estimating costs for a specific coal deposit should reflect its unique characteristics and project location factors. For example, the geology and other challenges such as infrastructure can differ from one place to another. These challenges affect the final cost estimates. For example, a coal seam with a large overburden requires higher capital for stripping as compared to a coal seam covered by smaller overburden. The variation of the size of overburden for the two deposits will result in different cost estimates.

Cost estimates for mine projects is done at different stages of a new mine project before commencing production. The stages involve a sequence of studies named feasibility studies. In this context, the discussion is limited to estimation of cost in the conceptual study, pre-feasibility and feasibility study of new mine project. The detail of each stage of cost estimate requires engineering analysis to be done based on the available information at each stage of the project.

The engineering analysis required for cost estimation of new mine project among others includes: selection of mine method, mine design and production schedules, selection and compiling the list of mining equipment, generating engineering drawing showing the mine layouts, estimate time for the implementation of the project, test work to develop the processing flow sheet(mineralogical test, metallurgical test), flow sheet design, planning of infrastructure such as access and service roads to name a few (Runge Pincock Minarco 2015; Darling 2011; ALLEN 2012).

The level of the engineering analysis for cost estimation in conceptual stage of the project is very low. Most of the parameters about the project are assumed, for example the resource is not well defined and production rate is assumed. The results are used for the decision of advancing or rejecting the project to proceed it to the pre-feasibility study. The pre-feasibility study has more engineering analysis including estimation of production rate, listing of mining equipment and preliminary selection of mine method. The cost estimate at this stage has more engineering inputs compared to the conceptual stage of the project. Feasibility study cost estimates is more detailed than pre-feasibility study estimates. It has more detailed engineering analysis such as the optimal mine design and plans are generated, reserve has been defined clearly, metallurgical testing and design of flow sheets have been performed among others. This intends for approval or reject of the project for investment (Runge Pincock Minarco 2015).

There are methods and tools that are frequently applied for cost estimation of new mine projects at different stages, they can be grouped into statistical methods, on-line tools, comparative approaches and itemized methods (Budeba et al. 2015). The first three approaches are mostly applied in the conceptual stage. The cost estimation in the pre-feasibility and feasibility studies in most cases uses itemized method that requires engineering analysis of the project which forms the bases of cost estimation. Frequently, quotes from tenders, equipment suppliers and other unit costs are used in estimation of the mine costs. The details of each method are discussed in the following subsections.

2.2.2.1 Statistical method

The methods for estimating capital and operating costs in this category include the O'Hara model, multiple regression based on principle component analysis, econometric models, and the use of single-variable regression models compiled in handbooks. Examples of handbooks that are used in estimating mining costs include Capital Cost (CAPCOSTS) for mining and mineral processing equipment costs and capital expenditures, Canada Center for Mineral and Energy Technology (CANMET) for estimation of pre-production and operating costs of small underground deposits, and the cost estimation handbook for the Australian mining industry

(Sayadi et al. 2012b).

The O'Hara method involves sets of graphs that plot the cost of production against the production rate for both mining and milling operations. The models were prepared in the 1980s, and they need to be updated to accommodate the escalation of costs in the estimates. This can be done using cost indices published by different sources, such as the Marshall Swift indices. The indices are used to adjust the costs, taking into account the effect of inflation and the adjustments needed to cover costs such as mining and milling, labour, machinery, and heavy equipment (Shafiee and Topal 2012). The general formula for updating the costs is represented by Equation (2.3) (Wellmer et al. 2007).

$$\text{Costs today} = \text{Cost in year } x \times \frac{\text{index today}}{\text{index in year } x} \tag{2.3}$$

The equations representing all costs using the O'Hara method are presented by Shafiee and Topal (2012). This can also be written into its components, as shown by Equations (2.4)–(2.10). Consider the following definitions:

T = tons mined and milled daily

T_w = million tons of overburden rock

T_s = million tons of overburden soil

T_d= tons of deposit and waste mined daily

By using the definitions above, the cost for each component of the mine project is as follows:

$$\text{Capital cost (US\$M)} = \$400000T^{0.6} \tag{2.4}$$

$$\text{Stripping cost soils (US\$M)} = \$800T_s^{0.5} \tag{2.5}$$

$$\text{Stripping cost rocks (US\$M)} = \$8500T_w^{0.5} \tag{2.6}$$

$$\text{Equipment cost (US\$M)} = \$6000T_d^{0.7} + \$5000T_d^{0.5} \tag{2.7}$$

$$\text{Maintenance cost (US\$M)} = \$150000T_d^{0.3} \tag{2.8}$$

$$\text{Labour cost (US\$)} = \$58.563T_d^{-0.5} + \$3.591T_d^{-0.3} \tag{2.9}$$

$$\text{Supplies cost (US\$)} = \$13.40T_d^{-0.5} + \$1.24T_d^{-0.3} + \$0.9T_d^{-0.2} \tag{2.10}$$

An illustration of the application of the O'Hara model is given in example 2.2.1.

Example 2.2.1 : Consider a new open pit coal mine, together with a washing plant, located in South Africa, which is planned to produce steam coal at 1.5 Mt/yr. If the mine would start in January 2016, what would be the estimated capital cost assuming that the mine operates 340 days/year? To compute the estimates of capital cost, Equation 2.4 is used and it gives the following answer:

$$\text{Capital cost}(US\$M) = 400000 \times (1500000/340)^{0.6} = \text{US\$61.5 million.}$$

A capital cost of US$61.5 million obtained from the calculation above is the cost for the project using the models developed in the 1980s. In order to update the cost to January 2016, the cost index for capital in mining, including coal, is applied. Given a cost index of 73.1 in 1984 and a cost index of 169.1 in January 2016 (US. Bureau of Labor Statistics 2016), the mine capital cost would be US$142.3 million using Equation (2.3), as shown in the following calculation.

$$\text{Costs for January 2016} = 61.5 \times \frac{169.1}{73.1} = \text{US\$142.3 million.}$$

CAPCOSTS, developed by Mullar and Poullin in 1998, the Cost Estimation System (CES), prepared by the US Bureau of Mines, and the cost estimation handbook for the Australian mining industry are among the available handbooks. They are based on single-variable regression for the estimation of costs. The models were developed using geometric regression and are presented in a general form as Equation (2.11).

$$\mathrm{Y} = A(X)^B \tag{2.11}$$

where A and B are coefficients, X is an independent variable that can refer to capacity or size, and Y is the cost.

In some projects, the models have been used for pre-feasibility types of economic evaluation to estimate the costs. Sayadi et al. (2012b) argue that most of these cost estimation models use single-variable regression to estimate mineral industry costs. The other significant independent variables affecting cost are simply ignored. For example, when these models are used for coal mine cost estimation, the depth of waste removal over the coal, the location of the mine from the market and services, and geologic variations such as thickness are excluded. However, these variables influence the capital costs of mining (Hartman 1992). Excluding them will change the final results of the cost estimates. Moreover, Sayadi et al. (2012b) claim that the models are obsolete by showing examples of those developed in the 1980s and 1990s, so updating them can cause errors in the cost estimates.

Long (2011), on the other hand, found that in estimating the cost of porphyry deposits, the only variables that could be used to estimate the capital cost before mining construction began were the mineral processing rate, the stripping ratio, and the distance from the nearest railway. In the case of the operating cost, the stripping ratio was the only variable that affected the cost.

In addition, Long (2011) claims that no model attempted to date has found any significant explanatory variable other than the stripping ratio, which at best explains only 40% of the variation in the costs. The variables tested included grade, mineral processing rate, ore hardness, and fuel and power costs for porphyry copper mines. In the case of coal mines, Shafiee et al. (2009) estimated the operating cost in a surface coal mine using capital, thickness of the deposit, production rate, and stripping ratio.

Multiple linear regression based on analysis of the principal component has been used in estimating the capital and operating costs for individual equipment, such as back-hoe loaders and shovels (Sayadi et al. 2012a; Oraee et al. 2011). In this method, costs are estimated from bucket size, digging depth, dump height, weight, and horsepower. Another form of multiple regression is the econometric model, proposed by Shafiee and Topal (Shafiee and Topal 2012; Shafiee et al. 2009). This model involves estimating the operating cost using capital cost, production rate, stripping ratio, and coal thickness as independent variables. This model is represented by Equation (2.12) as adapted from the Shafiee and Topal model. Example 2.2.2 shows the application of the model using one of the datasets adapted from the table generated by Shafiee and Topal (2012); Shafiee et al. (2009).

$$E = 8.744955 + 0.041556T + 1.658269S - 0.000459C - 0.041408P \tag{2.12}$$

where E is the estimated operating cost (cost per tonne), T is the average thickness of the deposit in metres, S is the stripping ratio, C is the capital cost (million dollars in 2008), and P is the daily production rate (1000t). The study indicated that the model can estimate costs to within ±20%.

Example 2.2.2 *:* Suppose an open pit coal mine is expected to produce 1.8 Mt/yr and has a coal seam with an average thickness of 12.3 m. The mine's Stripping Ratio (SR) is 10.2. What is the operating cost for the mine, given the capital cost of A$60.2 million to be spent for the project?

Substituting the values of the given information into Equation (2.12) gives the following result:

$$\mathrm{E} = 8.744955+0.041556\times 12.3+1.658269\times 10.2-0.000459\times 60.88-0.041408\times 1.8 = A\$25.05$$

An operating cost of A$25.05 is the result calculated using Equation (2.12), which was developed from data collected in 2009. This cost has to be adjusted using cost indices for the year in which the cost estimates are required.

Other cost guides for coal mines include the Australian Coal Cost Guide, which is internally generated by Costmine in Australia, and Coalval, a tool that was developed by the United States Geological Survey (USGS) for the valuation of coal properties. This does not consider the characteristics of the coal deposit, such as the geology of the seam, which is one of the variables that affects mining costs (Chan 2008).

2.2.2.2 Itemised method

The major concept of the itemised method has been discussed by Darling (2011). It involves three major steps. First, a conceptual mine plan is developed using the available information, such as pit outlines, routes, depth of the waste dump, and location of the processing plant. Second, parameters that are associated with costs are estimated, and third, the known unit costs of labour, equipment operation, and other facilities are applied to finalize the cost estimates. The itemised method is a detailed approach that depends on the production rate and conceptual mine plan. A change in the production rate and/or mine plan can affect the final cost estimates.

To illustrate an example of the itemised method, Table 2.1 represents the equipment capital cost estimate as adapted from Leinart and Schumacher (2010). The last column of the table is the capital cost (US$), which is the product of the units and the cost/unit columns. The total equipment cost is the sum of all capital costs.

Table 2.1: Equipment capital cost 2009 (Leinart and Schumacher 2010)

Equipment	**Specification**	**Units**	**Cost/unit**	**Capital cost (US$)**
Wheel Loader	800hp,11.5m3 bucket	1	1807100	1 807 100
Front Shovel	1550hp, 17.0m3	1	4396000	4 396 000
Haul trucks	938hp,90tonne,rig frame	15	1128000	1 692 000
Rotary blast hole drill	15.2cm hole,52.7m hole depth approximately 27000kg pull down	3	670000	2 010 000
Road maintainer	4.3 blade,270hp	1	445000	445 000
Dozers	4.5blade,270hp	4	731000	2 924 000
Bulk explosive trucks	459kg/min	1	74800	74 800
Lighting plant	16kw	4	21900	87 600
Fuel/Lube truck		2	55400	110 800
Mechanic's truck		2	67000	134 000
Tire service truck		1	158000	158 000
Centrifuge pump	65hp,3028lmp, 45m head	1	23670	23 670
Water truck	53000 litre	1	744000	744 000
Pick up trucks	3/ton, four wheel driver	13	23600	306 800
Total Equipment cost				**30 141 770**

Other components of the total costs such as operating costs are estimated using the same approach as indicated in Table 2.1. Summing all itemised costs for all components of mining gives the total mining costs. All costs from different tables recorded using the itemised method are combined to create a cost model of the mine project.

2.2.2.3 Comparative approach

The comparative approach involves the use of an existing mine with characteristics similar to those of the mine project under study. Most of the time, mine evaluators use the average costs of similar mining projects and operations, then adjust by a factor to account for specific site conditions, mining methods, commodity prices, and milling processes (Shafiee and Topal 2012). However, there are no clear guidelines on the cost adjustment needed to reflect the conditions of the mine under evaluation. Hustrulid and Kuchta (2006) refer to this method as being an analogous method of cost estimation. They suggest that comparing similar operations should be undertaken with care because accounting practices vary as well.

2.2.2.4 On-line tools

One on-line tool is the Mine and Mill cost calculator and mine costs. This is a database of mining and milling equipment. It can be used to calculate the cost of equipment to be used in the operation. Mine Cost is a second on-line tool, which consists of spreadsheets and curves that show capital and operating costs. On-line tools such as Mine Cost estimate the total mining cost of a specific mine. It is not clear how the costs are obtained; the tool simply generates the final total cost estimates (Shafiee and Topal 2012).

2.3 New mine production planning

Once the production rate and costs have been estimated, the optimisation process is conducted to determine the optimal shape of the cut for mining the coal. The major input variables used in the optimisation process include the estimated capital costs, operating costs, price, coal quality such as calorific value Calorific Value (CV), and the expected mining production rate, to name a few. The optimal cut consist of the tonnage and quality of coal that can be extracted to generate the maximum expected NPV. Illustration for optimization in coal project using Lerchs–Grossmann (LG) algorithm has been discussed in the study by Prentice (2005) and Stojanovic et al. (2014).

On the other hand, a commonly traditional approach to determine the pit limit that can be mined at profit is the use of a Break Even Stripping Ratio (BESR). The BESR refers to ratio at which the unit cost of producing a tonne of coal is equal to the unit of price of a tonne of coal. To illustrate the concept, consider I to be the revenue per tonne of ore, C_t to be the production cost per tonne of ore (including all costs to the point of sale, excluding stripping) and C_s to be the stripping cost per tonne of waste. Then, BESR can be represented by Equation (2.13) (Oraee et al. 2008).

$$BESR = \frac{I - C_t}{C_s} \tag{2.13}$$

After optimisation process either by using LG or the use BESR, design of the geometry of the pit, such as the height of the benches for the open pit method is done. Next is scheduling, which shows how the coal seam will be mined in the given optimal cut limit over a specific period. Following design and scheduling is the extraction of the coal. Some of the activities continuously repeat over the entire life of the mine, with updates at each stage if new information becomes available or when opening a new area that requires mining. Figure 2.4 shows the major repeating activities to ensure continuous delivery of coal for sales.

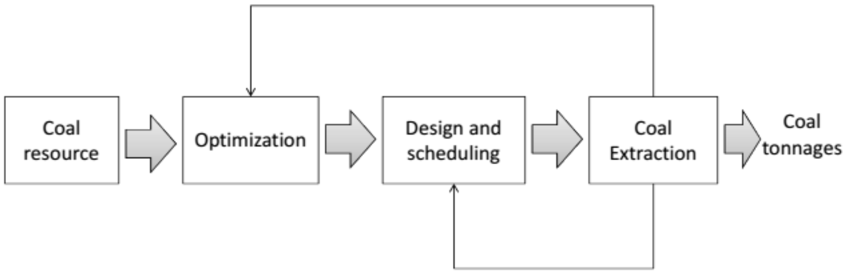

Figure 2.4: Activities for continuous production cycle.

Despite the effort that goes into the activities shown in Figure 2.4, problems have been identified for mines failing to achieving their objectives, such as tonnage required, once they are extracting coal. For example, a survey by Ernst Young (2014a) indicated that productivity is the first of the ten ranked risks in the mining industry. The traditional approach to productivity improvement, such as cost cutting, has been challenged. Instead, the report suggests that changes in mine plans are necessary, and it recommends that productivity should come from a whole-of-business, end-to-end transformation. For example, the survey showed that labour productivity dropped by 50% in Australia and 30% in the USA in the period between 2009 and 2012. One of the reasons established was cost cutting applied as a single-point solution, which consequently compromised the whole optimisation plan of the supply chain, hence affecting the productivity.

2.4 Technical efficiency and application in mines

Before discussing technical efficiency for the evaluation of mines, it is important to review the concept of productivity for entities such as mining operations. Productivity refers to the quantity of output per unit quantity of inputs used in producing the output. For example, labour productivity in surface coal mining can refer to the ratio of the total tonnes of coal produced to the number of the employees used. If the surface coal mine produces 2.7 Mt/annum of coal using 300 employees working in the mine, then the labour productivity would be 9000 tonnes/annum/employee.

Productivity can be used by a producer to compare its current use of inputs against the minimum quantity of inputs it could use to generate the present output, or it can be used to compare itself with other producers of similar commodities. When a producer is using multiple inputs and outputs, productivity can be calculated using aggregated outputs and inputs (Fried et al. 2008). The comparison of the current use of inputs relative to the minimum possible use of inputs to produce the present outputs reflects the efficiency of such a producer.

2.4.1 Concept of technical efficiency

Efficiency has been discussed by several authors. According to Farrell (1957), 'when one talks about the efficiency of a firm, one usually means its success in producing as large as possible an output from a given set of inputs. Provided all inputs and outputs were correctly measured, this usage would probably be accepted.' Efficiency has also been described by other scholars such as Markovits-Somogyi (2012, p.12), who defines it as the capacity of a company to realize its stated objectives and to use its available resources cost-effectively.

In addition, Joubert (2010) articulates that analysis of efficiency offers guidelines and benchmarks for both public and private enterprises to achieve maximum outputs with minimum inputs. Determining efficiency helps the enterprise to evaluate its ability to compete in a group of enterprises doing similar business.

The concept of efficiency can also be illustrated using Figure 2.5. This refers to a firm that uses two input resources and transforms them into two outputs. Such a firm is referred to as a Decision Making Unit (DMU). Therefore, the relationship between the outputs of the DMU to its inputs is used in determining the efficiency of the DMU.

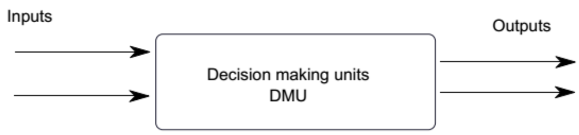

Figure 2.5: Input–output relationships for a DMU (Fried et al. 2008).

It is important for the management of both non-profitable and profitable enterprises to be able to assess the efficiency of their operations in such a way that they can identify areas needing improvement if the operations appear inefficient. They can also determine the size of the operation that will improve efficiency and make it competitive. For example, when a company operates in a group of firms consuming similar input resources and generating similar outputs, it would be interesting to know the best practices among all the firms and to measure their level of competitiveness. This would require determination of the relative efficiency of each firm involved.

It could be generalized that an operation that allocates and utilises minimum input resources to produce target outputs or maximizes the output using the same quantity of inputs is efficient. A surface coal mine is no exception; it needs to be efficient when using inputs such as capital and labour within the given geological, geographical, environmental, and legislative constraints to achieve its production outputs. An operating mine has to be efficient to deliver its targeted outputs in a competitive business environment and to generate a return on investment to the shareholders.

2.4.2 Methods of measuring technical efficiency

Methods of measuring efficiency can be classified into two main categories: parametric and non-parametric (Porcelli 2009; Budeba et al. 2014).

The parametric methods includes deterministic and stochastic. Deterministic method includes: Ordinary least square regression, Corrected Ordinary least Squares (COLS) and Modified Ordinary least Squares (MOLS). These methods differ on how the inefficiency is considered, deterministic methods considers the sources of inefficiency is the deviation of the observed production from the maximum feasible outputs while stochastic method considers the that the inefficiency is the result of two components that includes deviation of the observed outputs from the maximum feasible output and the random effects such as weather which can not be measures. The comprehensive discussion about these methods is found in the work by Kumbhakar and Lovel (2003). The advantage of the method is the ability to conduct statistical test for the parameters of the model. The disadvantage of these methods is mis-specification of the function and problems of estimating the parameters (Murillo-Zamorano and Vega-Cervera 2001).

The non-parametric methods, on the other hand, do not require the function to be predefined. Instead, the Linear Program (LP) technique is used to construct a piecewise envelope that represents the efficient DMUs. These methods are used to measure the efficiency of a firm using data. They can thus be regarded as data-driven. Under these methods, no function form is assumed; instead, an LP is used to determine the envelope of best practices. The leading non-parametric method for measuring the efficiency of DMUs with multiple inputs and outputs is Data Envelopment Analysis (DEA) Despotis et al. (2010); Hollingsworth (2003); this is deterministic, with the assumption that all inputs or outputs are discretionary. It was first introduced by Charnes and Cooper in 1978. DEA can be input-oriented or output-oriented. The input-oriented method aims at minimizing the inputs while satisfying at least the given output, whereas the output-oriented method aims at maximizing the output without using more of the inputs (Cooper et al. 2007, p.41).

The basic DEA models are those of Charnes–Cooper–Rhodes (CCR) and Banker–Cooper–Charnes (BCC). The CCR model assumes that an increase in input results in a proportional increase in output. Thus, for example, an increase in input by $\alpha\%$ will increase the output by

$\alpha\%$ as well. This refers to Constant Return to Scale (CRS), whereas the BCC model assumes that an increase in input will result in a greater or lesser proportionate increase in output; it is considered to be variable return to scale, VRS.

The relationship between the CCR and BCC models is that the CCR model can be decomposed into pure technical efficiency, denoted by BCC, and scale efficiency, denoted by Scale Efficiency (SE). The pure technical efficiency is referred to simply as technical efficiency, which measures the ability of management to utilize the input resources of the DMU in producing the outputs. The scale efficiency measures the ability of management to choose the optimal size of the DMU. The product of BCC and SE gives CCR (Kumar and Gulati 2008).

DEA method has the following advantages: it is used to measure efficiency considering multiple inputs and outputs of a unit under study, it does not require assuming the function form to use for estimating relative efficiency of each unit, it use inputs and output even if there differences of unit of measurements Charnes (1994). The main disadvantage of the DEA methods does not separate the difference between technical inefficiency and statistical noise effects (Murillo-Zamorano and Vega-Cervera 2001). The method depend on the accurate measurement of data with no errors. Incorrect conclusion can be reached about the technical efficiency of a DMU for the observation of the inputs or outputs consisting of measurement errors.

The parametric and non-parametric methods of efficiency measurements for mine projects are illustrated using example 2.4.1.

Example 2.4.1 *:* Suppose a group of surface coal mines named A, B, C, D, E, F, G, H, I, J, K, and L are producing coal using capital as a major input and generate tonnes of coal as an output. Each mine's capital used and coal tonnage produced are shown in Table 2.2. To illustrate and compare the use of parametric and non-parametric methods for measuring efficiency, the data provided in Table 2.2 are plotted in Figures 2.6–2.8. Figure 2.6 shows each mine and its capital and coal production. Figure 2.7 applies the parametric (regression) method, and Figure 2.8 applies the non-parametric DEA method.

The parametric method requires a regression equation to be specified using data. The equation is assumed to be an optimal plane that applies to each data, but in reality the equation represents the average plane Charnes (1994).

Table 2.2: Input and output of each mine

Mine	Capital (US$M)	Coal (Mt/yr)
A	100	4.50
B	190	6.00
C	55	2.00
D	80	3.00
E	90	4.80
F	85	2.80
G	95	4.00
H	35	1.00
I	45	2.40
J	115	5.00
K	50	2.20
L	230	6.00

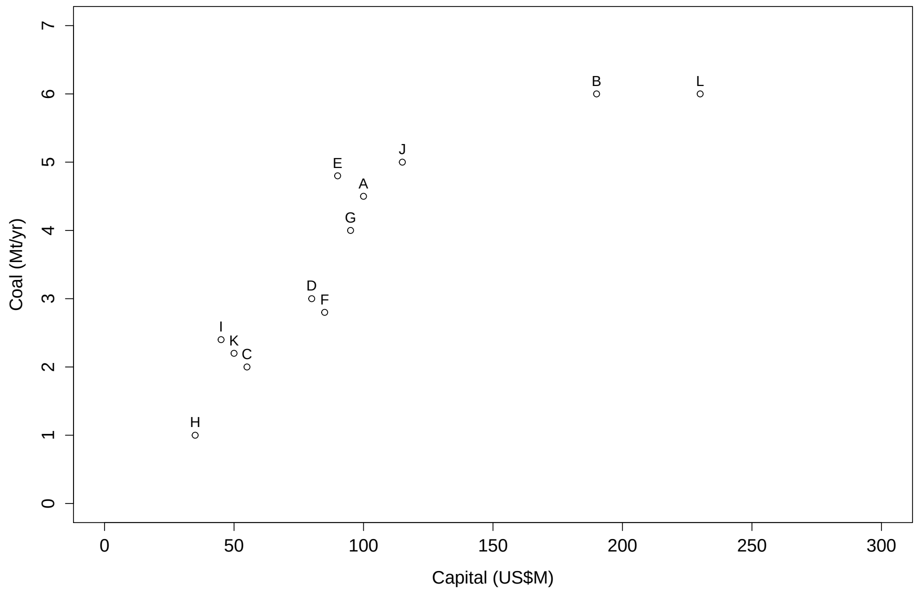

Figure 2.6: Single input–output for mines.

When the parametric method is applied, a plane is fitted on the points, as shown in Figure 2.6, relating each mine's coal tonnes as a dependent variable and capital as an independent variable. The resulting plot is indicated in Figure 2.7. In fitting the plane, it is assumed that the errors are identically and independently normally distributed. The mines on the plane represent those operating at an average efficiency, those above the plane represent mines operating at an efficiency above the average, and those below the plane represent mines operating at an efficiency below the average.

Therefore, this method does not indicate how relative the mine is efficient because the residuals are used to compare how inefficient a mine is from the average efficiently operating mine. The efficiency using this approach does not reflect the correct measure because the residuals may consist of random noise outside the control of the mine.

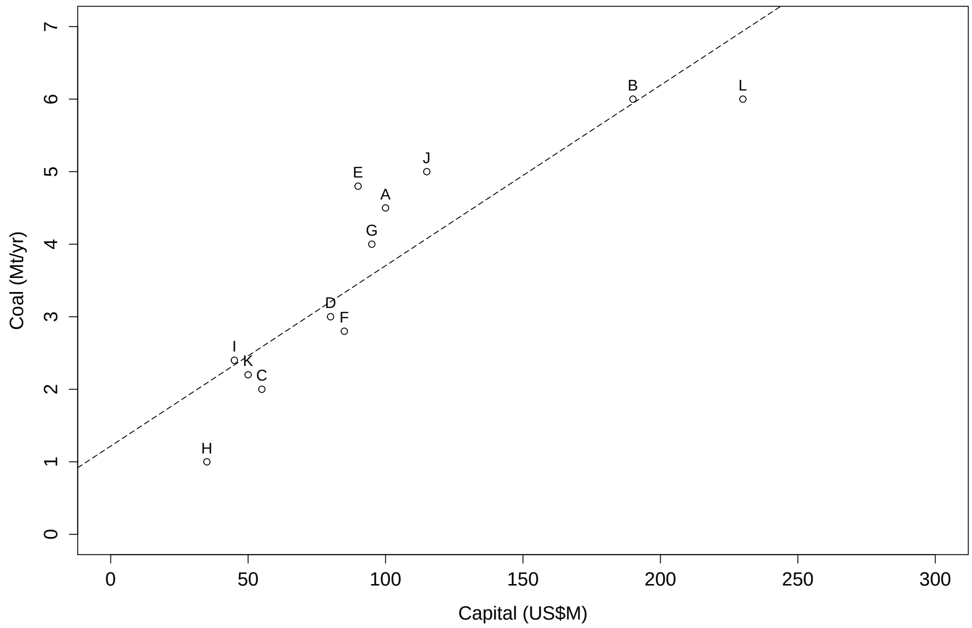

Figure 2.7: Parametric method.

The non-parametric DEA method was also applied to the data given in Table 2.2. The values of CRS and Variable Return to Scale (VRS) technical efficiency were calculated using R Core Team (2015), and the results are shown in Table 2.3. The efficiency scores under the CRS assumption were calculated in the software using the following Equation 2.14:

$$\text{Efficiency score} = \frac{\text{Productivity of a DMU}}{\text{Maximum productivity}} \tag{2.14}$$

For example, the productivity of mine A is 0.0450 tonnes/yr/U$M and the maximum productivity of all mines is 0.05333. Therefore, the efficiency score of mine A under CRS is 0.84375. The values under VRS were calculated in the software using the linear programming formulation of the problem as implemented in (Bogetoft and Otto 2015).

Table 2.3: Efficiency scores for each mine

Mine	Capital (US$M)	Coal (Mt/yr)	CRS	VRS
A	100	4.50	0.8437	0.8438
B	190	6.00	0.5921	1.0000
C	55	2.00	0.6818	0.7662
D	80	3.00	0.7031	0.7031
E	90	4.80	1.0000	1.0000
F	85	2.80	0.6176	0.6176
G	95	4.00	0.7895	0.7895
H	35	1.00	0.5357	1.0000
I	45	2.40	1.0000	1.0000
J	115	5.00	0.8152	0.9275
K	50	2.20	0.8250	0.8714
L	230	6.00	0.4891	0.8261

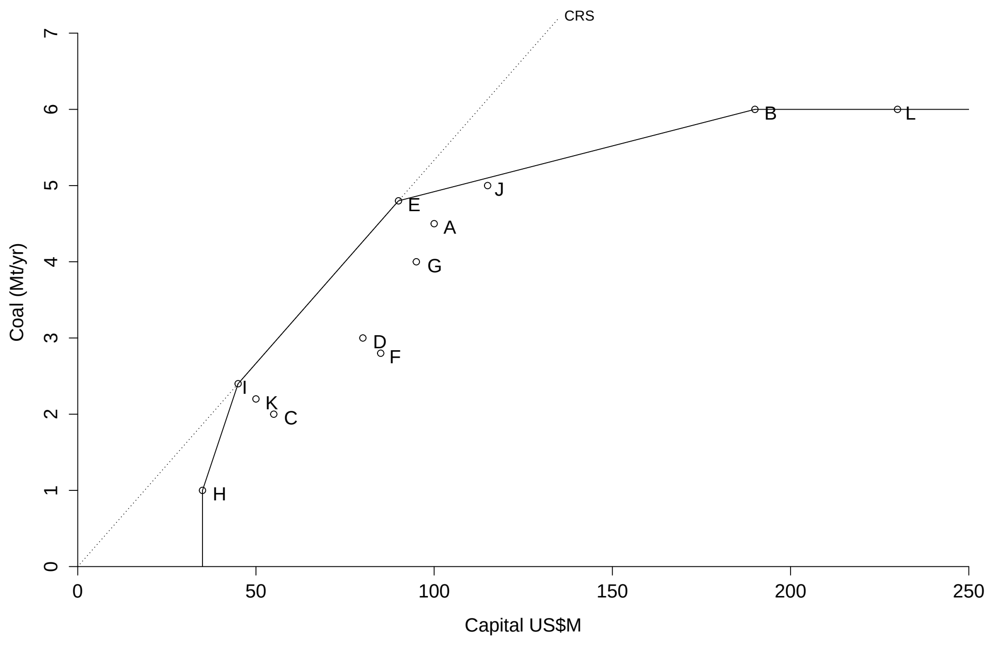

Figure 2.8: Non-parametric method.

From Figure 2.8, mines L, B, E, H, and I under the VRS assumption form the envelope of best practices. In the case of CRS, the efficient mines are I and E, with efficiency scores of 1. The inefficient mines have efficiency scores less than 1. They are A, C, D, F, G, K and J. These mines need to improve their efficiency by reducing their inputs to reach the envelope of best practices while producing their present coal tonnage.

A comparison of the two methods between Figure 2.7 and Figure 2.8 indicates that the parametric method represents the average plane, which is assumed to apply to all mines, whereas the non-parametric DEA shows the envelope representing the efficient mines.

An illustration of using more than one input in mining operations is presented in example 2.4.2.

Example 2.4.2 *:* To illustrate the use of more than one input to generate one output, consider 12 operating open pit mines, each of which produces 1.5 Mt/yr of coal, using capital and labour as inputs. The datasets for the inputs and output for each mine are given in Table 2.4. We can apply DEA to identify the mines that are efficient and those that are not efficient.

Table 2.4: Mine design scenarios

Mine	Capital (US$M)	Employees	Coal (Mt/yr)	VRS
D1	58	300	1.5	1.0000
D2	124	107	1.5	0.7628
D3	130	75	1.5	1.0000
D4	98	100	1.5	0.9342
D5	144	120	1.5	0.6657
D6	112	98	1.5	0.8424
D7	77	150	1.5	1.0000
D8	106	95	1.5	0.8861
D9	95	80	1.5	1.0000
D10	115	122	1.5	0.7896
D11	200	75	1.5	1.0000
D12	145	180	1.5	0.6042

The data provided in Table 2.4 were used to calculate the efficiency scores of each mine under VRS using the benchmarking package Bogetoft and Otto (2015) in R software and also illustration of dea for isoquant (units having same quantity of outputs) by Behr (2016) and included in Table 2.4. A graphical representation of each mine is shown Figure 2.9.

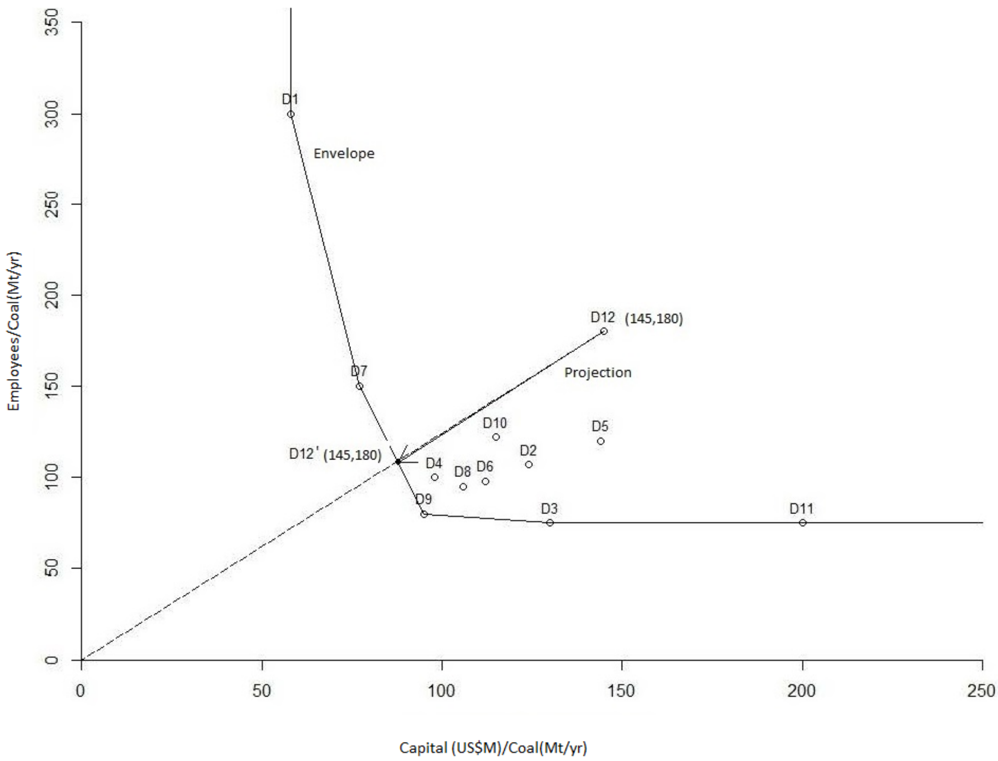

Figure 2.9: Two inputs–one output using the DEA method.

From Figure 2.9, the mines that form the envelope of best practices are $D1$, $D7$, $D9$, $D3$, and $D11$. These have efficiency scores of 1. They are using the minimum amount of inputs to produce the present output of 1.5 Mt/year. It can be noted that $D11$ is efficient, but it is using excess capital as compared to $D3$, which is using less capital to attain the same production. Reducing its capital to that of mine $D3$ will still generate the production target. The other remaining mines are inefficient, i.e., they are using an excess of both capital and employees to produce the 1.5 Mt/yr of coal. They can only be efficient if they reduce their capital and labour proportionally to be efficient. For example, $D12$ will only be efficient if it reduces its capital and employees to that of $D12'$ on the envelope.

Representing more than two inputs and outputs on a single graph and showing the efficiency scores of each DMU for the DEA study is a challenge. The only frequent approach used is to compute the efficiency scores using mathematical models developed for the DEA method. These models are basic and their results are always recorded in tabular form and then interpreted. The models are implemented in some software such as R Core Team (2015), as used for the illustrations in examples 2.4.1 and 2.4.2.

2.4.3 Mathematical representation of the basic DEA models

To represent the mathematical models for the DEA method, we assume a set of $\mathbf{J}$=$\{1,...,n\}$ production or service-providing units and each of them is regarded to be a DMU. Consider that each $j \in \boldsymbol{J}$ uses m different inputs to generate s different outputs. Let us assume the set of inputs is $\mathbf{I}$=$\{1,...,m\}$ and the set of outputs is $\mathbf{R}$=$\{1,...,s\}$ and use the following definitions:

$x_{ij} \triangleq$ the given usage of input $i \in \boldsymbol{I}$ by DMU $j \in \boldsymbol{J}$.

$v_i \triangleq$ the weight of input $i \in \boldsymbol{I}$.

$y_{rj} \triangleq$ the given output $r \in \boldsymbol{R}$ generated by DMU $j \in \boldsymbol{J}$.

$u_r \triangleq$ the weight of output $r \in \boldsymbol{R}$.

Then, from the definition of efficiency given by Talluri (2000) in Equation (2.15), the relative technical efficiency of each DMU can be written in fractional form shown by Equations (2.16)–(2.18). The DMU under evaluation is referred to as DMU$_o$.

$$\text{Efficiency} = \frac{\text{weighted sum of outputs}}{\text{weighted sum of inputs}} \tag{2.15}$$

$$\max h_o = \frac{\sum_{r\in\boldsymbol{R}} u_r y_{ro}}{\sum_{i\in\boldsymbol{I}} v_i x_{io}} \tag{2.16}$$

subject to

$$\frac{\sum_{r\in\boldsymbol{R}} u_r y_{rj}}{\sum_{i\in\boldsymbol{I}} v_i x_{ij}} \leq 1 \qquad \forall j \in \boldsymbol{J} \tag{2.17}$$

$$v_i, u_r \geq 0 \qquad \forall i \in \boldsymbol{I}, r \in \boldsymbol{R} \tag{2.18}$$

The Fractional Program (FP) in Equations (2.16)–(2.18) is non-linear and can be transformed using the Charnes and Cooper transformation. LP is used in this method to determine the set of weights that maximise the efficiency of DMU$_o$. To carry out the transformation from Fractional Program (FP) to LP, let t be the transformation parameter given by Equation (2.19).

$$t = \frac{1}{\sum_{i\in\boldsymbol{I}} v_i x_{io}} \qquad \forall i \in \boldsymbol{I} \tag{2.19}$$

where $\sum_{i\in\boldsymbol{I}} v_i x_{io}$ is the sum of the weighted inputs of DMU$_o$. When we apply the transformation t to the FP, the new weight for the output becomes $\mu_r = tu$ and the weight for the inputs becomes $\nu_i = tv$. Therefore, the resulting LP is given by Equations (2.20)–(2.23).

$$\max g_o = \sum_{r\in \boldsymbol{R}} \mu_r y_{ro} \tag{2.20}$$

subject to

$$\sum_{r\in \boldsymbol{R}} \mu_r y_{rj} - \sum_{i\in \boldsymbol{I}} \nu_r x_{ij} \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{2.21}$$

$$\sum_{i\in \boldsymbol{I}} \nu_i x_{io} = 1 \tag{2.22}$$

$$\nu_i, \quad \mu_r \geq 0 \qquad \forall i \in \boldsymbol{I}, r \in \boldsymbol{R} \tag{2.23}$$

In using the equations above, the DMU is regarded as efficient if the efficiency score $g_o = 1$ and optimal values of the multipliers $(\nu^*, \mu^*) > 0$. Otherwise, if the efficiency score is $g_o < 1$, the DMU is considered inefficient. Charnes (1994) points out that if (u^*, v^*) is optimal for Equations 2.16–2.18, then for $t > 0$, the variables (tu^*, tv^*) are also optimal.

The above LP can be written in the form indicated by Equations (2.24)–(2.27). This is known as dual optimisation of the primal problem, shown in Equations (2.20)–(2.23). The dual optimisation problem involves fewer of variables to solve and still attains the same optimal solution as the primal problem.

$$\min \theta \tag{2.24}$$

subject to

$$\theta x_{io} - \sum_{j\in \boldsymbol{J}} \lambda x_{ij} \geq 0 \qquad \forall i \in \boldsymbol{I} \tag{2.25}$$

$$\sum_{j\in \boldsymbol{J}} \lambda_j y_{rj} \geq y_{ro} \qquad \forall r \in \boldsymbol{R} \tag{2.26}$$

$$\lambda_j \geq 0 \tag{2.27}$$

When considering the existence of excess input or output shortfalls (slack), the CCR model can be expressed using Equations (2.28)–(2.31), where the excess of inputs is s^- and the output shortfalls are s^+.

$$\min \left\{ \theta - \varepsilon \left(\sum_{i\in \boldsymbol{I}} s^- + \sum_{r\in \boldsymbol{R}} s^+ \right) \right\} \tag{2.28}$$

subject to

$$\theta x_{io} = \sum_{j\in \boldsymbol{J}} \lambda x_{ij} + s^- \qquad \forall j \in \boldsymbol{J} \tag{2.29}$$

$$y_{ro} = \sum \lambda y_{rj} - s^+ \qquad \forall r \in \boldsymbol{R} \tag{2.30}$$

$$s^{-}, s^{+}, \varepsilon, \lambda \geq 0 \tag{2.31}$$

The equation that involves slack can be solved in two stages by first solving θ without slack and then using the value of θ obtained in the first stage to solve the optimal values of the slack in the second stage.

In the BCC model, a condition, $\sum_{j \in \boldsymbol{J}} \lambda_j = 1$ is added to the constraints, in addition to the constraints used in the CCR model. These models are basic, and they can be used to develop more complex models for measuring the efficiency of a supply chain system.

Determining the envelope of best practice firms using the DEA method can be done as follows:

- Compute the efficiency score of the first DMU by determining the set of weights of both inputs and outputs that maximize the relative efficiency of such a DMU.
- Repeat the step above for each DMU.
- Select all DMUs with efficiency scores equal to 1. These form the envelope of best practices, whereas those with efficiency scores less than 1 form a set of inefficient DMUs.

2.4.4 Previous applications of DEA in coal mine projects

A survey of the application of DEA in the mining industry was done to identify how the DEA method has been used in coal mining operations. Three major areas of the mining business were identified from the previous work of other scholars. These include coal mine production, evaluating the safety aspects of coal mines, and measuring the performance of coal enterprises. In most of the studies reviewed, the DEA method was used as a black box because the inputs and outputs were generalized as a whole without detailing the components of the mine as a system. Most of the previous studies concentrated on one area of the mine independently.

However, Cook et al. (2010) proposed a generic multistage DEA for general consideration of intermediate measures for many industries. These authors gave an example of the use of a multistage DEA. ROM would be considered as an output from the first stage, where the efficiency of the coal mining process would be measured. The second stage would involve washing the coal, and then the washed coal would be delivered to the market. But these authors did not formulate models specific to the operation structures of coal mines, nor did they take into account the influence of non-discretionary variables that have an impact on operations. The formulation of a DEA model that would consider the market structures of coal and account for the non-discretionary variables could help a mine evaluate its competitiveness given the unique characteristics of the mine, such as the climate in which it operates.

Reddy et al. (2013) and Kulshreshtha and Parikh (2002) discussed the efficiency measurement in coal mine production. Kulshreshtha and Parikh (2002) described the efficiency of both open-cast and underground coal mining firms in India covering a period between 1985 and 1997. The inputs and outputs were as follows:

- The inputs for open-cast mining were mining machinery, cranes, and dumpers man-shifts and the output was the overburden removal.
- The inputs for underground mining included mining machinery and rope haulage man-shifts and the output was the tonnage of coal.

The authors found that between 1985 and 1997, open-cast mining did not show more productivity growth than the underground mining as previously believed. Also, the underground mining was shown to have more efficiency practices, whereas open-cast mining showed more technical changes over the study period. The authors applied a Malmquist index by decomposing productivity changes into efficiency and technical changes. Reddy et al. (2013), on the other hand, discussed benchmarking of open-cast coal mines in India. The aim was to rank the mines according to their efficiency and determine the improvements required for the inefficient mines. The inputs for the study were wage costs, store costs, overburden removal costs, and other costs, whereas the output was saleable coal. In addition, the two studies applied the basic DEA model, but they had different inputs.

Another area of study has been safety evaluation in coal mines. The application of DEA for safety is discussed by Lei and Ding (2008), who assessed the safety inputs for a coal mine in China covering a period between 2001 and 2005. The intention was to assess how safety inputs are used and to optimize the resource allocation. The assessment was done using a CCR output-oriented model of the operating mines, considering the inputs as outlays for safety: funds, staff, and time. These inputs included personal protective equipment, training, and management costs. The authors developed seven indices that were related per man in the mine, which were used as inputs. The output index used for illustration was the ratio of reducing accidents in a given year as compared to the previous year. On the other hand, Shu-Ming (2011) used different inputs and output, but all of them were generated from the costs allocated for safety at the mine. This study covered the period between 1996 and 2005.

Fang et al. (2009) described the application of DEA for measuring the performance of coal enterprises in the energy sector. The study compared 8 listed coal mining companies in the US and 17 listed Chinese coal mining companies using CCR and BCC. The authors found that despite China being a large producer of coal, the Chinese coal companies had much lower technical efficiency as compared to their American counterparts. To point out one cause of inefficiency among others, as discussed by Fang et al. (2009), most of the Chinese companies were state-owned coal mining firms, with a large number of redundant employees. The author used the following inputs for comparison: operating costs, total assets, and numbers of employees; the output variables consisted of earnings per share, operating revenue, and net profit before taxes.

2.5 Conclusion

This chapter reviewed methods of estimating the production rate and costs of a new mine. It then discussed methods for measuring the technical efficiency of a new mine. The findings indicated that estimating the production rate and costs requires determination of the efficiency of the mine before starting to operate. This will minimize the risk of failing once the new mine starts operation. The new mine will face challenges and competitiveness from other producers of the same commodity. Its survival will depend on being efficient in regard to best practices; otherwise, it will fail and incur financial losses owing to overuse of inputs such as capital.

The concept of lowest-cost producer, on the other hand, does not necessarily ensure efficiency of a mine operation. The need for end-to-end evaluation of the mine efficiency was identified. This will help the new mine to iteratively optimise its technical inputs to achieve best practices.

A method based on DEA was reviewed and proposed for measuring the technical efficiency of existing and new mines (Budeba et al. 2015, 2014). The reasons for its choice over other methods were that the method allows evaluation of the efficiency of mines by considering multiple inputs and multiple outputs, it does not assume mathematical functions, and it can model efficiency for each subsystem of given system of the mine producing and supplying coal to the market.

The DEA method has been applied in surface coal mines studies, but much of the work has been based on general inputs and outputs, which make it something of a black box. Some of the applications from previous research include evaluating mine production using inputs such as machinery to produce the output tonnage. Most studies have not detailed an end-to-end investigation of coal mining production as an efficient system considering its subsystems. In addition, the influence of non-controllable variables requires investigation.

It can be concluded that the variables that were used in previous DEA studies for surface coal mines and those influencing the efficiency of mine competitiveness identified in Chapter 1 will be used in model formulation in Chapter 4 and data compilation and simulation of the mining supply systems used in this study in Chapter 3.

Chapter 3

Source of data and simulation

The use of the Data Envelopment Analysis (DEA) method requires data for evaluation purposes. For this research, data were subscribed from the Raw Material Group (RMG) database for coal commodity. RMG is database that collects data from most metal and energy mining companies in the world. It is a reliable database enough for this study.

To extract the data from RMG for this study, a list of the variables that influences the efficiency and performance of surface coal mining was compiled by the author from different sources including literature study of the published articles showing the variables influencing the coal mine production (Shafiee and Topal 2012; Shafiee et al. 2009; Chan 2008; Haftendorn et al. 2012; Höök et al. 2010). Others include: technical reports (International Energy Agency 2011), mining annual reports, feasibility studies reports for coal mine projects and governments and analyst reports (CPA Australia 2011). The compiled list includes those variables that influence the efficiency of extraction, washing and sales of coal to the port.

In this research, a total of 16 coal mines consisting of those producing and supplying coal to both export and local markets, export only, and local markets only were extracted from the RMG database. Some mines were found to have most of the required variables, although there were a few missing values, which were supplemented from mining reports, company websites, and government reports. The extracted data were given DMU numbers in consideration of the sensitivity of some of the information to the mining industry.

The number of mines extracted from the RMG database was insufficient for use in the DEA method for this research because the number of variables of the extracted mines was larger than the number of mines themselves. This caused poor discrimination among the mines when evaluating them using the DEA method. The results of a study using fewer Decision Making Units (DMUs) than the number of variables would show all the DMUs to be efficient.

In this research, for example, there were 7 mines producing and supplying coal to both export and local markets, where the number of variables for each of the DMUs were 10 inputs and 2 outputs. This limited the use of the DEA method for the current research, and hence there was a need to increase the number of DMUs for this study to attain a number of samples larger than the minimum required for the study. This was achieved with the use of a simulation that generated data that maintained the characteristics of the collected samples such as correlation among the variables of the samples. The simulated data representing real-world mines was used for the DEA method for this research.

To determine the minimum number of DMUs required for the DEA study, Cooper et al. (2007) explain one of the rules of thumb for determining the minimum number of DMUs in relation to the number of inputs and outputs needed for the DEA study: for a given n number of DMUs, m number of inputs, and s number of outputs, the authors state that the condition $n \geq \max[m \times s, 3 \times (m + s)]$ should be applied. For example, if we apply this rule to producing and supplying coal to both local and export markets by mines, the minimum number of DMUs required for use in the DEA method should be 36.

It is also the case that coal mines are situated in regions with similar characteristics. For example, in South Africa, most mines are located in Mpumalanga, whereas in Australia, most mines are located in Queensland and New South Wales. A few mines can represent the generic characteristics of other mines in the regions from which they are drawn, and therefore these few samples can be used in simulations.

3.1 Data compilation

In this research, the data subscribed by the candidate in 2013–2014 are shown in Tables 3.1–3.6, and a declaration of subscription is attached in Appendix A (IntierraRMG 2014). The data consist of variables for mines producing thermal coal and supplying it to the stated markets. Tables 3.1 and 3.2 give the data extracted from mines producing and supplying thermal coal to both export and local markets. Tables 3.3 and 3.4 show the data extracted from mines producing and supplying coal to the export market only and Tables 3.5 and 3.6 represent some mines producing coal for the local market only.

Table 3.1: Mine design and coal production variables

DMUs	CAPEX (US$M)	Stripping ratio	Run-of-Mine (Mt/yr)	Capacity of plant (Mt/yr)	Age (yrs)	Local supply (Mt/yr)	Export supply (Mt/yr)	Employees
DMU1	631.33	2.0	11.07	15.0	8	7.70	4.80	550
DMU2	502.71	5.0	8.89	6.0	19	2.00	5.40	325
DMU3	15.88	10.3	14.80	11.0	17	1.81	9.90	704
DMU4	62.18	13.2	8.31	7.8	9	0.84	6.99	682
DMU5	2.00	7.8	0.60	4.0	2	0.30	0.30	94
DMU6	8.31	2.4	1.24	1.2	4	1.28	0.58	210
DMU7	424.99	5.2	4.00	2.4	3	0.55	2.30	263

Table 3.2: Coal deposit-specific variables

DMUs	Ash (%)	Moisture (%)	Distance to port (Km)	Precipitation (mm)	Calorific value (MJ/kg)	Thickness (m)
DMU1	26.5	9.0	262.0	630	26.1	15.0
DMU2	10.1	11.0	275.0	656	27.8	5.5
DMU3	5.5	15.5	41.5	2809	25.8	8.7
DMU4	6.0	16.0	517.0	2905	28.9	5.0
DMU5	5.5	13.5	98.0	2121	27.6	3.5
DMU6	25.0	10.0	951.2	688	20.0	10.0
DMU7	13.3	2.9	570.0	683	27.8	3.0

Table 3.3: Mine design and coal production variables for mining supply system to export only

DMUs	CAPEX (US$M)	Stripping ratio	Run-of-Mine (Mt/yr)	Capacity of plant (Mt/yr)	Age (yrs)	Export supply (Mt/yr)	Employees
DMU1	90.3	7.0	2.80	2.8	9	2.8	279
DMU2	1355.8	3.2	8.21	12.0	4	8.2	887
DMU3	2.7	4.2	1.20	1.5	6	1.0	54
DMU4	260.7	7.0	2.00	1.8	4	1.4	550
DMU5	167.7	7.0	4.15	5.5	3	3.5	500

Table 3.4: Coal deposit-specific variables for mining supply system to export only

DMUs	Ash (%)	Moisture (%)	Distance to port (Km)	Precipitation (mm)	Calorific value (MJ/kg)	Thickness (m)
DMU1	6.00	3.5	410	676	31.00	3.2
DMU2	9.50	14.5	380	663	27.90	38.0
DMU3	8.75	11.0	320	643	29.50	11.5
DMU4	11.00	8.0	380	673	27.85	1.5
DMU5	14.00	9.0	120	640	26.20	10.4

Table 3.5: Selected mine design and coal production variables for mining supply system to local market only

DMUs	CAPEX (US$M)	Stripping ratio	Run-of-Mine (Mt/yr)	Age (yrs)	Employees
DMU1	62.76	10	2.80	7	–
DMU2	31.42	4	0.72	7	–
DMU3	4.96	4	1.20	6	–
DMU4	50.98	3	5.40	10	–

Table 3.6: Coal deposit-specific variables for mining supply system to local market only

DMUs	Ash (%)	Moisture (%)	Precipitation (mm)	Calorific value (MJ/kg)	Thickness (m)
DMU1	25.0	8.3	623.0	21.58	3.0
DMU2	24.4	3.7	693.0	22.50	5.0
DMU3	23.1	3.5	634.5	23.40	3.0
DMU4	13.2	3.0	689.0	19.20	5.6

3.2 Statistics of data collected and choice of simulation method

To conduct simulations of data for each mining supply system, statistics for inferring the distribution of each of the variables was done in order to choose the optimum method for simulating the data. The statistics estimated were skewness and kurtosis. Skewness measures the symmetry of data, whereas kurtosis explains the flatness of the distribution. These were used as normality test instead of other methods because the sample size was too small that required bootstrapping to infer the distribution of the samples and then test for normality. Bootstrap generates repetitive values which are known as ties with the samples, these are problems to methods such as Anderson-Darling test of normality which severely affected by these ties due to poor precision Machiwal and Jha (2012).

Bootstrapping is the sampling process used to estimate the distribution of a certain statistic of interest, such as the mean of a variable, by using the original sample obtained from a given population without assuming the distribution of that population. The original sample is assumed to be a population, and it is resampled with replacement to create other samples known as bootstrap samples. The size of each bootstrap sample is the same as that of the original sample. The bootstrap samples are then used to estimate the statistic of interest (Hesterberg 2015). A bootstrap technique was used to estimate skewness and kurtosis for the data collected.

The rule of thumb for interpreting skewness and kurtosis is, if the value of skewness is zero, then the distribution of the variable is normal; otherwise, it is a non-normal distribution. If the kurtosis value is 3, then the distribution of the variable is normal. A kurtosis value of either less than or greater than 3 indicates a non-normal distribution (DeCarlo 1997).

To illustrate the determination of the skewness and kurtosis of each variable of the collected data, consider a sample of mining supply systems $\boldsymbol{X}$ of size n number of DMUs and $\boldsymbol{X} = \{x_1, x_2, ..., x_n\}$. Each DMU has multiple variables. Algorithm 3.1 is applied to estimate the skewness and kurtosis of each variable shown for each mining supply system presented in subsection 3.1. The algorithm was created by adapting the method presented by Efron and Tibshirani (1994).

Algorithm 3.1 Bootstrap for testing normality of data

Require: Input file of the original sample data

1: Take a sample $\boldsymbol{X}_i = \{x_1, x_2, ..., x_n\}$. Consider $\boldsymbol{X}_i$ as a population. Each element of sample $\boldsymbol{X}_i$ has a $\frac{1}{n}$ probability of being chosen.
2: **for** $b = 1 : B$ **do**
3: Sample the members of $\boldsymbol{X}_i$ with replacement to generate bootstrap sample $\boldsymbol{X}_b^*$.
4: Compute the skewness and kurtosis of each variable for each bootstrap sample b.
5: **end for**
6: Generate the approximate distribution of skewness and kurtosis using the results of B samples obtained in step 4.
7: Estimate the mean skewness and kurtosis and interpret the results.

The results from the application of Algorithm 3.1 are presented in Tables 3.7–3.9. The results show that mean skewness of the variables was either negative or positive greater than 0, indicating non-normality. The mean kurtosis values were positive and greater or less than 3, which indicates that the variables of the data were obtained from non-normal distributions of their populations. The property of non-normality of the variables suggests the choice of a multivariate non-normal simulation method for generating data that are conditioned to the properties of samples of mining supply systems presented in Tables 3.1–3.6.

Table 3.7: Results for 1000 bootstrap variables for mining supply system for both export and local markets

Variable	Mean skewness	Mean kurtosis
Run-off mine	0.02	1.99
Calorific value	-0.88	2.85
Thickness	0.56	2.24
Plant capacity	0.40	2.14
CAPEX	0.45	1.97
Stripping ratio	0.34	2.05
Precipitation	0.41	1.96
Age	0.50	2.15
Ash	0.65	2.23
Moisture	-0.37	2.24
Distance-port	0.46	2.24
Export	0.29	2.18
Local supply	0.96	2.97

Table 3.8: Results for 1000 bootstrap variables for mining supply system for export market only

Variable	Mean skewness	Mean kurtosis
Run-off mine	0.56	2.11
Calorific value	0.14	1.96
Thickness	0.56	2.19
Plant capacity	0.73	2.18
CAPEX	0.71	2.28
Stripping ratio	-0.43	1.97
Precipitation	-0.12	1.84
Age	0.52	2.09
Ash	0.09	2.00
Moisture	-0.04	2.03
Distance-port	-0.66	2.23
Export	0.60	2.12

Table 3.9: Results for 1000 bootstrap variables for mining supply system for local market only

Variables	Mean skewness	Mean kurtosis
Run off mine	0.40	1.97
Calorific value	-0.32	1.97
Thickness	0.10	1.84
CAPEX	-0.23	1.90
Striping ratio	0.35	2.24
Precipitation	0.01	1.84
Age	0.33	2.13
Ash	-0.60	2.16
Moisture	0.40	2.23

3.3 Simulation of multivariate data for the research

The method applied in this research to generate multivariate non-normal data for the DEA method was proposed by Ruscio and Kaczetow (2008), which is known as *Sample and Iterate* (Sample and Iterate (SI)). This method generates data that produces a correlation matrix that is approximate to the correlation matrix of the original samples used.

The reasons for the choice of this method were as follows: It does not require specification of the moments (mean, variance, kurtosis, and skewness); rather, it iterates to restore the correlation matrix; it can be used in sampling discrete distributions to distinguish populations with equal moments; it does not require boundary conditions for defining moments; and it can handle distribution with undefined moments.

The simulation of data in this research was carried out using Algorithm 3.2.

Algorithm 3.2 Simulation algorithm for evaluating the DEA model

Require: Input file with raw data

1: Prepare the dataset in a format for R-software compatibility.
2: Determine the correlation among the variables in the original dataset, referred to as a target correlation matrix.
3: **for** $i = 1 : n$ **do**
4: Generate non-normal data using the iterative method proposed by Ruscio and Kaczetow (2008) and compute the correlation matrix of the variables of the resulting data.
5: Compare correlation matrices obtained in steps 2 and 4 based on Root Mean Square Residuals (RMSR).
6: **end for**
7: Extract the data that generate minimum RMSR in step 5.

3.3.1 Simulation of data for mining supply systems for both local and export markets

Generating primary data starts with the determination of the correlation matrix obtained from the variables of the collected samples of mining supply systems indicated in Tables 3.1 and 3.2. The correlation matrix is an input into the SI method. Table 3.10 shows the matrix of correlation of variables of the original samples (target matrix). The strongest correlation is between Run-of-Mine (ROM) and export tonnage, which is 0.943, meaning that the run-off mine is related to the export tonnage by 94.3%. The matrix of the original sample was used to simulate 60 DMUs, which were extracted for use in this research. The matrix resulting from the simulated datasets is presented in Table 3.11. The simulated samples that gave a minimum RMSR were chosen and extracted for use in the study.

Table 3.10: Correlation of sample mines' variables for both local and export markets

	Run-of-Mine	Calorific value	Thickness	Plant capacity	CAPEX	Stripping ratio	Precipitation	Age	Ash	Moisture	Distance-port	Export	Local supply
Run-of-Mine	1												
Calorific value	0.267	1											
Thickness	0.426	-0.515	1										
Plant capacity	0.834	0.259	0.654	1									
CAPEX	0.28	0.286	0.301	0.376	1								
Stripping ratio	0.28	0.541	-0.495	0.099	-0.561	1							
Precipitation	0.298	0.346	-0.277	0.217	-0.693	0.936	1						
Age	0.791	0.204	0.14	0.458	0.194	0.234	0.173	1					
Ash	-0.14	-0.657	0.747	0.101	0.437	-0.85	-0.747	-0.302	1				
Moisture	0.335	0.107	0.018	0.314	-0.614	0.694	0.817	0.405	-0.525	1			
Distance-port	-0.521	-0.584	0.036	-0.58	-0.1	-0.337	-0.42	-0.44	0.54	-0.4	1		
Export	0.943	0.336	0.192	0.679	0.034	0.553	0.529	0.799	-0.379	0.501	-0.463	1	
Local supply	0.49	-0.086	0.868	0.794	0.663	-0.498	-0.372	0.164	0.645	-0.143	-0.201	0.205	1

Table 3.11: Correlation matrix for simulated data for both local and export markets

	Run-of-Mine	Calorific value	Thickness	Plant capacity	CAPEX	Stripping ratio	Precipitation	Age	Ash	Moisture	Distance-port	Export	Local supply
Run-of-Mine	1												
Calorific value	0.411	1											
Thickness	0.333	-0.435	1										
Plant capacity	0.897	0.273	0.502	1									
CAPEX	0.179	0.215	0.194	0.317	1								
Stripping ratio	0.239	0.408	-0.426	0.048	-0.486	1							
Precipitation	0.305	0.266	-0.195	0.14	-0.525	0.856	1						
Age	0.689	0.419	0.155	0.616	0.084	0.318	0.319	1					
Ash	-0.171	-0.559	0.641	-0.016	0.374	-0.739	-0.568	-0.298	1				
Moisture	0.387	0.07	-0.021	0.269	-0.606	0.636	0.663	0.333	-0.543	1			
Distance-port	-0.642	-0.588	0.08	-0.483	-0.132	-0.421	-0.369	-0.446	0.468	-0.323	1		
Export	0.857	0.415	0.16	0.741	-0.041	0.463	0.524	0.641	-0.353	0.466	-0.6	1	
Local supply	0.427	-0.035	0.686	0.595	0.538	-0.422	-0.337	0.159	0.529	-0.261	-0.09	0.175	1

The resulting RMSR between the correlation matrices for the samples and simulated data is 0.094. This implies that the correlation matrix of the simulated data has an average error of 9.4% from the target correlation matrix.

3.3.2 Simulation of data for mining supply systems for export markets

The correlation matrix of the original datasets representing mines producing coal and supplying it to the export market only is shown in Table 3.12, whereas that for 50 simulated DMUs is shown in Table 3.13. The strongest relationship in the original sample of 0.995 was between ROM and Export variables. A comparison between the correlation matrices indicates an RMSR of 0.122, which implies an error of 12.2% for the matrix of simulated data from the matrix of the original datasets collected.

Table 3.12: Correlation for sample mines' variables for export markets

	Run-of-Mine	Calorific value	Thickness	Plant capacity	CAPEX	Stripping ratio	Precipitation	Age	Ash	Moisture	Distance-port	Export
Run-of-Mine	1											
Calorific value	-0.367	1										
Thickness	0.886	-0.257	1									
Plant capacity	0.994	-0.386	0.925	1								
CAPEX	0.93	-0.276	0.91	0.933	1							
Stripping ratio	-0.508	-0.044	-0.844	-0.58	-0.654	1						
Precipitation	0.056	0.481	-0.157	-0.026	0.198	0.24	1					
Age	-0.371	0.967	-0.349	-0.404	-0.37	0.132	0.451	1				
Ash	0.14	-0.968	0.038	0.161	0.026	0.203	-0.569	-0.91	1			
Moisture	0.594	-0.488	0.843	0.667	0.704	-0.846	-0.435	-0.64	0.334	1		
Distance-port	-0.002	0.711	0.065	-0.033	0.259	-0.248	0.815	0.568	-0.815	-0.119	1	
Export	0.995	-0.272	0.9	0.99	0.933	-0.549	0.091	-0.283	0.041	0.578	0.07	1

Table 3.13: Correlation for simulated data for export markets

	Run-of-Mine	Calorific value	Thickness	Plant capacity	CAPEX	Stripping ratio	Precipitation	Age	Ash	Moisture	Distance-port	Export
Run-of-Mine	1											
Calorific value	-0.325	1										
Thickness	0.925	-0.316	1									
Plant capacity	0.889	-0.413	0.925	1								
CAPEX	0.872	-0.187	0.926	0.856	1							
Stripping ratio	-0.678	0.097	-0.621	-0.697	-0.713	1						
Precipitation	-0.053	0.58	-0.041	-0.07	0.037	-0.124	1					
Age	-0.307	0.896	-0.325	-0.406	-0.181	0.1	0.542	1				
Ash	0.148	-0.915	0.134	0.207	0.002	0.081	-0.593	-0.787	1			
Moisture	0.787	-0.571	0.736	0.828	0.662	-0.598	-0.174	-0.573	0.406	1		
Distance-port	-0.081	0.727	-0.037	-0.07	0.042	-0.062	0.523	0.55	-0.883	-0.286	1	
Export	0.874	-0.231	0.908	0.875	0.86	-0.719	-0.027	-0.243	0.01	0.69	0.079	1

3.3.3 Simulation of data for mines producing coal for local markets only

The resulting correlation matrices for the datasets collected from mines producing coal for the local market only and those collected from 30 generated DMUs are presented in Table 3.14 and Table 3.15, respectively. The strongest correlation in the original datasets is highlighted, which is 0.999, which indicates a relationship between Stripping Ratio (SR) and moisture. The RMSR between the matrices for the original and simulated data is 0.12. This indicates that the average error of the matrix of simulated data from the target matrix is 12%.

Table 3.14: Correlation matrix for simulated mines supplying coal to local markets

	Run-of-Mine	Calorific value	Thickness	CAPEX	Stripping ratio	Precipitation	Age	Ash	Moisture
Run-of-Mine	1								
Calorific value	-0.955	1							
Thickness	0.45	-0.651	1						
CAPEX	0.624	-0.708	0.222	1					
Stripping ratio	-0.062	0.111	-0.644	0.568	1				
Precipitation	0.207	-0.432	0.966	0.031	-0.718	1			
Age	0.907	-0.974	0.784	0.556	-0.331	0.598	1		
Ash	-0.863	0.845	-0.688	-0.222	0.558	-0.523	-0.915	1	
Moisture	-0.032	0.072	-0.608	0.605	0.999	-0.689	-0.292	0.533	1

Table 3.15: Correlation matrix for simulated mines supplying coal to local markets

	Run-of-Mine	Calorific value	Thickness	CAPEX	Stripping ratio	Precipitation	Age	Ash	Moisture
Run-of-Mine	1								
Calorific value	-0.751	1							
Thickness	0.545	-0.489	1						
CAPEX	0.636	-0.661	0.063	1					
Stripping ratio	-0.052	0.227	-0.671	0.391	1				
Precipitation	0.323	-0.448	0.667	-0.058	-0.728	1			
Age	0.737	-0.812	0.571	0.494	-0.309	0.418	1		
Ash	-0.678	0.734	-0.707	-0.246	0.521	-0.592	-0.818	1	
Moisture	-0.044	0.198	-0.6	0.413	0.936	-0.677	-0.294	0.47	1

3.4 Secondary attributes of the mining supply systems

Secondary attributes were generated through computation using the available developed mathematical equations. The attributes included the number of employees, the actual price of coal based on adjustments for the specified calorific value of the coal, and the net revenue obtained from coal sales.

3.4.1 Number of employees

The number of personnel for each system is related to the ROM or tonnage of material used in the specific section of the mining supply system. The approach used for each of the mining supply systems was that proposed by Hustrulid and Kuchta (2006) for soft rocks. Therefore, the number of personnel was computed using the following information:

T is the total tons of coal crushed/day.

$T_{\mathbf{p}}$ is the total tons of waste and coal mined/day.

$T_{\mathbf{o}}$ is the total tons of coal mined/day.

$T_{\mathbf{w}}$ is the total tons of overburden mined/day.

$N_{\mathbf{op}}$ is the number of mine personnel (operators).

$N_{\mathbf{ml}}$ is the process crew size.

$N_{\mathbf{sv}}$ is the number of service personnel required for open pit mining.

$N_{\mathbf{at}}$ is the number of administrative and technical personnel required for mining, crushing, and services.

Assume that the mine operates 5 days/week and the mill operates 7 days/week. The relationship between tons and the number of personnel is given by Equations 3.1–3.5 (Hustrulid and Kuchta 2006).

$$N_{\mathrm{op}} = 0.024T_{\mathrm{p}}^{0.8} \quad \text{for competent soft rock in this case is coal} \tag{3.1}$$

$$T = 0.71 \times T_{\mathrm{o}} \tag{3.2}$$

$$N_{\mathrm{ml}} = 7.2T^{0.3} \tag{3.3}$$

$$N_{\mathrm{sv}} = 0.254(N_{\mathrm{op}} + N_{\mathrm{ml}}) \tag{3.4}$$

$$N_{\mathrm{at}} = 0.11(N_{\mathrm{op}} + N_{\mathrm{ml}} + N_{\mathrm{sv}}) \tag{3.5}$$

3.4.2 Revenue and carbon emission

In most cases, export coal is sold on the basis of the Free on Board (FOB) price. This is the price given to the seller upon delivering the coal to the buyer's vessel. The FOB price is corrected to take into account the energy content expressed in terms of Calorific Value (CV). The approach to correcting the FOB price was highlighted by Docker (2011). The calculation of CO_2 emissions is shown in Equation 3.8 (Australian Government Department of the Environment 2014). Consider the following definitions:

CV_s is the minimum calorific value required for export (MJ/kg).

CV is the calorific value of the tonnage of coal produced (MJ/kg).

FOB_s is the price that can be offered for a specified standard calorific value (US$/tonne).

CO_2 is the allowed limit of carbon emission. The excess emission is a penalty that is imposed by reducing the gross revenue obtained from selling thermal coal.

EF is the emission factor.

Therefore, the Net Revenue (NR) (US$M) is given by Equation (3.6), the FOB price for coal of specific quality can be obtained by using Equation (3.7), and the CO_2 Emission Penalty (EP) is represented by Equation (3.9).

$$NR = \text{Total revenue} - \text{emission penalty} - \text{royalty} \tag{3.6}$$

$$FOB = \frac{CV_s}{CV} \times FOB_s \tag{3.7}$$

$$\mathrm{CO_2} \quad \text{emission in tonnes} = ROM \times EF \tag{3.8}$$

$$EP = \text{Carbon tax per tonne} \times \text{excess} \quad \mathrm{CO_2} \quad \text{emission in tonnes} \tag{3.9}$$

The price of thermal coal for three years covering the period for which the data were subscribed was used. The average price of coal used was $90/tonne for the period from 2012 to 2014, as shown in the Figure 3.1 (Kolesnikov 2015). This was used as the benchmark price for all mines in the study. The average price for exported coal was adjusted to account for the calorific value and to determine the NR. For local market supply, the coal price used was US$37/tonne, which is an average of US$35, US$37, and US$39 for the period from 2012 to 2014 (Angloamerican 2013, 2012). The minimum calorific value specification for exported coal was considered to be 24.5 MJ/kg (5,850 kcal/kg NCV) (globalCOAL 2015).

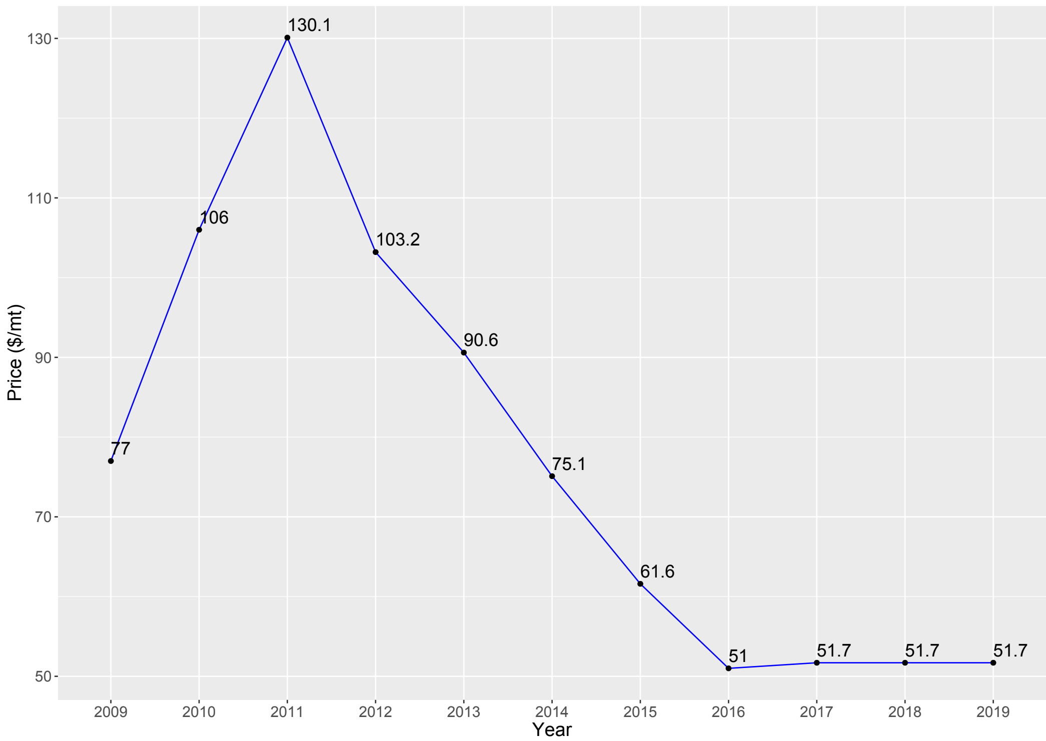

Figure 3.1: Coal prices for export in Australia (Source: IMF commodity price forecasts data).

Both primary simulated data and the secondary attributes of the mining supply systems were compiled to form a set of data, which are given in Appendices B.1–B.3 of this report.

3.5 Conclusion

Simulated data for the major variables influencing the competitiveness of surface coal mines producing and supplying thermal coal to the market were generated in this chapter. The data were generated using the SI method for non-normal multivariate data. The intention was to generate data that maintained the characteristics of the original samples using correlation matrices and to provide the number of DMUs that satisfy the minimum requirements for applying the DEA method in order to achieve discrimination among the DMUs.

For all datasets generated by the simulations, two samples of DMUs for coal supply systems for both local and export markets were found to be outliers because their export capacities were higher than the maximum capacity of a washing plant per year. For this reason, they were excluded from the study, leaving 58 DMUs for use in the study. The other simulated data had no outliers. The used simulated samples for the extraction and supply of coal to export markets only were 50 DMUs and 30 DMUs for the supply of coal to local markets only.

Data envelopment analysis requires data for measuring the efficiency of an entity such as coal mining operations. The data simulated in this chapter and real mines extracted from RMG database forms the group of mines for DEA study. Therefore, this chapter will be followed by Chapter 4 which uses the variables of compiled data to formulated the models for measuring technical efficiency of surface coal mining. The variables of generated data in this chapter are used in Chapter 4 about the formulation of the models. Then, the data will be used to evaluate the models as discussed in Chapter 5.

Chapter 4

Modelling the efficiency and performance of a surface coal mine

Modelling refers to the process of producing an imitation or an abstract representation or artefact of a specific system or object. A model is a close approximation of a real system, but it is simpler than the system itself (Maria 1997; Gupta and Grover 2013). Models are used to solve complex problems about a given system. This is done through abstraction and construction of a model of such a system using its properties and then developing solutions for a stated problem. There two main types of models. The first is a physical model, which is a prototype of a real object or system. For example, dragline mining equipment can be represented as a three-dimensional physical object mimicking the features and operational properties of an actual dragline. The second type is an analytical model, which seeks to explain the behaviour of a system in mathematical language or in the form of a computer program.

This chapter describes the mathematical modelling of the technical efficiency and performance of a surface mine producing and supplying thermal coal to the market. The chapter is organised into two parts. First, Data Envelopment Analysis (DEA) is applied to formulate the mathematical models for evaluating the relative technical efficiency of a mine producing and supplying thermal coal to the market. Secondly, mathematical models are developed to predict the technical efficiency and performance of a mine.

Mine production and supply systems of thermal coal are considered to consist of subsystems that include extraction, washing, and delivery to the port where the coal is sold to the export market. In this research, the structure of the production and supply of thermal coal will be referred to as a *mining supply system*. The system is considered to use discretionary input variables that can be controlled by mine management to generate target outputs while operating under non-discretionary conditions that cannot be controlled. Both discretionary and non-discretionary variables will be considered in the formulation of the models. To formulate the models, the following structures of mining supply systems are first explained:

Combined System for Local and Export (CSLE) A system that involves the extraction of coal, washing to remove ash and other inorganic matter, and transporting it to the port for the export market. It also supplies coal to nearby power plants for electricity generation.

Export Coal Mine Supply (ECMS) A system that involves the extraction of coal, washing, and then supplying it to the export market only.

Local Coal Mine Supply (LCMS) A system that extracts coal and supplies it to nearby power plants that produce electricity. The supply of coal to buyers at the mine is also known as a *mine mouth system.*

4.1 Formulation of models for measuring the technical efficiency of mining supply systems

Consider Figure 4.1 to represent the generic structure of a system producing and supplying thermal coal to the markets. It is referred to as a mining supply system, and it consists of extraction, washing, and port operations as subsystems. This structure is used to describe the models for evaluating the technical efficiency of a given mining supply system, as detailed in subsections 4.2–4.5.

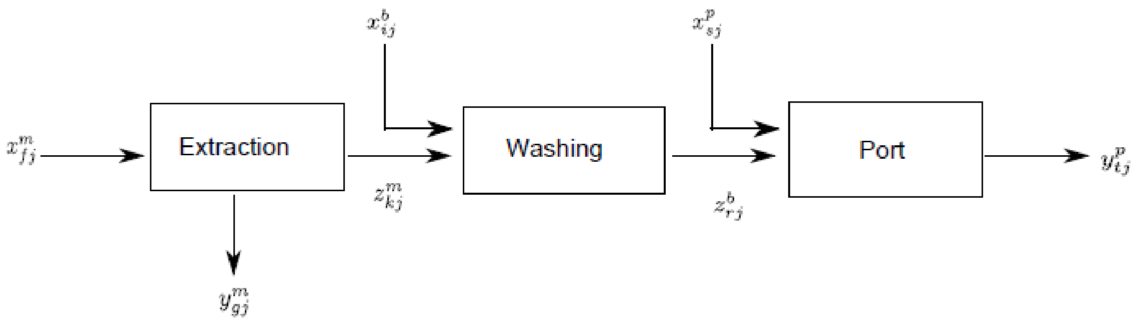

Figure 4.1: Generic coal mining supply system for mine $j \in \boldsymbol{J}$.

We can represent the set of subsystems with Figure 4.1. Consider the extraction subsystem to be denoted by superscript m, the washing subsystem to be represented by superscript b, and the port subsystem to be denoted by superscript p. Let $\boldsymbol{F} = \{1, ..., F\}$ denote the set of inputs to the extraction subsystem and $\boldsymbol{K} = \{1, ..., K\}$ the set of intermediate outputs from the extraction subsystem to the washing subsystem. Also, let $\boldsymbol{I} = \{1, ..., I\}$ denote the set of inputs at the beginning of the washing subsystem and $\boldsymbol{R} = \{1, ..., R\}$ the set of intermediate outputs from the washing subsystem to the port subsystem. Finally, let $\boldsymbol{S} = \{1, ..., S\}$ denote the set of inputs at the beginning of the port subsystem and $\boldsymbol{T} = \{1, ..., T\}$ the set of outputs from the port subsystem.

Consider a set of surface mining supply systems $\boldsymbol{J} = \{1, ..., n\}$. Each mining supply system, $j \in \boldsymbol{J}$, is regarded to be a Decision Making Unit (DMU) that produces coal and supplies it to a specific market. The definitions of variables shown in Figure 4.1 and other notations used to explain the models are as follows Budeba et al. (2016):

x_{fj}^m is the given input $f \in \boldsymbol{F}$ by the extraction subsystem (m) of mining supply system $j \in \boldsymbol{J}$. The inputs include: Capital Expenditure (CAPEX) (US\$M), number of employees in the mining operation, and Stripping Ratio (SR).

υ_f^m is the weight of input $f \in \boldsymbol{F}$.

y_{gj}^m is the given output $g \in \boldsymbol{G}$ generated by the extraction subsystem (m) of a mining supply system $j \in \boldsymbol{J}$ to the local market. The output is revenue (US\$M) from the sale of coal.

ν_g^m is the weight of output $g \in \boldsymbol{G}$.

z_{kj}^m is the given intermediate output $k \in \boldsymbol{K}$ from the extraction subsystem (m) of mining supply system $j \in \boldsymbol{J}$ that will be used in the washing subsystem (b). The intermediate outputs in this subsystem include Run-of-Mine (ROM) (Mt/yr), ash (%), and moisture (%).

η_k^m is the weight of intermediate output $k \in \boldsymbol{K}$.

x_{ij}^b is the given input $i \in \boldsymbol{I}$ at the beginning of the washing subsystem (b) of a mining supply system $j \in \boldsymbol{J}$. (Examples of these inputs are plant capacity (Mt/yr), administration, and number of employees in the washing plant.)

υ_i^b is the weight of input $i \in \boldsymbol{I}$.

z_{rj}^b is the given intermediate output $r \in \boldsymbol{R}$ from the washing subsystem (b) and the usage to the port subsystem (p) of a coal mining supply system $j \in \boldsymbol{J}$. The intermediate output in this subsystem is the tonnage of clean coal for export (Mt/yr).

η_r^b is the weight of intermediate output $r \in \boldsymbol{R}$.

x_{sj}^p is the given input $s \in \boldsymbol{S}$ at the beginning of the port subsystem (p) of a mining supply system $j \in \boldsymbol{J}$. The input at the beginning of the port is the allowable carbon emission (tonnes).

υ_s^p is the weight of input $s \in \boldsymbol{S}$.

y_{tj}^p is the given output $t \in \boldsymbol{T}$ generated by the port subsystem (p) of a mining supply system $j \in \boldsymbol{J}$. The output is the revenue (US$M) from the sales of coal to the export market.

ν_t^p is the weight of output $t \in \boldsymbol{T}$.

θ_j is the overall efficiency of a mining supply system $j \in \boldsymbol{J}$.

θ_j^m is the efficiency of the extraction subsystem (m) of a mining supply system $j \in \boldsymbol{J}$.

θ_j^b is the efficiency of the washing subsystem (b) of a mining supply system $j \in \boldsymbol{J}$.

θ_j^p is the efficiency of the port subsystem (p) of a mining supply system $j \in \boldsymbol{J}$.

σ_j^m is the ratio of the weighted inputs in the extraction subsystem (m) to the total weighted inputs of the whole mining supply system $j \in \boldsymbol{J}$.

σ_j^b is the ratio of the weighted inputs in the washing subsystem (b) to the total weighted inputs of the whole mining supply system $j \in \boldsymbol{J}$.

σ_j^p is the ratio of the weighted inputs in the port subsystem (p) to the total weighted inputs in the whole mining supply system $j \in \boldsymbol{J}$.

4.2 Combined System for Local and Export (CSLE) model

The CSLE model is composed of surface coal mines producing and supplying coal to local and export markets. The coal supply chain for this model consists of extraction, washing (cleaning), and port operations. These subsystems are considered to be interdependent, and they determine the overall technical efficiency of the mining supply system.

Cook et al. (2010) suggested that the overall technical efficiency of a system involving subsystems is the convex linear combination of stage-level measures. This implies that the sum of the weighted efficiency of each subsystem gives the overall efficiency of the system. The weight of a subsystem is the ratio of the inputs of that subsystem to the overall inputs of the whole system.

In this case, the mathematical representation of the CSLE model for measuring the technical efficiency of the mining supply system is derived from the generic structure of the surface coal mine supply system, represented in Figure 4.1. The technical efficiency of the system and its subsystems is defined using Equation (4.1).

$$\text{Efficiency} = \frac{\text{Weighted sum of outputs}}{\text{Weighted sum of inputs}} \tag{4.1}$$

To formulate the models, consider the composition of the weighted inputs and outputs and other conditions for a mining supply system denoted by DMU_o as follows:

1. The total weighted inputs to the extraction subsystem are given by Equation (4.2).

$$\sum_{f \in \boldsymbol{F}} \upsilon_f^m x_{fo}^m \tag{4.2}$$

2. The total weighted inputs to the washing subsystem are the combination of the weighted inputs at the beginning of the washing subsystem and the weighted output from the extraction subsystem, given by Equation (4.3).

$$\sum_{i \in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum_{k \in \boldsymbol{K}} \eta_k^m z_{ko}^m \tag{4.3}$$

3. The total weighted inputs to the port subsystem are the sum of the weighted inputs at the beginning of the port subsystem and the weighted output from the washing subsystem, given by Equation (4.4).

$$\sum_{s \in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum_{r \in \boldsymbol{R}} \eta_r^b z_{ro}^b \tag{4.4}$$

4. The total of weighted inputs of the whole surface coal mining supply system is given by Equation (4.5).

$$\sum_{f \in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum_{k \in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum_{i \in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum_{s \in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum_{r \in \boldsymbol{R}} \eta_r^b z_{ro}^b \tag{4.5}$$

5. The value of the efficiency of the whole system and each of subsystem is valid under the conditions given by $0 \leq \theta_j \leq 1$.

6. For each subsystem to be in operation, the total weighted inputs entering such a subsystem must be greater than or equal to a minimum value, say α. This ensures that all subsystems are functioning—if the value of the total weighted inputs is zero, such a subsystem collapses and is deemed as not functioning.

7. The technical efficiency of a coal mine system is influenced by non-discretionary variables, in addition to the discretionary variables that are controllable by the management.

The overall efficiency is formulated from the efficiency of each of the subsystems involved in the supply chain (extraction, washing, and port) using information from items 1–7 above. In this case, the overall efficiency of the mining supply system under investigation, denoted by DMU_o, and the weighted input of each of its subsystems is determined as follows:

$$\text{The overall efficiency of} \quad \mathrm{DMU}_o = \theta_o = \sigma_o^m \theta_o^m + \sigma_o^b \theta_o^b + \sigma_o^p \theta_o^p \tag{4.6}$$

where

$$\sigma_o^m = \frac{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fo}^m}{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b} \tag{4.7}$$

$$\sigma_o^b = \frac{\sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{io}^b}{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b} \tag{4.8}$$

$$\sigma_o^p = \frac{\sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b}{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b} \tag{4.9}$$

Consider the input-oriented Variable Return to Scale (VRS) for each of the subsystems. Substituting the values σ_o^m, σ_o^b, and σ_o^p into Equation (4.6) gives Equation (4.10). Thus, the CSLE model for the determination of technical efficiency is represented by the Fractional Program (FP) in Equations (4.10)–(4.19). FP is difficult to solve, and thus it needs to be transformed into a Linear Program (LP) that can be solved more easily.

$$\max \theta_o = \frac{\sum\limits_{g\in \boldsymbol{G}} \nu_g^m y_{go}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b + \sum\limits_{t\in \boldsymbol{T}} \nu_t^p y_{to}^p + \pi_o^m + \pi_o^b + \pi_o^p}{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b} \tag{4.10}$$

subject to

$$\frac{\sum\limits_{g\in \boldsymbol{G}} \nu_g^m y_{gj}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{kj}^m + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{rj}^b + \sum\limits_{t\in \boldsymbol{T}} \nu_t^p y_{tj}^p + \pi_j^m + \pi_j^b + \pi_j^p}{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fj}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{kj}^m + \sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{ij}^b + \sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{sj}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{rj}^b} \le 1 \qquad \forall j \in \boldsymbol{J} \tag{4.11}$$

$$\frac{\sum\limits_{g\in \boldsymbol{G}} \nu_g^m x_{gj}^m + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{kj}^m + \pi_j^m}{\sum\limits_{f\in \boldsymbol{F}} \upsilon_f^m x_{fj}^m} \le 1 \qquad \forall j \in \boldsymbol{J} \tag{4.12}$$

$$\frac{\sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{rj}^b + \pi_j^b}{\sum\limits_{i\in \boldsymbol{I}} \upsilon_i^b x_{ij}^b + \sum\limits_{k\in \boldsymbol{K}} \eta_k^m z_{kj}^m} \le 1 \qquad \forall j \in \boldsymbol{J} \tag{4.13}$$

$$\frac{\sum\limits_{t\in \boldsymbol{T}} \nu_t^p y_{tj}^p + \pi_j^p}{\sum\limits_{s\in \boldsymbol{S}} \upsilon_s^p x_{sj}^p + \sum\limits_{r\in \boldsymbol{R}} \eta_r^b z_{rj}^b} \le 1 \qquad \forall j \in \boldsymbol{J} \tag{4.14}$$

$$\sum_{f\in \boldsymbol{F}} \upsilon_f^m x_{fj}^m \geq \alpha \qquad \forall j \in \boldsymbol{J} \tag{4.15}$$

$$\sum_{i\in \boldsymbol{I}} \upsilon_i^b x_{ij}^b + \sum_{k\in \boldsymbol{K}} \eta_k^m z_{kj}^m \geq \alpha \qquad \forall j \in \boldsymbol{J} \tag{4.16}$$

$$\sum_{s\in \boldsymbol{S}} \upsilon_s^p x_{sj}^p + \sum_{r\in \boldsymbol{R}} \eta_r^b z_{rj}^b \geq \alpha \qquad \forall j \in \boldsymbol{J} \tag{4.17}$$

$$\upsilon_f^m, \nu_g^m, \eta_k^m, \upsilon_i^b, \eta_r^b, \upsilon_s^p, \nu_t^p, \geq \epsilon \tag{4.18}$$

$$\pi_j^m, \pi_j^b, \pi_j^p \quad \text{are free in sign} \tag{4.19}$$

The variables π_j^m, π_j^b, and π_j^p account for the measure of the return to scale for the extraction, washing, and port subsystems, respectively. A value greater or lesser than zero indicate that a subsystem or the system is undergoing VRS; a value of zero indicates a Constant Return to Scale (CRS).

The transformation from FP to LP is done using the same form as that of the Charnes and Cooper transformation described in subsection 2.4.3. In this case, let

$$t = \frac{1}{\sum_{f\in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum_{k\in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum_{i\in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum_{s\in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum_{r\in \boldsymbol{R}} \eta_r^b z_{ro}^b} \tag{4.20}$$

and therefore, the new weights are obtained after multiplying t by each equation from (4.11)–(4.14) in the numerator and denominator and the right- and left-hand sides of Equations (4.15)–(4.17). The following weights are obtained: $\gamma_k^m = t\eta_k^m$, $\gamma_r^b = t\eta_r^b$ and $\mu_t^p = t\nu_t^p$, $\omega_f^m = t\upsilon_f^m$, $\omega_i^b = t\upsilon_i^b$, $\mu_g^m = t\nu_g^m$ and $\omega_s^p = t\upsilon_s^p$, $u^m = t\pi^m$, $u^b = t\pi^b$, and $u^p = t\pi^p$. The resulting LPs are shown by Equations (4.21)–(4.31).

$$\max \theta_o = \sum_{g\in \boldsymbol{G}} \mu_g^m y_{go}^m + \sum_{k\in \boldsymbol{K}} \gamma_k^m z_{ko}^m + \sum_{r\in \boldsymbol{R}} \gamma_r^b z_{ro}^b + \sum_{t\in \boldsymbol{T}} \mu_t^p y_{to}^p + u_o^m + u_o^b + u_o^p \tag{4.21}$$

subject to

$$\sum_{g\in \boldsymbol{G}} \mu_g^m y_{gj}^m + \sum_{t\in \boldsymbol{T}} \mu_t^p y_{tj}^p - \sum_{f\in \boldsymbol{F}} \omega_f^m x_{fj}^m - \sum_{i\in \boldsymbol{I}} \omega_i^b x_{ij}^b - \sum_{s\in \boldsymbol{S}} \omega_s^p x_{sj}^p + u_j^m + u_j^b + u_j^p \leq 0 \tag{4.22}$$

$$\sum_{f\in \boldsymbol{F}} \omega_f^m x_{fo}^m + \sum_{i\in \boldsymbol{I}} \omega_i^b x_{io}^b + \sum_{s\in \boldsymbol{S}} \omega_s^p x_{so}^p + \sum_{k\in \boldsymbol{K}} \gamma_k^m z_{ko}^m + \sum_{r\in \boldsymbol{R}} \gamma_r^b z_{ro}^b = 1 \tag{4.23}$$

$$\sum_{g\in \boldsymbol{G}} \mu_g^m y_{gj}^m + \sum_{k\in \boldsymbol{K}} \gamma_k^m z_{kj}^m - \sum_{f\in \boldsymbol{F}} \omega_i^m x_{fj}^m + u_j^m \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{4.24}$$

$$\sum_{r\in \boldsymbol{R}} \gamma_r^b z_{rj}^b - \sum_{i\in \boldsymbol{I}} \omega_i^b x_{ij}^b - \sum_{k\in \boldsymbol{K}} \gamma_k^m z_{kj}^m + u_j^b \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{4.25}$$

$$\sum_{t \in \boldsymbol{T}} \mu_t^p y_{tj}^p - \sum_{s \in \boldsymbol{S}} \omega_s^p x_{sj}^p - \sum_{r \in \boldsymbol{R}} \gamma_r^b z_{rj}^b + u_j^p \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{4.26}$$

$$\sum_{f \in \boldsymbol{F}} \omega_f^m x_{fj}^m \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.27}$$

$$\sum_{i \in \boldsymbol{I}} \omega_i^b x_{ij}^b + \sum_{k \in \boldsymbol{K}} \gamma_k^m z_{kj}^m \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.28}$$

$$\sum_{s \in \boldsymbol{S}} \omega_s^p x_{sj}^p + \sum_{r \in \boldsymbol{R}} \gamma_r^b z_{rj}^b \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.29}$$

$$\omega_f^m, \mu_g^m, \gamma_k^m, \omega_i^b, \gamma_r^b, \omega_s^p, \mu_t^p \geq \epsilon \tag{4.30}$$

$$u_j^m, u_j^b, u_j^p \quad \text{are free in sign} \tag{4.31}$$

The variables u_j^m, u_j^b, and u_j^p account for the measure of the return to scale for the extraction, washing, and port subsystems, respectively. These variables are free in sign, and the value of each can be negative or positive. If any variable or all of them in total equal zero, then the system in undergoing CRS; otherwise, it indicates VRS.

The model in Equations (4.21)–(4.31) should then be solved $\|\boldsymbol{J}\|$ times to compute the technical efficiency of each of the $j \in \boldsymbol{J}$ mining supply systems used in the study. The LP model computes the optimal weights of each DMU, which maximises its technical efficiency score relative to the other mining supply systems used in the study. The DMU is termed *efficient* if the efficiency score is equal to 1 and *inefficient* if it is greater than 0 but less than 1.

4.3 Combined System for Local and Export (CSLE) with non-discretionary variables

Non-discretionary variables are those variables that cannot be changed at the discretion of the management. These variables affect the output of the production unit, which in turn influences the efficiency of the DMU. Various approaches to account for the influence of non-discretionary variables on the technical efficiency of a DMU have been discussed in the DEA literature, one of which mentions entering non-discretionary inputs in the objective function of the multiplier DEA model. This is achieved by subtracting the weighted sum of the non-discretionary variables from the weighted sum of outputs of a DMU in the DEA model. The multipliers associated with the non-discretionary variable can be zero if such a variable does not affect the efficiency; otherwise, if the multiplier is greater than zero, such a variable will affect the efficiency score (Lotfi et al. 2007; Cooper et al. 2007).

Another approach is the use of the two-stage method, which involves the computation of the efficiency scores in the first stage using discretionary variables in the DEA model and applying the regression method in the second stage to account for the influence of non-discretionary variables (Banker and Natarajan 2008). Examples of regression methods used in the second stage include Tobit regression and ordinary least squares.

A third approach is the use of the bootstrap technique. This method involves two stages. The first stage is the computation of the efficiency scores using discretionary inputs only; the second stage involves sampling the resulting efficiency scores of each DMU with replacement. The aim is to create independent efficiency scores by eliminating the dependency among the efficiency scores that violates the regression assumption. The bootstrapped efficiency scores are used for truncated regression relating them to the non-discretionary variables. More details about the application of the bootstrap technique for non-discretionary variables can be found in the scholarly work by Xue et al. (1999); Simar and Wilson (2007); Afonso and Aubyn (2006); Nedelea and Fannin (2013).

In the mining context, the observed output is a result of the discretionary inputs and the influence of the non-discretionary variables such as rainfall (precipitation), which can interfere with production operations. This research applies the approach discussed by Lotfi et al. (2007) and Cooper et al. (2007). The method ensures that the resulting efficiency scores account for both discretionary and non-discretionary variables.

Consider the coal mining supply system to be operating in an environment in which there are non-discretionary external variables, denoted by superscript e, and each variable is indexed by $h \in \boldsymbol{H}$. The weighted sum of non-discretionary variables is given by $\sum_{h \in \boldsymbol{H}} \omega_h^e x_{ho}^e$, which is subtracted from the overall weighted output of the mining supply system in Equations (4.32) and (4.33).

$$\max \theta_o = \sum_{g \in \boldsymbol{G}} \mu_g^m y_{gj}^m + \sum_{k \in \boldsymbol{K}} \gamma_k^m z_{ko}^m + \sum_{r \in \boldsymbol{R}} \gamma_r^b z_{ro}^b + \sum_{t \in \boldsymbol{T}} \mu_t^p y_{to}^p - \sum_{h \in \boldsymbol{H}} \omega_h^e x_{ho}^e + u_o^m + u_o^b + u_o^p \tag{4.32}$$

subject to

$$\sum_{g \in \boldsymbol{G}} \mu_g^m y_{gj}^m + \sum_{t \in \boldsymbol{T}} \mu_t^p y_{tj}^p - \sum_{f \in \boldsymbol{F}} \omega_f^m x_{ij}^m - \sum_{i \in \boldsymbol{I}} \omega_i^b x_{ij}^b - \sum_{s \in \boldsymbol{S}} \omega_s^p x_{sj}^p - \sum_{h \in \boldsymbol{H}} \omega_h^e x_{hj}^e + u_j^m + u_j^b + u_j^p \leq 0 \quad \forall j \in \boldsymbol{J} \tag{4.33}$$

$$\sum_{f \in \boldsymbol{F}} \omega_f^m x_{fo}^m + \sum_{k \in \boldsymbol{K}} \gamma_k^m z_{ko}^m + \sum_{i \in \boldsymbol{I}} \omega_i^b x_{io}^b + \sum_{r \in \boldsymbol{R}} \gamma_r^b z_{ro}^b + \sum_{s \in \boldsymbol{S}} \gamma_s^p x_{so}^p = 1 \tag{4.34}$$

$$\sum_{g \in \boldsymbol{G}} \mu_g^m y_{gj}^m + \sum_{k \in \boldsymbol{K}} \gamma_k^m z_{kj}^m - \sum_{f \in \boldsymbol{F}} \omega_f^m x_{fj}^m + u_j^m \leq 0 \quad \forall j \in \boldsymbol{J} \tag{4.35}$$

$$\sum_{r \in \boldsymbol{R}} \gamma_r^b z_{rj}^b - \sum_{i \in \boldsymbol{I}} \omega_i^b x_{ij}^b - \sum_{k \in \boldsymbol{K}} \gamma_k^m z_{kj}^m + u_j^b \leq 0 \quad \forall j \in \boldsymbol{J} \tag{4.36}$$

$$\sum_{t \in \boldsymbol{T}} \mu_t^p y_{tj}^p - \sum_{s \in \boldsymbol{S}} \omega_s^p x_{sj}^p - \sum_{r \in \boldsymbol{R}} \gamma_r^b z_{rj}^b + u_j^p \leq 0 \quad \forall j \in \boldsymbol{J} \tag{4.37}$$

$$\sum_{f \in \boldsymbol{F}} \omega_f^m x_{ij}^m \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.38}$$

$$\sum_{i \in \boldsymbol{I}} \omega_i^b x_{ij}^b + \sum_{k \in \boldsymbol{K}} \gamma_k^m z_{kj}^m \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.39}$$

$$\sum_{s \in \boldsymbol{S}} \omega_s^p x_{sj}^p + \sum_{r \in \boldsymbol{R}} \gamma_r^b z_{rj}^b \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.40}$$

$$\omega_{hj}^e \geq 0 \qquad \forall h \in \boldsymbol{H}, \quad \forall j \in \boldsymbol{J} \tag{4.41}$$

$$\omega_f^m, \mu_g^m, \gamma_k^m, \omega_i^b, \gamma_r^b, \omega_s^p, \mu_t^p \geq \epsilon \tag{4.42}$$

$$u_j^m, u_j^b, u_j^p \quad \text{are free in sign} \tag{4.43}$$

4.4 Export Coal Mine Supply (ECMS) model

The Export Coal Mine Supply (ECMS) model is a special case of the CSLE model. It deals with surface coal mines that produce coal, wash it to improve its quality by removing ash and other organic matter that lowers the energy content of the coal, and then sells clean coal tonnage to gain revenue (US$M) at the port from export markets only. The ECMS model differs from the CSLE model in one respect: no coal is supplied to the local market.

The overall technical efficiency for ECMS is determined from the linear convex combination of the technical efficiency of its subsystems, indicated in Figure 4.2, which is part of the generic mining supply system shown in Figure 4.1. The resulting FP is represented by Equations (4.44)–(4.53).

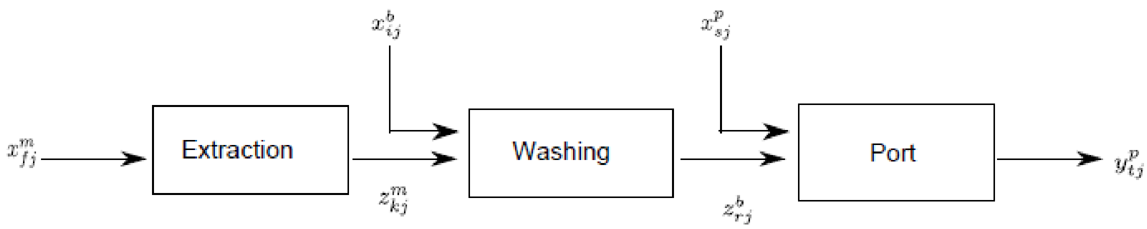

Figure 4.2: Coal mining supply system for export market.

$$\max \theta_o = \frac{\sum_{k \in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum_{r \in \boldsymbol{R}} \eta_r^b z_{ro}^b + \sum_{t \in \boldsymbol{T}} \nu_t^p y_{to}^p + \pi_o^m + \pi_o^b + \pi_o^p}{\sum_{f \in \boldsymbol{F}} \upsilon_f^m x_{fo}^m + \sum_{i \in \boldsymbol{I}} \upsilon_i^b x_{io}^b + \sum_{s \in \boldsymbol{S}} \upsilon_s^p x_{so}^p + \sum_{k \in \boldsymbol{K}} \eta_k^m z_{ko}^m + \sum_{r \in \boldsymbol{R}} \eta_r^b z_{ro}^b} \tag{4.44}$$

subject to

$$\frac{\sum_{k\in\boldsymbol{K}}\eta_k^m z_{kj}^m + \sum_{r\in\boldsymbol{R}}\eta_r^b z_{rj}^b + \sum_{t\in\boldsymbol{T}}\nu_t^p y_{tj}^p + \pi_j^m + \pi_j^b + \pi_j^p}{\sum_{f\in\boldsymbol{F}}\upsilon_f^m x_{fj}^m + \sum_{i\in\boldsymbol{I}}\upsilon_i^b x_{ij}^b + \sum_{s\in\boldsymbol{S}}\upsilon_s^p x_{sj}^p + \sum_{k\in\boldsymbol{K}}\eta_k^m z_{kj}^m + \sum_{r\in\boldsymbol{R}}\eta_r^b z_{rj}^b} \tag{4.45}$$

$$\frac{\sum_{k\in\boldsymbol{K}}\eta_k^m z_{kj}^m + \pi_j^m}{\sum_{f\in\boldsymbol{F}}\upsilon_f^m x_{fj}^m} \leq 1 \qquad \forall j \in \boldsymbol{J} \tag{4.46}$$

$$\frac{\sum_{r\in\boldsymbol{R}}\eta_r^b z_{rj}^b + \pi_j^b}{\sum_{i\in\boldsymbol{I}}\upsilon_i^b x_{ij}^b + \sum_{k\in\boldsymbol{K}}\eta_k^m z_{kj}^m} \leq 1 \qquad \forall j \in \boldsymbol{J} \tag{4.47}$$

$$\frac{\sum_{t\in\boldsymbol{T}}\nu_t^p y_{tj}^p + \pi_j^p}{\sum_{s\in\boldsymbol{S}}\upsilon_s^p x_{sj}^p + \sum_{r\in\boldsymbol{R}}\eta_r^b z_{rj}^b} \leq 1 \qquad \forall j \in \boldsymbol{J} \tag{4.48}$$

$$\sum_{f\in\boldsymbol{F}}\upsilon_f^m x_{fj}^m \geq \alpha \qquad \forall j \in \boldsymbol{J} \tag{4.49}$$

$$\sum_{i\in\boldsymbol{I}}\upsilon_i^b x_{ij}^b + \sum_{k\in\boldsymbol{K}}\eta_k^m z_{kj}^m \geq \alpha \qquad \forall j \in \boldsymbol{J} \tag{4.50}$$

$$\sum_{s\in\boldsymbol{S}}\upsilon_s^p x_{sj}^p + \sum_{r\in\boldsymbol{R}}\eta_r^b z_{rj}^b \geq \alpha \qquad \forall j \in \boldsymbol{J} \tag{4.51}$$

$$\upsilon_f^m, \eta_k^m, \upsilon_i^b, \eta_r^b, \upsilon_s^p, \nu_t^p, \geq \epsilon \tag{4.52}$$

$$\pi_j^m, \pi_j^b, \pi_j^p \quad \text{are free in sign} \tag{4.53}$$

FP was transformed to LP using the same form as that in the Charnes and Cooper transformation using the same approach presented in subsection 4.2. The resulting LP model, after considering the effects of non-discretionary variables, is given in Equations (4.54)–(4.65).

$$\max \theta_o = \sum_{k\in\boldsymbol{K}}\gamma_k^m z_{ko}^m + \sum_{r\in\boldsymbol{R}}\gamma_r^b z_{ro}^b + \sum_{t\in\boldsymbol{T}}\mu_t^p y_{to}^p - \sum_{h\in\boldsymbol{H}}\omega_h^e x_{ho}^e + u_o^m + u_o^b + u_o^p \tag{4.54}$$

subject to

$$\sum_{t\in\boldsymbol{T}}\mu_t^p y_{tj}^p - \sum_{f\in\boldsymbol{F}}\omega_f^m x_{fj}^m - \sum_{i\in\boldsymbol{I}}\omega_i^b x_{ij}^b - \sum_{s\in\boldsymbol{S}}\omega_s^p x_{sj}^p - \sum_{h\in\boldsymbol{H}}\omega_h^e x_{hj}^e + u_j^m + u_j^b + u_j^p \leq 0 \tag{4.55}$$

$$\sum_{f\in\boldsymbol{F}}\omega_f^m x_{fo}^m + \sum_{i\in\boldsymbol{I}}\omega_i^b x_{io}^b + \sum_{s\in\boldsymbol{S}}\omega_s^p x_{so}^p + \sum_{k\in\boldsymbol{K}}\gamma_k^m z_{ko}^m + \sum_{r\in\boldsymbol{R}}\gamma_r^b z_{ro}^b = 1 \tag{4.56}$$

$$\sum_{k\in\boldsymbol{K}}\gamma_k^m z_{kj}^m - \sum_{f\in\boldsymbol{F}}\omega_f^m x_{fj}^m + u_j^m \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{4.57}$$

$$\sum_{r\in\boldsymbol{R}} \gamma_r^b z_{rj}^b - \sum_{i\in\boldsymbol{I}} \omega_i^b x_{ij}^b - \sum_{k\in\boldsymbol{K}} \gamma_k^m z_{kj}^m + u_j^b \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{4.58}$$

$$\sum_{t\in\boldsymbol{T}} \mu_t^p y_{tj}^p - \sum_{s\in\boldsymbol{S}} \omega_s^p x_{sj}^p - \sum_{r\in\boldsymbol{R}} \gamma_r^b z_{rj}^b + u_j^p \leq 0 \qquad \forall j \in \boldsymbol{J} \tag{4.59}$$

$$\sum_{f\in\boldsymbol{F}} \omega_f^m x_{fj}^m \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.60}$$

$$\sum_{i\in\boldsymbol{I}} \omega_i^b x_{ij}^b + \sum_{k\in\boldsymbol{K}} \gamma_k^m z_{kj}^m \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.61}$$

$$\sum_{s\in\boldsymbol{S}} \omega_s^p x_{sj}^p + \sum_{r\in\boldsymbol{R}} \gamma_r^b z_{rj}^b \geq \beta \qquad \forall j \in \boldsymbol{J} \tag{4.62}$$

$$\omega_{hj}^e \geq 0 \qquad \forall h \in \boldsymbol{H}, \quad \forall j \in \boldsymbol{J} \tag{4.63}$$

$$\omega_f^m, \gamma_k^m, \omega_i^b, \gamma_r^m, \omega_s^m, \mu_t^p, \geq \epsilon \tag{4.64}$$

$$u_j^m, u_j^b, u_j^p \quad \text{are free in sign} \tag{4.65}$$

4.5 Local Coal Mine Supply (LCMS) model

The LCMS model is a second special case of the CSLE model. It involves extraction and the supply of thermal coal to local customers such as power plants that buy coal from the mine for electricity generation. The output considered in this market is the revenue (US$M) generated from the quantity of thermal coal sold at the mine mouth.

The model is formulated considering Figure 4.3, which is a subsystem of the CSLE model indicated in Figure 4.1. In this model, no outputs are generated for use as inputs to another subsystem, which means that the output tonnage is directly consumed by local customers. The efficiency of the mine supplying coal to this market is formulated from the definition given by Equation (4.1).

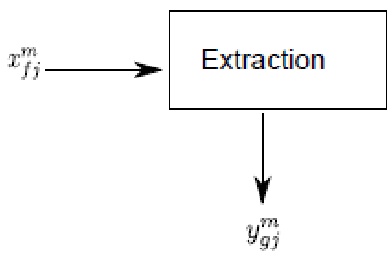

Figure 4.3: Coal mining supply system for local market.

Each mining supply system $j \in \boldsymbol{J}$ only supplies coal to the local market to generate revenue (US$M). Consider that the mine uses different inputs $f \in \boldsymbol{F}$ to produce different outputs $g \in \boldsymbol{G}$. Then, the relative technical efficiency of a surface coal mine denoted by the

DMU_o under evaluation is represented by Equations (4.66)–(4.69).

$$\max \theta_o = \frac{\sum_{g \in \boldsymbol{G}} \nu_g^m y_{go}^m + \pi_o^m}{\sum_{f \in \boldsymbol{F}} \upsilon_f^m x_{fo}^m} \tag{4.66}$$

subject to

$$\frac{\sum_{g \in \boldsymbol{G}} \nu_g^m y_{gj}^m + \pi_j^m}{\sum_{f \in \boldsymbol{F}} \upsilon_f^m x_{fj}^m} \leq 1 \qquad \forall j \in \boldsymbol{J} \tag{4.67}$$

$$\nu_g^m, \quad \upsilon_f^m \geq 0 \tag{4.68}$$

$$\pi_j^m \quad \text{is free in sign} \qquad \forall j \in \boldsymbol{J} \tag{4.69}$$

Transforming from FP to LP according to the Charnes and Cooper transformation gives Equations (4.70)–(4.74), which consider the influence of non-discretionary variables. The transformation used is $t = \frac{1}{\sum_{f \in \boldsymbol{F}} \upsilon_f^m x_{fo}^m}$, which gives new weights $\omega = t\upsilon$ and $\mu = t\nu$.

$$\max \theta_o = \sum_{g \in \boldsymbol{G}} \mu_g^m y_{go}^m - \sum_{h \in \boldsymbol{H}} \omega_h^e x_{ho}^e + u_o^m \tag{4.70}$$

subject to

$$\sum_{g \in \boldsymbol{G}} \mu_g^m y_{gj}^m - \sum_{h \in \boldsymbol{H}} \omega_i^e x_{hj}^e + u_j^m \leq \epsilon \qquad \forall j \in \boldsymbol{J} \tag{4.71}$$

$$\sum_{f \in \boldsymbol{F}} \omega_f^m x_{fo}^m = 1 \tag{4.72}$$

$$\mu_g^m, \quad \omega_f^m \geq \epsilon \tag{4.73}$$

$$u_j^m \quad \text{is free in sign} \qquad \forall j \in \boldsymbol{J} \tag{4.74}$$

where u^m is used to account for the return to scale in the mining operation system, as elaborated in subsection 4.2. The ϵ value is a small positive number that is used to ensure that a DMU with an efficiency score of one must be efficient (Zhu 2008). The model has to be solved $\|\boldsymbol{J}\|$ times, once for each DMU.

4.6 Models for the improvement to the best practices

This section covers the approach to the projection of the inputs to the optimum value for inefficient DMUs using the CSLE model. It shows the mathematical equations that can be used to determine the difference between the actual amount of inputs used by the inefficient mining supply system and the optimum amount of inputs for each variable of the DMU that could produce the same outputs. The difference between the actual input and the optimum amount that could be used to produce the present outputs is referred to as *excess input* in this study. In addition to the definitions used in the formulation of the CSLE model, the following notations are used to explain the equations:

Δx_{fj}^{m} is the magnitude by which excess inputs at the beginning of the extraction subsystem $f \in \boldsymbol{F}$ have to be reduced in order to improve the overall efficiency of the coal mining system $i \in \boldsymbol{J}$.

Δz_{kj}^{m} is the magnitude by which excess intermediate outputs $k \in \boldsymbol{K}$ have to be reduced in order to improve the overall efficiency of the coal mining system $i \in \boldsymbol{J}$.

Δx_{ij}^{b} is the magnitude by which excess inputs at the beginning of the washing subsystem $i \in \boldsymbol{I}$ have to be reduced in order to improve the overall efficiency of the coal mining system $i \in \boldsymbol{J}$.

Δz_{rj}^{b} is the magnitude by which excess intermediate outputs $r \in \boldsymbol{R}$ have to be reduced in order to improve the overall efficiency of the coal mining system $i \in \boldsymbol{J}$.

Δx_{sj}^{p} is the magnitude by which excess inputs at the beginning of port subsystem $s \in \boldsymbol{S}$ have to be reduced in order to improve the overall efficiency of the coal mining system $i \in \boldsymbol{J}$.

The following equations can be applied to determine the magnitude of excess inputs in the specific subsystem of the coal mine system:

1. Improvement for inputs in the extraction subsystem

$$\Delta x_{fj}^{m} = x_{fj}^{m} - \theta_j^m x_{fj}^{m} \tag{4.75}$$

2. Improvement for inputs in the washing subsystem

$$\Delta x_{ij}^{b} = x_{ij}^{b} - \theta_j^b x_{ij}^{b} \tag{4.76}$$

3. Improvement for inputs in the port subsystem

$$\Delta x_{sj}^{p} = x_{sj}^{p} - \theta_j^p x_{sj}^{p} \tag{4.77}$$

4. Improvement for intermediate outputs

$$\Delta z_{kj}^{m} = z_{kj}^{m} - \theta_j^b z_{kj}^{m} \tag{4.78}$$

$$\Delta z_{rj}^{b} = z_{rj}^{b} - \theta_j^p z_{rj}^{b} \tag{4.79}$$

For all efficient DMUs, the excess inputs Δx_{fj}^{m}, Δz_{kj}^{m}, Δx_{ij}^{b}, Δz_{rj}^{b}, and Δx_{sj}^{p} are all zero. If any of them is greater than zero, then the overall mining supply system becomes inefficient. Equations (4.75)–(4.77) are used to determine the magnitude of excess usage of the inputs. An illustration of the computation of the excess inputs for the inefficient DMUs is discussed in Chapter 5.

4.7 Limitation of the formulated DEA models

The CSLE model and the special case models developed in this chapter have some limitations. First, the models assume the ratio of linear sum of weighted outputs to the sum weighted inputs. The expression of total outputs indicates the value function of the outputs, the same applies to the inputs. This assumption does not account for the nonlinear of the marginal values for some of outputs or inputs when they exist. For example if there is diminishing of value function for the increase in output for some variables indicate the presence of nonlinearity. In case of the nonlinearity value functions the models require modifications in order to improve their performance. To consider the nonlinearity among the partial value function of either total weighted output or inputs, Despotis et al. (2010) propose and discus the use of the DEA models that accounts for existence of nonlinearity of partial value functions of the total weighted inputs or outputs which can be adapted and used in the models developed in this chapter.

Second, the models do not account of the correlation among the inputs or outputs in determining the technical efficiency for the given set of DMUs. These can influence the results that will be obtained upon evaluation.

Third, the mathematical models representing the improvement for each subsystems in Equation (4.79)–(4.75) support the process of decision making. For example the models show the result of a zero value of improvement for all subsystems when the DMU is efficient implying nor further improvement is required. But if the DMU is inefficient, the results for the improvement suggested by the models for each subsystem in a linear function form can not be directly used. The management is still required to focus on those variables that it can control and then re-evaluate the efficiency. This stage requires iterative process beginning with the suggested improvement by these models and then varies either some of inputs or/and outputs to achieve the improved efficiency of the inefficient DMU.

4.8 Conclusion

Models for measuring the technical efficiency of surface coal mining supply systems were formulated in this chapter. The models were formulated using the DEA method considering the structure of the mining supply system. The structure of the mining supply system used consisted of extraction, washing, and port subsystems. Each subsystem contributed to the overall efficiency of the mining supply system. One main model for computing the efficiency scores of mining supply systems selling coal to both local and export markets, CSLE, was created and two special-case models for the export market only and the local market only were presented.

The CSLE model and the special-case models can help a mining supply system do a comparative analysis of its competitiveness among the existing producers. This will be achieved by using the efficiency scores calculated from the models. When the mining supply system realises that it is inefficient, the controllable variables are managed to improve efficiency. The models can be applied in the evaluating surface coal mines that have same scope; they use similar inputs and outputs which can vary in quantity. For example one can use CSLE to evaluate a group surface coal mines extracting coal, washing and selling to the market considering that the variables are homogeneous across the mines. To illustrate this for examples all mines must have equal set of variables such as CAPEX and labour as inputs and tonnages of coal as outputs.

It should be noted that it is difficult to model each detail of the system, and thus the focus of this chapter was on the major components of the mining supply system: extraction, washing, and port subsystems. The data simulated in Chapter 3 will be used to illustrate the application and evaluation of the performance of the models in Chapter 5.

Chapter 5

Evaluation of the models for the technical efficiency of a mine

In this chapter, the application of the Combined System for Local and Export (CSLE) model and the special-case models developed in Chapter 4 for measuring the technical efficiency of mining supply systems will be evaluated using the data that were simulated in Chapter 3. The data will be used in the models to compute the relative efficiency scores for each mining supply system and its subsystems, including extraction, washing, and port. The resulting efficiency scores will then be used to identify the best practices, inefficient mining supply systems, and their corresponding subsystems.

The models will thus show how each mining supply system transforms the inputs into outputs relative to similar mining supply systems of thermal coal. The efficient mining supply systems will be those attaining an efficiency score of 1 and the inefficient will be those with efficiency scores in the range of $0 \leq \theta < 1$, where θ is the efficiency score. The mining supply system will be deemed to exhibit best practices if all its subsystems have an efficiency score of 1.

Inefficient mining supply systems imply that the company is not competitive, which is characterized by overuse of their inputs. For example, the use of higher capital than optimal to deliver the same amount of outputs causes the mining supply system to be inefficient. The mining management should determine which part of the system and which input variables need to be optimised to ensure competitiveness and realise the target outputs, such as revenue from the sale of thermal coal.

The focus of this chapter is twofold. First, it discusses the application of the models, including the assumptions that were made, it lists the input and output variables, and it provides the procedure for computing the efficiency scores using a computer program coded for implementing the models. Secondly, it presents and interprets the results of the evaluation of each model.

5.1 Application of the technical efficiency models

The application of the CSLE and special-case models is illustrated using the data provided in Appendix B.1. The CSLE mining supply systems were solved using the computer program coded by the candidate using R open source software (R Core Team 2015) and the input from work that was introduced by Pessanha et al. (2013). The computer program for implementing the CSLE model is provided in Appendix C. In order to solve the special-case models, the computer code was modified to conform with the mathematical model representing the Export Coal Mine Supply (ECMS) and Local Coal Mine Supply (LCMS) models.

The assumptions that were made in evaluating the models for the technical efficiency of mining supply systems included the following:

- The mines were considered to operate in regions where they are subjected to the same conditions (conditions such as economics, for example, the price of coal for a specific quality, and legislation).
- The supply of thermal coal to power plants is done at the mine site. This is known as *mine mouth trading*.
- The allowable maximum carbon emission was assumed to be 25000 tonnes of CO_2 equivalent per annum (CPA Australia 2011).
- The excess carbon emission charges is US$23/tonne (CPA Australia 2011).
- The port capacity is not a limit to mines producing and supplying thermal coal to the export market. The capacity can always be expanded to accommodate increasing numbers of suppliers of coal for the export market.

The variables used to illustrate the application of the models consisted of both discretionary and non-discretionary variables. These included the inputs and outputs at each subsystem of the mining supply system, which are shown in Figure 5.1.

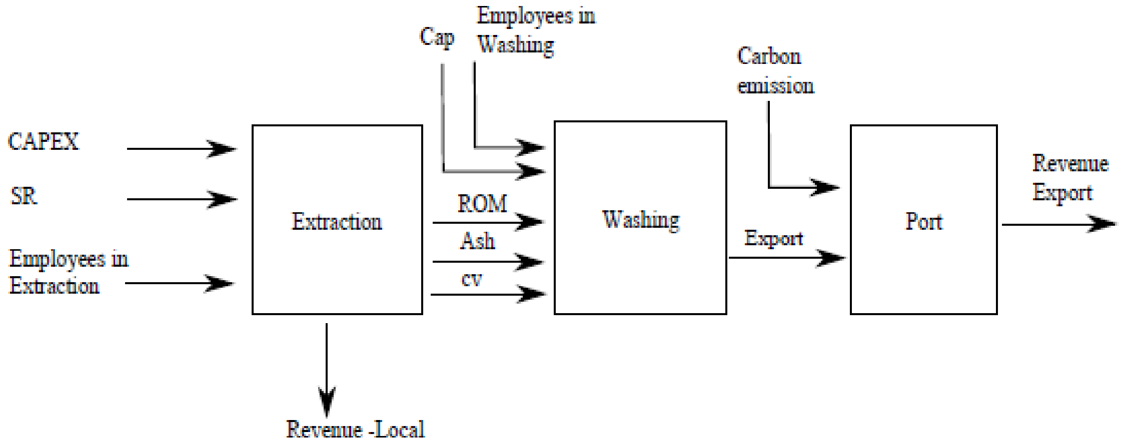

Figure 5.1: Inputs and outputs for computations of efficiency of a mining supply system.

[1] The optimum parameters used in solving the models are $\beta = 0.005$ for CSLE and $\beta = 0.05$ for ECMS. These parameters represent the minimum weighted inputs for each subsystem in order for it to operate. The minimum weight value of each input variable in the models was considered to be $\epsilon = 1 \times 10^{-6}$ (Ramanathan 2003). The procedure for solving the models is as follows:

1. Prepare the data and import into R software.

2. Load the packages supporting the execution of the computer code. These packages include *lpsolve* Berkelaar et al. (2015) for solving linear programming problems and *ggplot2* Wickham (2009), *gridExtra* Auguie (2015), and *grid* Murrell (2002) for generating plots, *reshape2* Wickham (2007) for transforming data between wide and long formats of data frames, *stargazer* Hlavac (2015) for exporting the results, and others, as indicated in the computer code developed.

3. Execute the computer programs to solve the models and extract the results of efficiency scores and graphs for interpretation.

5.2 Results and interpretations

The results obtained after solving the CSLE and special-case models are presented in graphical form. The graphs display the efficiency scores of the mining supply systems and their

[1] Carbon emission indicated in Figure 5.1 represent the excess emission above the allowable carbon emission which is considered to be a penalty charged for export coal, since input oriented Data Envelopment Analysis (DEA) aims at minimizing the inputs while maintaining the present revenue. It also represent input technology for coal production

subsystems. The efficiency scores are plotted against each mining supply system, referred to as a Decision Making Unit (DMU). In addition, comparative analysis of the efficiency scores for the simulated data and the combined set of data of both simulated and original samples for the CSLE model are presented. The results for the special-case models are also presented.

5.2.1 Results of efficiency scores for Combined System for Local and Export (CSLE) model

The results after solving CSLE are shown in Figure 5.2, which shows the efficiency scores of 58 DMUs ranked from the least efficient score of 0.402 for DMU_{28} and ending with all those that are efficient with a score of 1. The results suggest that out of the 58 simulated mining supply systems evaluated, only 8 were efficient, implying that these DMUs exhibit 'best practices' relative to the others. The best practices form the envelope of efficient coal mining supply systems.

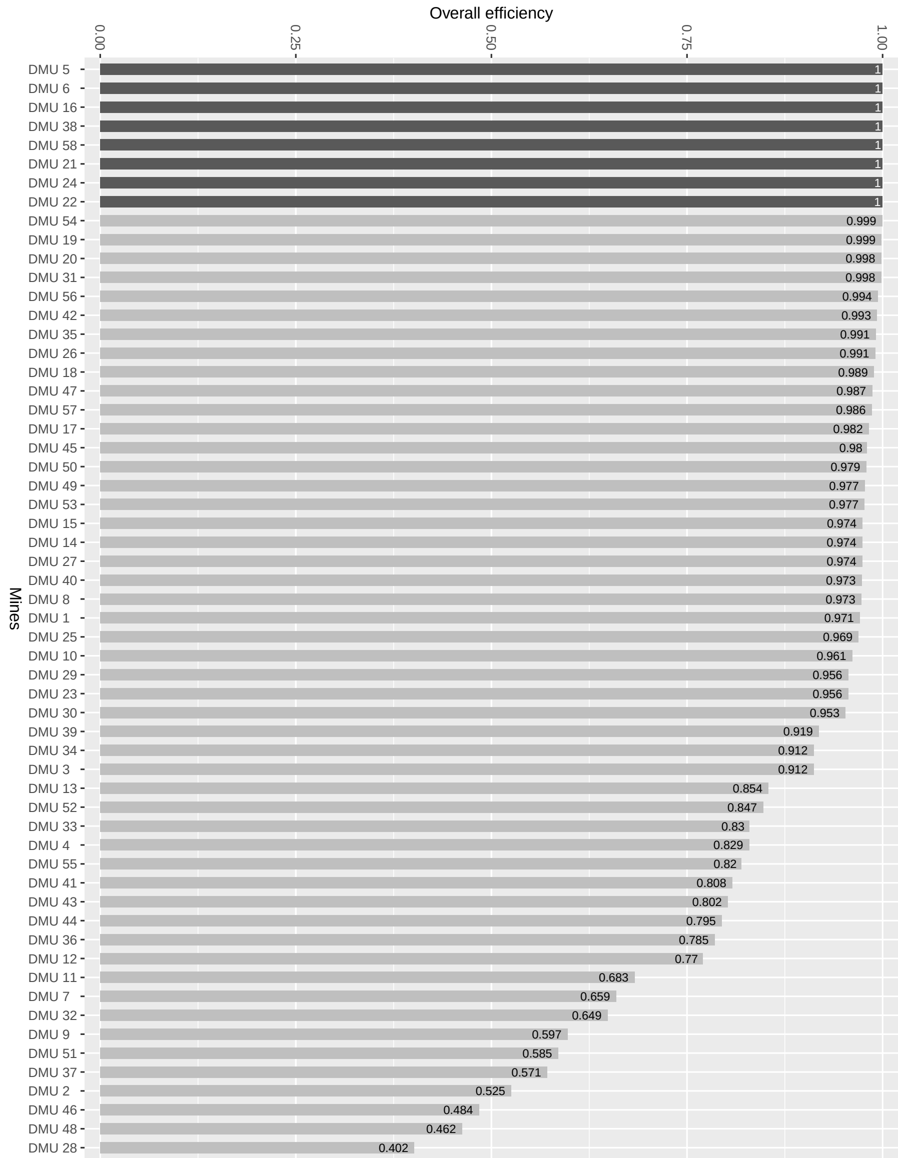

Figure 5.2: Overall efficiency scores for each DMU for the CSLE model.

The efficient coal mining supply systems were DMU_5, DMU_6, DMU_{16}, DMU_{21}, DMU_{22}, DMU_{24}, DMU_{38}, DMU_{54}, and DMU_{58}. The remaining coal mining supply systems were inefficient; however, they can improve their efficiency by reducing their inputs and still achieve their present outputs.

To illustrate the influence of the original samples on the simulated data, the 58 simulated mining supply systems were combined with the original 7 samples extracted from Raw Material Group (RMG) to form 65 DMUs. The efficiency scores of the combined data of DMUs using the CSLE model are presented in Figure 5.3.

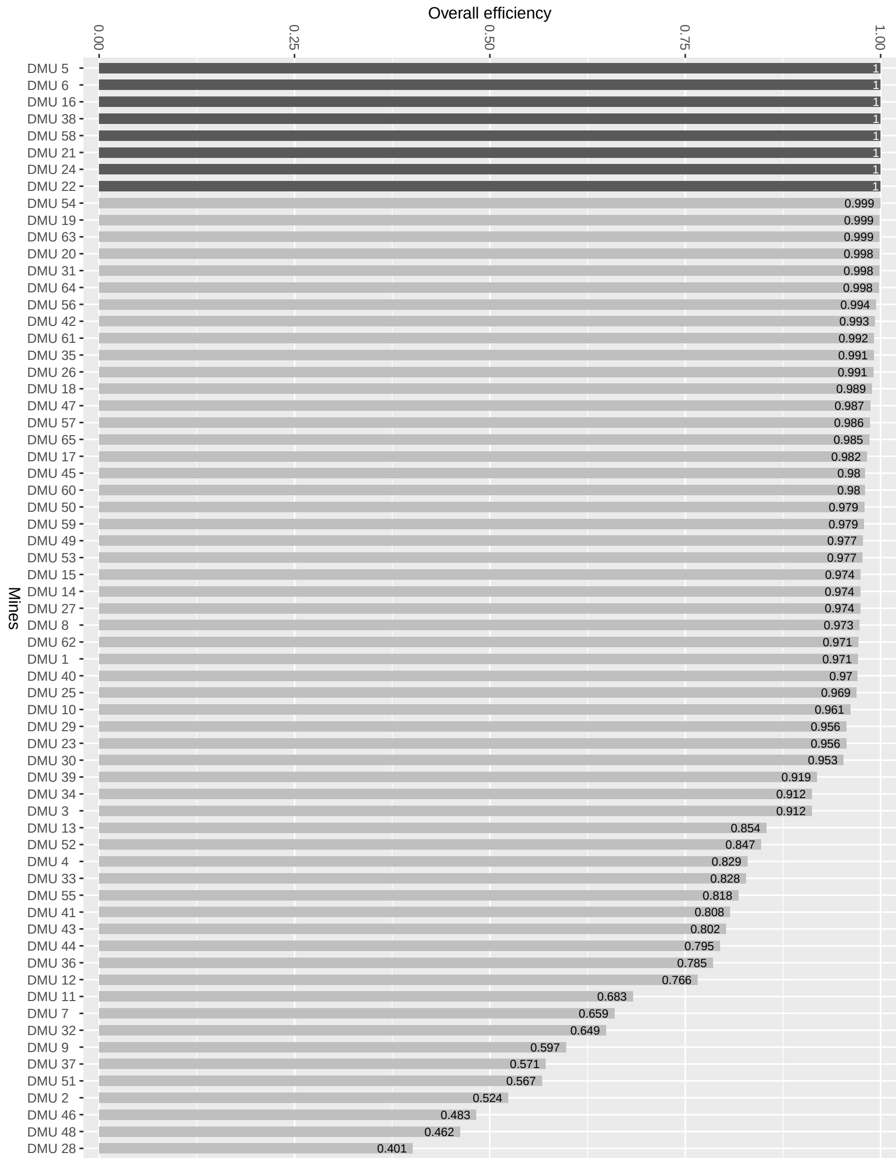

Figure 5.3: Overall efficiency scores for each DMU for the CSLE model.

The results from Figure 5.3 show that the efficient DMUs remained efficient. Some of the efficiency scores of the inefficient DMUs slightly decreased, whereas others increased. For example, the efficient score of DMU_{28} decreased from 0.402 to 0.401.

Considering the results of the simulated 58 mining supply systems shown in Figure 5.2, the efficiency scores of the extraction, washing, and port subsystems of all evaluated mining supply systems are presented in Figures 5.4–5.6 and are discussed below.

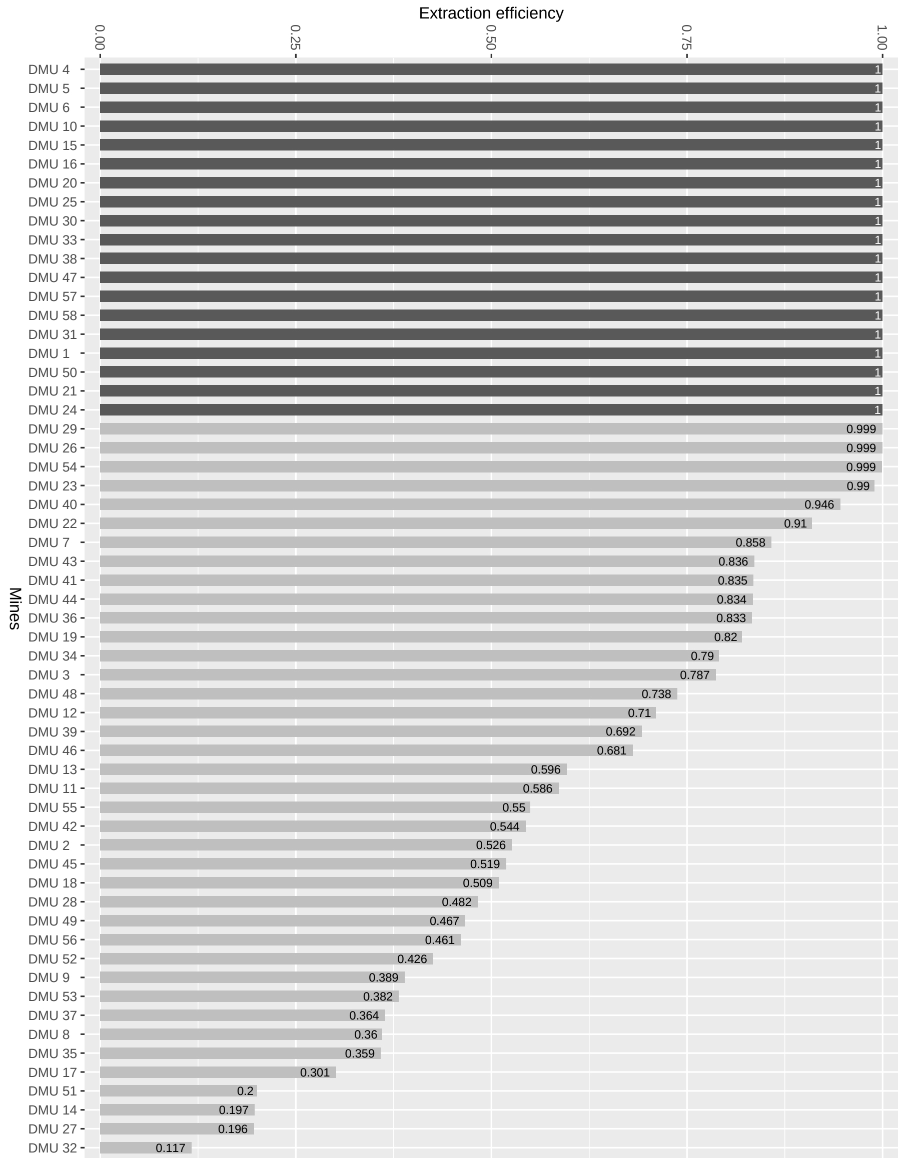

Figure 5.4: Efficiency scores for the extraction subsystem for each DMU of the CSLE model.

Figure 5.4 shows the efficiency scores of the extraction subsystem that were obtained after solving the models to determine the efficiency scores of the overall system of surface coal mines producing and supplying coal for the export market. Of 58 coal mining supply systems, only 19 had efficient extraction subsystems: DMU_1, DMU_4, DMU_5, DMU_6, DMU_{10}, DMU_{15}, DMU_{16}, DMU_{20}, DMU_{21}, DMU_{24}, DMU_{25}, DMU_{30}, DMU_{31}, DMU_{33}, DMU_{38}, DMU_{47}, DMU_{50}, DMU_{57}, and DMU_{58}. The efficient extraction subsystems use minimum Capital Expenditure (CAPEX) (US$M), stripping ratios (SRs), and number of employees in mining to achieve the present outputs, including Run-of-Mine (ROM) (Mt/yr) of a given ash (%) and moisture content (%) and supplying some of the coal tonnage produced to local power plants.

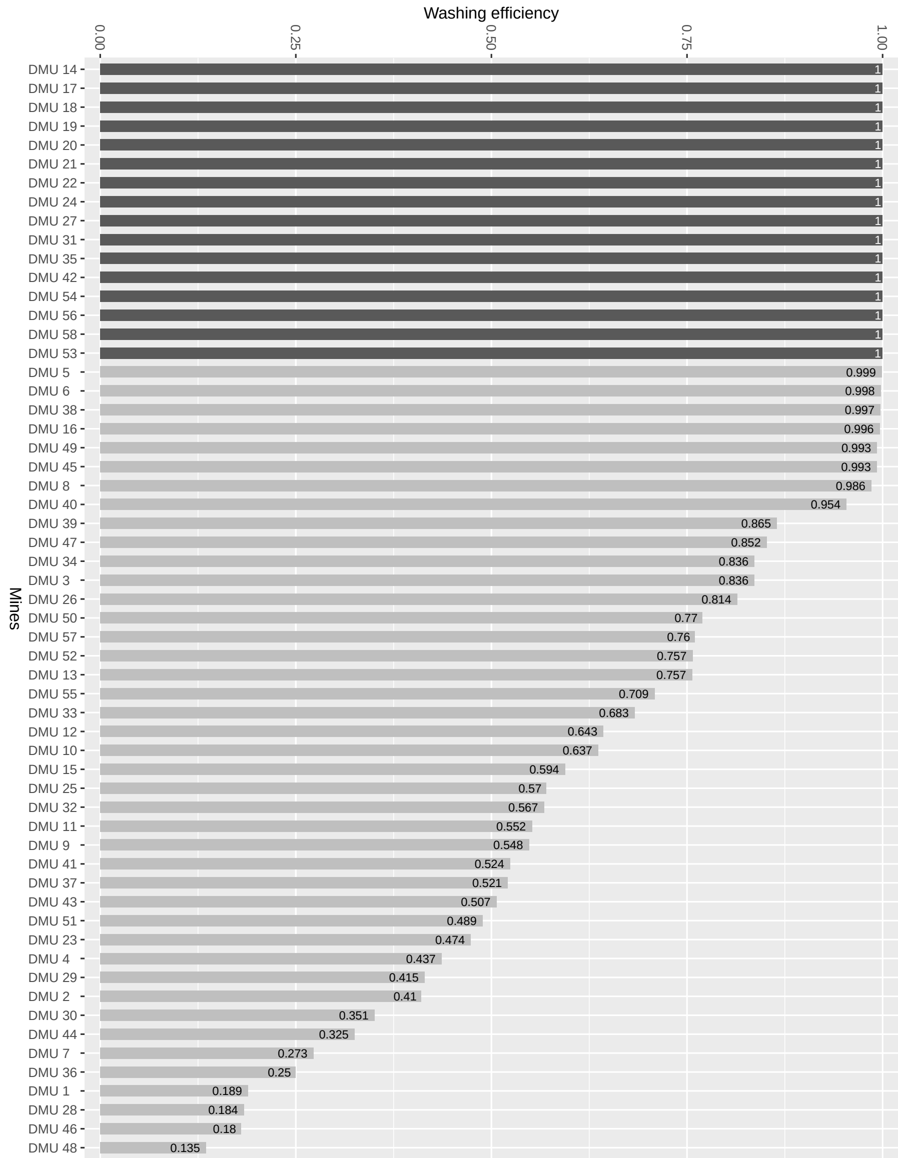

Figure 5.5: Efficiency scores for washing subsystem for each DMU of CSLE model.

Figure 5.5 shows the efficiency scores of the washing plant subsystem of each mining supply system. The results show that only 16 of the 58 coal mining supply systems evaluated had efficient washing subsystems, i.e., an efficiency score of 1. These subsystems use ROM (Mt/yr), ash (%), moisture (%), and capacity of the plant (Mt/yr) as part of the overall coal mine system inputs, and they generate coal tonnage for export. The least efficient is the washing subsystem of DMU_{44}, with an efficiency score of 19.2%.

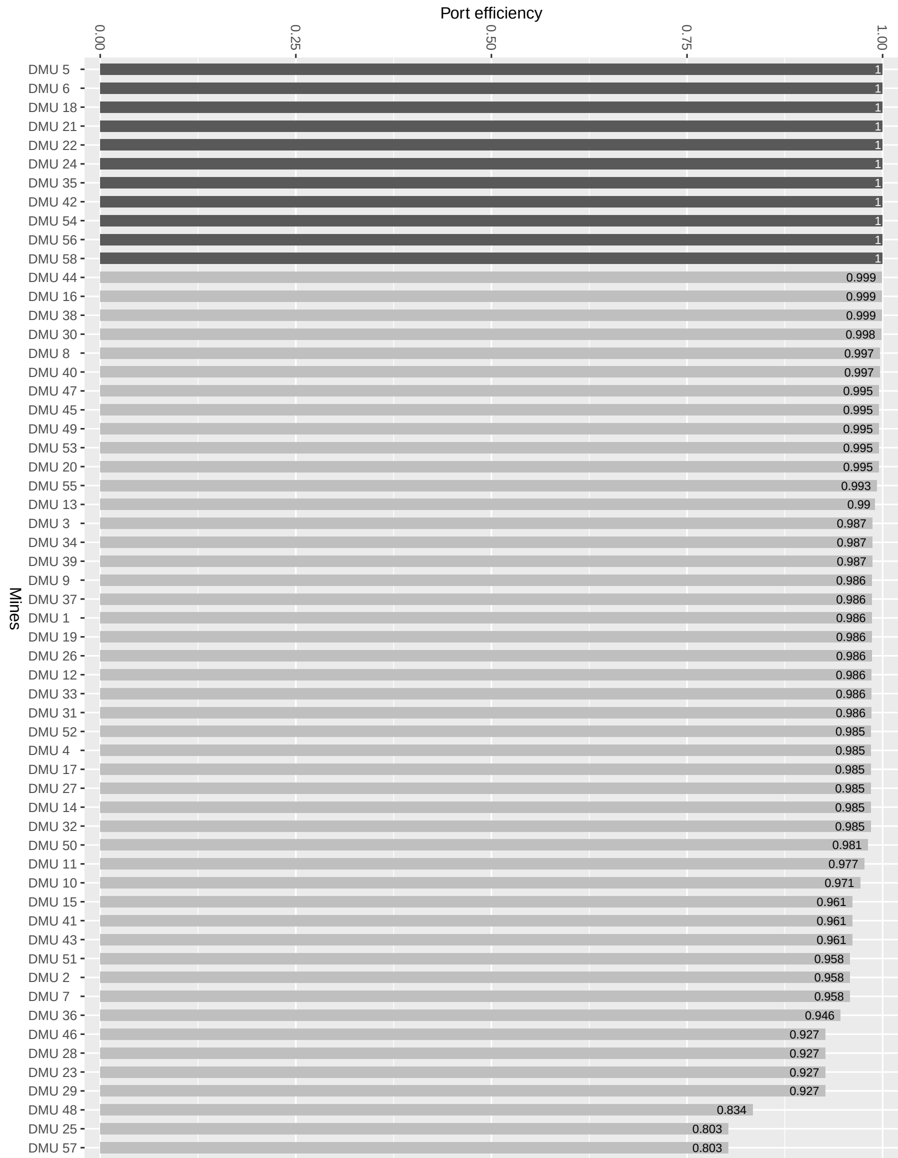

Figure 5.6: Results for port subsystem of each DMU of CSLE.

Figure 5.6 gives the results of the port subsystems of all mining supply systems studied. The results show 11 DMUs with efficiency scores of 1. These are the only efficient port subsystems out of all 58 DMUs. The least efficient port subsystems are those for DMU_{25} and DMU_{57}, each of which has an efficiency score of 0.803. The efficient DMUs supply optimal export coal tonnage at an allowable carbon emission limit to generate revenue from the sale of coal to the export market.

Furthermore, a comparative analysis was done between the efficiency scores of CSLE using discretionary inputs only and the model using both discretionary and non-discretionary inputs. It was found that the non-discretionary variables had a very slight effect on efficiency scores. The reason is suggested to be a slight variation of non-discretionary variables among the mining systems because they were simulated using data from similar operating conditions, and thus they may have had slight differences in their non-discretionary variables. The maximum deviation of the efficiency scores of the CSLE model with discretionary variables only from that with combined discretionary and non-discretionary variables was found to be $\pm 5 \times 10^{-5}$.

5.2.2 Results of efficiency scores for Export Coal Mine Supply (ECMS) model

The evaluation of the application of the ECMS model was done by solving it using 50 simulated coal mining supply systems (see attached Table B.2) that represent surface coal mines producing and supplying coal to the export market only. The model is a special case of the CSLE model. The efficiency scores of the overall system for each DMU are presented in Figure 5.7. The efficiency scores for each subsystems are shown in Figures 5.8–5.10.

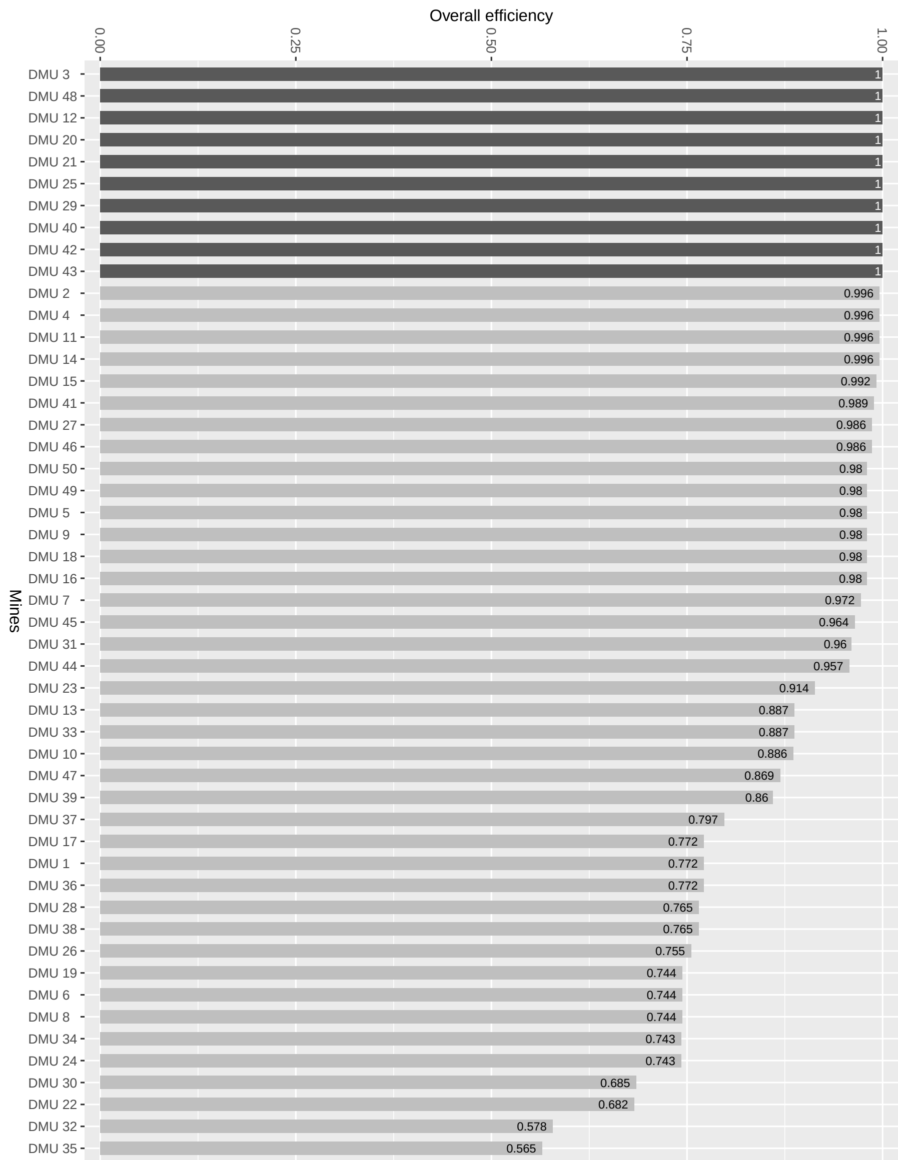

Figure 5.7: Overall efficiency scores of each DMU for the ECMS model.

Figure 5.7 shows the overall efficiency score of each coal mining supply system. The results suggest that the efficient DMUs are DMU_3, DMU_{12}, DMU_{15}, DMU_{20}, DMU_{21}, DMU_{25}, DMU_{29}, DMU_{40}, DMU_{42}, DMU_{43}, and DMU_{48}. These are the ones that efficiently use the optimum quantity of inputs from each of their subsystems.

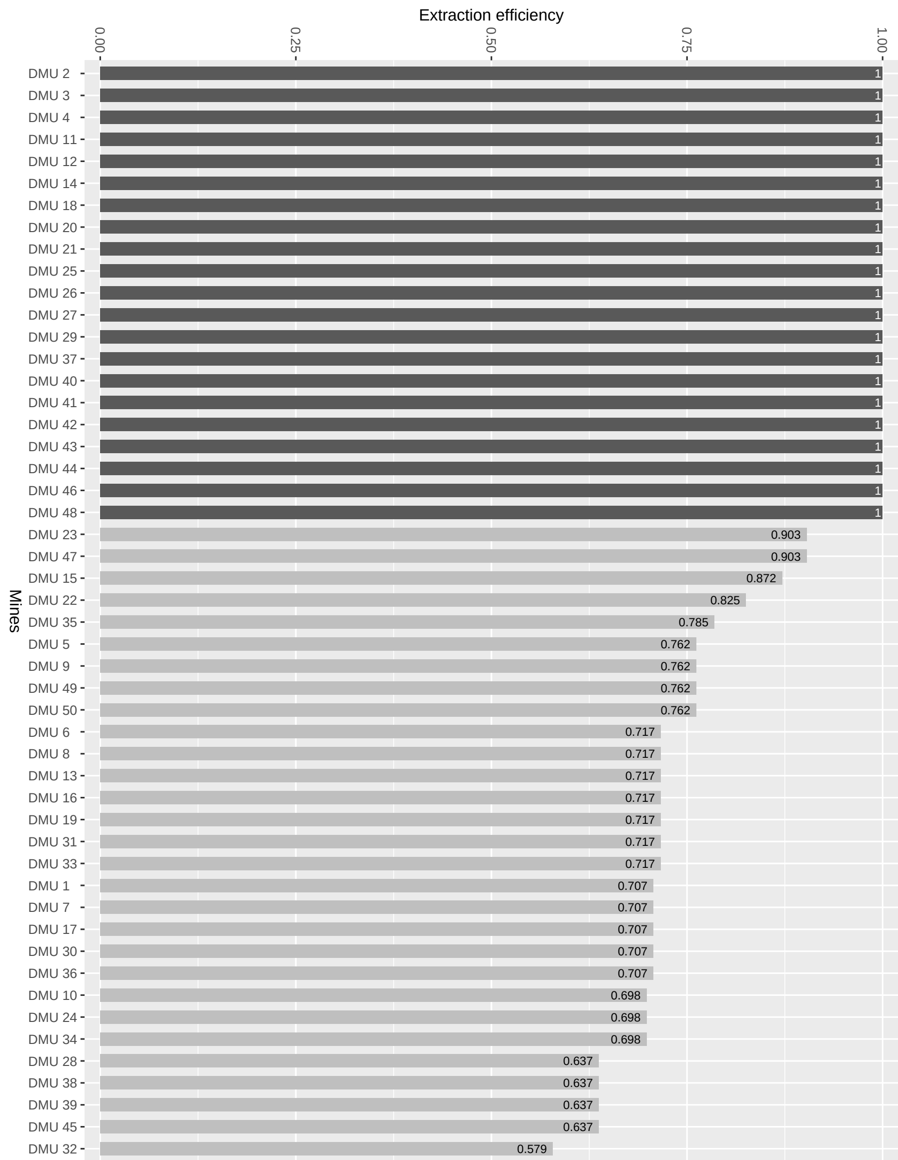

Figure 5.8: Efficiency scores for mining subsystem of each DMU of the ECMS model.

The efficiency scores of the extraction subsystem for the ECMS model are indicated in Figure 5.8. Efficient extraction subsystems were found for DMU_2, DMU_3, DMU_4, DMU_{11}, DMU_{12}, DMU_{14}, DMU_{18}, DMU_{20}, DMU_{21}, DMU_{25}, DMU_{26}, DMU_{27}, DMU_{29}, DMU_{37}, DMU_{40}, DMU_{41}, DMU_{42}, DMU_{43}, DMU_{44}, DMU_{46}, and DMU_{48}. These subsystems use minimum CAPEX, Stripping Ratio (SR), and number of employees to produce ROM (Mt/yr) of a given ash (%) and moisture content (%).

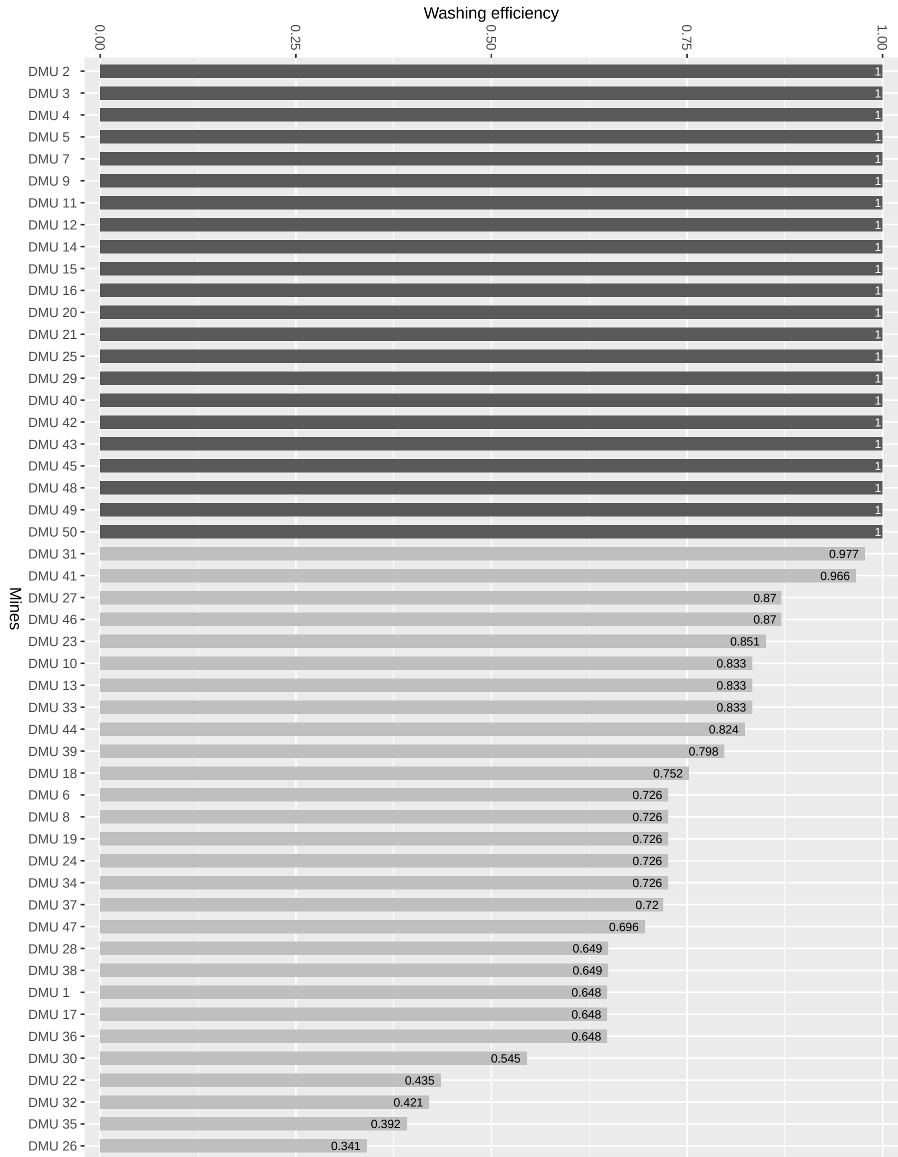

Figure 5.9: Efficiency scores for washing subsystem for each DMU of the ECMS model.

Figure 5.9 shows the efficiency scores for the washing subsystem of the 50 coal mining supply systems evaluated. The efficiency scores are ranked in decreasing order of magnitude. It can be seen that the efficient washing subsystems are those with efficiency scores of 1, such as the washing subsystem of DMU_4. The efficient washing subsystem utilizes a minimum number of inputs as part of the overall inputs of the coal mining supply system to generate the export tonnage. The inputs to the washing subsystem include ROM (Mt/yr) of a given moisture (%) and ash content(%) and the capacity of the washing plant (Mt/yr).

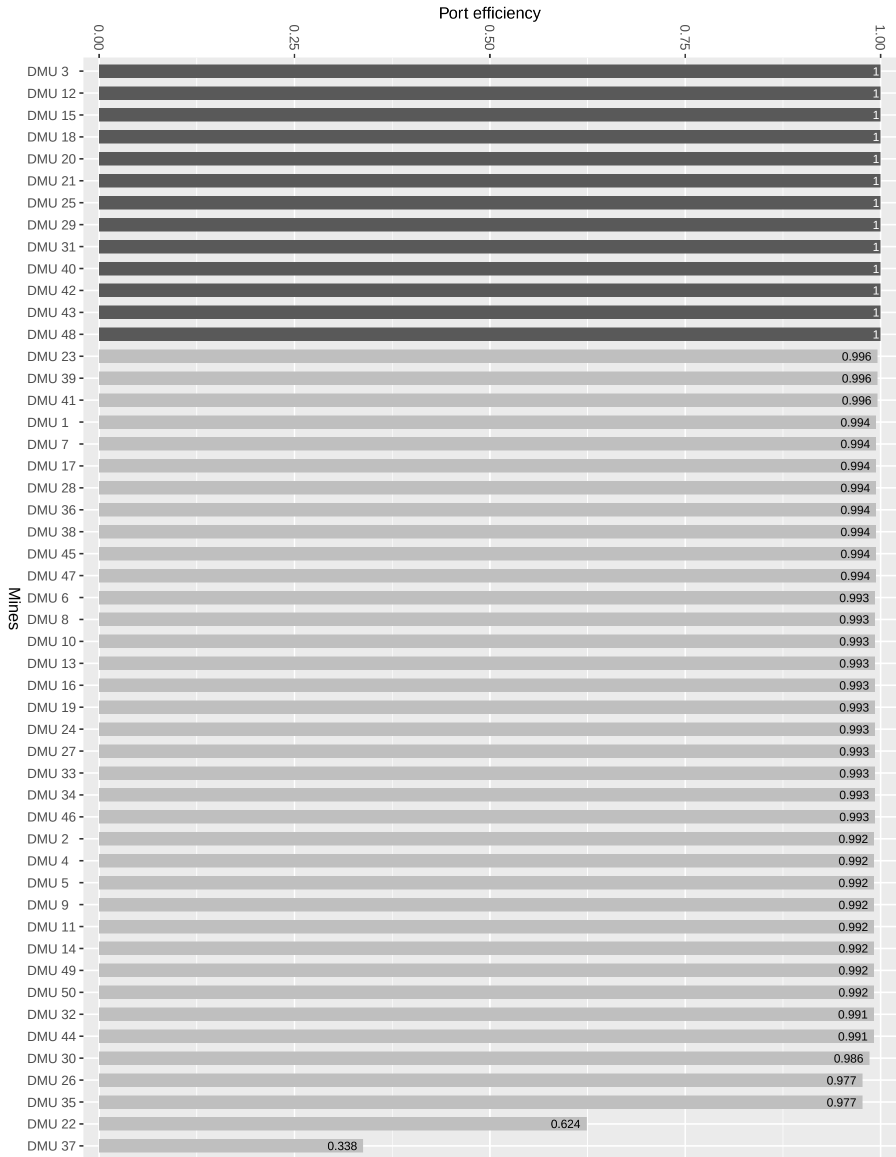

Figure 5.10: Efficiency scores for port subsystem for each DMU of the ECMS model.

Figure 5.10 shows the results for the port subsystem of the 50 DMUs. Of the 50 DMUs, 13 have efficient subsystems with a score of 1. Moreover, the vast majority of the port subsystems have higher efficiency scores except DMU_{22} and DMU_{37}, which have low efficiency scores of 0.624 and 0.338, respectively. The efficient port subsystems efficiently export tonnage at a given specific carbon emission limit to generate revenue from the sale of coal to the export market.

5.2.3 Results of efficiency scores for the Local Coal Mine Supply (LCMS) model

The results for the application of the LCMS model are represented in Figure 5.11 through computation of the efficiency scores of 30 (see Table B.3) simulated DMUs. This is the second special case of the CSLE model.

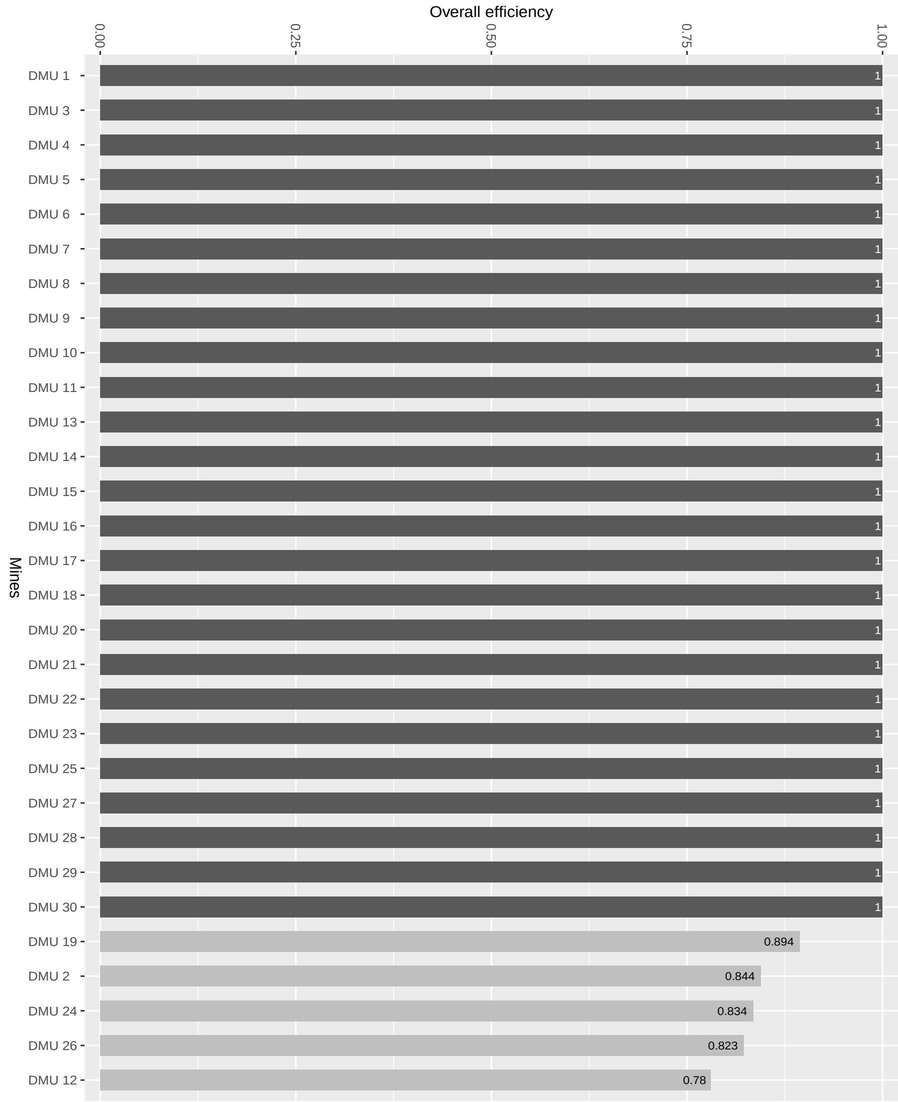

Figure 5.11: Results of DEA efficiency for each DMU ECMS.

Figure 5.11 shows that DMU_2, DMU_{12}, DMU_{19}, DMU_{24}, and DMU_{26} are inefficient surface coal mines. The remaining 25 DMUs are efficient, with efficiency scores of 1. These mines use optimal inputs: CAPEX, stripping ratio, and number of employees to generate revenue from the tonnage of coal sold at the mine mouth to power plants.

5.3 A use case of the models

Decision making in operating or new mining operations require analysis of options. Different scenarios are evaluated, the evidence and effects of the change are then used to support management decisions. For example, the report by Ernst Young (2014b) shows that BHP Billiton for all its operations decided to put on a common management information platform to improve the operations to the best practices and operational performance by generating more volume using the available equipment and lowering costs. It was achieved by improving availability and utilization of equipment. This reduced controllable cash cost of a total of US$2.7b in 2013 financial year across the group.

In this research, the CSLE model and the two special-case models have been developed and illustrations of the applications have been discussed up to this stage. But one would ask the question 'so what?' This can be answered through a use case of the models to support decision making for the management of mining companies in their operations and in formulating quantitative benchmarks.

Consider for example DMU_{12}, which appears inefficient among the simulated DMUs that were evaluated using the CSLE model. This mining supply system had an overall efficiency score of 0.770. The efficiency of its extraction subsystem was 0.710, that of the washing subsystem was 0.643, and that of the port subsystem was 0.986. The operating variables of DMU_{12} are indicated in Table 5.1. The results of the efficiency scores from the CSLE model suggest that an overall improvement of 23% is needed to attain the efficiency of the best practices. What could the management of DMU_{12} do to improve the efficiency of this mining supply system?

Table 5.1: Base case discretionary variables of DMU_{12}

Mine Variables	**Amount**
CAPEX (US$M)	631.33
Stripping ratio	7.8
Number of employees in extraction	1024
Number of employees in washing	129
Run-off mine (Mt/yr)	14.8
Ash (%)	15.5
Moisture (%)	13.5
Plant capacity (Mt/yr)	15
Export (Mt/yr)	6.99
Excess carbon emission (Mt/yr)	0.641
Local supply (Mt/yr)	2.0
Revenue from export (US$M)	564.03
Revenue for local supply (US$M)	68.08
Extraction efficiency	0.710
Washing efficiency	0.643
Port efficiency	0.986
Overall efficiency	0.770

Improving the efficiency of DMU_{12} requires to vary the controllable variables at the discretion of the mining management and compute the technical efficiency for each variation. This is a repetitive process to evalute the improvements. In the improvement process, an input or output cannot be increased or decreased without affecting one or more other variables of the mining supply system. The inputs are interdependent variables of the mining supply system, i.e., one output becomes an input into another subsystem.

To help the decision making for DMU_{12}, two options are suggested to the mining management to improve the efficiency. The first option is to reduce the discretionary inputs while maintaining the present amount of one or both outputs. The second option is to increase one or both outputs using the present amount of inputs. The process is iterative, and it requires management to focus on the variable that can be reduced or increased at their discretion. The results for the two options are indicated in Table 5.2.

Table 5.2: Options for improving the efficiency of DMU_{12}

Mine variables	**Base case**	**Option 1**	**Option 2**
CAPEX (US$M)	631.33	631.33	631.33
Stripping ratio	7.8	$\boxed{7.0}$	7.8
Employees in extraction	1024	$\boxed{650}$	1024
Employees in plant	129	$\boxed{100}$	129
Run-off mine (Mt/yr)	14.8	14.8	14.8
Ash (%)	5.5	5.5	5.5
Moisture (%)	13.5	13.5	13.5
Plant capacity (Mt/yr)	15.0	$\boxed{7.2}$	15.0
Export (Mt/yr)	6.99	6.99	$\boxed{10.0}$
Excess carbon emission (Mt/yr)	0.641	0.641	0.641
Local supply	2	2	$\boxed{4}$
Revenue from export (US$M)	564.03	564.03	**896.1**
Revenue for local supply (US$M)	68.08	68.08	**136.2**
Extraction efficiency	0.710	**0.851**	**0.624**
Washing efficiency	0.643	**1.000**	**1.000**
Port efficiency	0.986	0.986	**1.000**
Overall efficiency	0.770	**0.989**	**0.992**

From the results in Table 5.2, the first option suggests retrenchment of the total number of

403 people, generating mine designs that result in an SR of 7 [2], and decreasing the maximum capacity of the plant to 7.2 Mt/yr. These parameters will increase the overall efficiency of DMU_{12} from 0.770 to 0.989 while producing the present revenue of US$564 M.

The second option suggests expanding the mine's supply of coal to the local and export markets. The export should increase from 6.99 Mt/yr to 10 Mt/yr and the local output supply should increase from 2 to 4 Mt/yr. The revenue is expected to increase from US$564 M to US$896.1 M for export and from US$ 68 M to US$136.2 M for local sales. The efficiency will improve from 0.770 to 0.992.

Comparing the two options, the first option involves retrenching a total of 403 people from both the extraction section and in the washing plant while achieving the same revenue. The second option requires the availability of coal demand, which allows for an extra supply of coal; otherwise, it is becomes difficult to implement this option.

In addition, the above two options are not the only ones. One could vary the controllable variables and re-evaluate the efficiency scores using the CSLE model. The results would then be assessed to determine the viability of such options in improving the efficiency of DMU_{12}.

The verification of the models can be done by applying the DEA models and identifying the real DMUs representing the actual mines, and perform a qualitative check in the report of their performance for a given period and comparing their technical efficiencies computed from the model. For example the mines can be inefficient in the times of the events of floods which affects the use of inputs to generate the target coal outputs. This can be compared with the results generated by the model to validate the performance of the model.

5.4 Conclusion

The application of the models for measuring the efficiency of mining supply systems of thermal coal was illustrated in this chapter by using simulated data and computer code developed for the implementation of the models by the candidate.

All models provided the ability to differentiate between the efficient and inefficient mining supply systems. The efficient mining supply systems form the envelope of best practices. The inefficient mining supply systems and the corresponding subsystems can be identified. This gives an informed decision to identify the inputs and subsystems for the improvement of the technical efficiency to conform to best practices. Consequently, mine management can use this approach to identify realistic, quantitative benchmarks to work towards.

[2] This refers to the process of reducing SR which can be through reviewing the assumption for mine optimization and improving the design of the mine such as steepening the slope of the pit to attain the required SR from 7.8 to 7

In evaluating a new mine, the models require data to be collected from existing producing mining supply systems and application of the models to compute the efficiency scores of the new mine, together with those of the existing producers of thermal coal. The resulting efficiency scores of the new mine can help it to position itself competitively. Moreover, the use case based on simulated data described in this chapter illustrates how the models can assist mining management to make decisions about the project.

In case there is insufficient data for computing the efficiency scores or to simulate mining supply systems to include with the new mine, a predictive model is proposed and will be developed in Chapter 6 for use in predicting the efficiency scores of new mining supply systems. Furthermore, a predictive model for the initial selection of the mine production rate will also be formulated.

Chapter 6

Predictive modelling of the efficiency and performance of a mine

Predictive modelling refers to the process of developing a model that allows us to understand and quantify the model prediction accuracy for the future when given new data that have not been seen before (Kuhn and Johnson 2013). The process uses information (data) to specify the parameters of the model. It involves the use of input variables, known as *predictors*, and output variables, referred to as *response variables*. Two stages are involved in predictive modelling. The first stage is formulating the model and the second stage is to use data to estimate the parameters of the model using a supervised learning method known as Ordinary Least Squares (OLS). The parameters of the model are obtained by minimizing the sum of the squares error between the actual and estimated values of the response variables.

In this research, two models are proposed and formulated. One is a predictive model for the technical efficiency of mining supply systems and the second is a predictive model for initial selection of the production rate of coal mines using predetermined input variables. The proposed predictive models will serve the purpose of assessing the technical efficiency of a new mine before commencing operation. The models are developed from the efficiency scores obtained in Chapter 5 of each mining supply system and their corresponding input variables. These models can be used in cases where there is an insufficient number of mining supply systems to include with the new one in the computation of the relative efficiency scores for competitiveness.

6.1 Formulation of the technical efficiency predictive model

To formulate the predictive model for the technical efficiency of new mine projects, the input variables of each subsystem and the non-discretionary variables are used. Consider the data for 65 Decision Making Units (DMUs) (combined original and simulated mines), each of which has an efficiency score that is the maximum result obtained by solving the Combined System for Local and Export (CSLE) model as presented in Section 5.2. The set of efficiency scores is linearly related to the non-discretionary variables of the mining supply systems using the OLS method (Xue et al. 1999). In this research, the discretionary variables were combined with non-discretionary variables for use in OLS. The following assumptions for OLS were applied (Kurkiewicz et al. 2013):

- There is a linear relationship between the response and predictors. The linearity is in the regression parameters.
- There is constant variance.
- The errors are normally distributed.
- The errors are independent and have zero mean.
- The response variables are measured without error.

For modelling purposes, the relationship between the inputs and outputs is represented by Figure 6.1. Consider $\varphi(.)$ to be the function that maps the input variables to the output efficiency scores of a mining supply system $j \in \boldsymbol{J}$. The inputs to the model are the discretionary and non-discretionary variables and the outputs are the efficiency scores of the overall mining supply system.

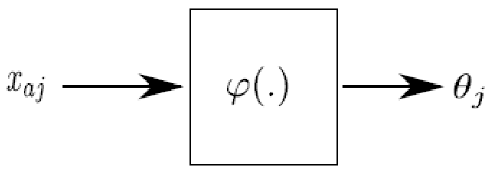

Figure 6.1: Simplified mapping of the inputs and outputs of a DMU.

Six sets of data are used as inputs: extraction, washing, and port subsystems and non-discretionary variables. All are represented by vector $\boldsymbol{x_{aj}}$.

Let $\quad \boldsymbol{x_{aj}} = \left\{ x^m_{fj}, x^b_{ij}, x^p_{sj}, x^e_{hj}, z^m_{kj}, z^b_{rj} \right\}$

where

x^m_{fj} is input variable $f \in \boldsymbol{F}$ in extraction subsystem (m) for mining supply system $j \in \boldsymbol{J}$.

x^b_{ij} is input variable $i \in \boldsymbol{I}$ in washing subsystem (b) for mining supply system $j \in \boldsymbol{J}$.

x^p_{sj} is input variable $s \in \boldsymbol{S}$ in port subsystem (p) for mining supply system $j \in \boldsymbol{J}$.

z^m_{kj} is intermediate input variable $k \in \boldsymbol{K}$ in washing subsystem (b) for mining supply system $j \in \boldsymbol{J}$.

z^b_{rj} is intermediate input variable $r \in \boldsymbol{R}$ in port operation (p) for mining supply system $j \in \boldsymbol{J}$.

x^e_{hj} is non-discretionary input variable $h \in \boldsymbol{H}$ of a mining supply system $j \in \boldsymbol{J}$.

To explain the model, the following notations are used:

θ_j is the efficiency score of mining supply system $j \in \boldsymbol{J}$ obtained from the CSLE model in Chapter 5.

$\hat{\theta}_j$ is the estimate of efficiency score θ_j.

$\hat{\beta}_a$ is the estimate of the model parameter β_a.

The formulation of the model is based on minimizing the overall sum of the squares error between the predicted and observed values of the efficiency scores. Consider ϵ_j to represent the error (residual) between the actual and estimated value of efficiency scores given by Equation (6.1).

$$\epsilon_j = [\theta_j - \varphi(x_{aj}, \beta_a)] \tag{6.1}$$

where φ represents the mean function for estimating the efficiency score θ_j. The form of Linear Program (LP) for determining the coefficients of the predictive model of the technical efficiency is represented by Equations (6.2)–(6.4) (Kong 2007).

$$\min \quad Z_j = \sum_{j \in n} \epsilon_j^2 \tag{6.2}$$

subject to

$$\varphi(x_{aj}, \beta_a) + \epsilon_j = \theta_j \tag{6.3}$$

$$\beta_a \geq 0 \tag{6.4}$$

Equation (6.3) for estimating the efficiency score can be represented by equation (6.5).

$$\theta_j = \beta_o + \sum_{j \in n} \beta_a x_{aj} + \epsilon_j \quad \text{where:} \quad a = \{1, ...d\} \quad \text{and} \quad j = \{1, ...n\} \tag{6.5}$$

The dimension of set $\boldsymbol{x_{aj}}$ is $n \times d$, where n is the number of DMUs and d is the total number of predictor variables. The dimension of β_a is $d \times 1$ and the dimension of the residual ϵ_j is $n \times 1$. The matrix form of the Equation (6.5) is as follows:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_d \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

The above matrix can be solved using available software such as R Core Team (2015) by determining the coefficients of the model that minimize the sum of the square errors between the technical efficiency obtained from the CSLE model and estimates from fitting Equation 6.5.

The efficiency score of each DMU (θ_j) is dependent on other DMUs because each of them was computed considering the inputs and outputs of other DMUs using the CSLE model. These efficiency scores cannot be directly used to specify the model parameters because one of the assumptions of OLS, which states that the errors in prediction are assumed to be independent, is violated. In this case, the errors cannot be independent because the efficiency score (response) variables are dependent on one another. Therefore, using these efficiency scores without creating independence among them will generate incorrect model parameters and the specified model may produce incorrect predictions when applied.

To create independent efficiency scores θ_j to overcome this problem, Xue et al. (1999) discuss the use of the bootstrap method of sampling of the efficiency scores with replacement, resulting in independent samples that are then used in ordinary least squares regression. Therefore, the model that will be used for estimation of the efficiency scores is given by Equation (6.6).

$$\hat{\theta_j} = \hat{\beta}_o + \hat{\beta}_1 x_{1j} + \hat{\beta}_2 x_{2j} +, ..., \hat{\beta}_a x_{aj} \tag{6.6}$$

where $\hat{\beta}_o$, $\hat{\beta}_1$, and $\hat{\beta}_2$ through $\hat{\beta}_a$ are estimates of the model parameters obtained by applying the bootstrap technique. They are also the parameters that will give minimum sum of the square errors of the efficiency score estimates.

6.1.1 Estimation of parameters of technical efficiency model

The method used to estimate the parameters of the model in Equation (6.6) is the bootstrap approach of the stepwise Akaike Information Criterion (AIC) proposed by Austin and Tu (2004). It applies a bootstrap resampling technique for automated variable selection methods. The algorithm for implementing this method is known as bootStepAIC, which was developed as a package in R software by Rizopoulos (2009). The algorithm selects the most significant predictor variables and omits the ones that are not significant by relating them to the technical efficiency scores as response variables.

To specify the parameters of the predictive model for technical efficiency, the original set of mining system were combined with the simulated mines to give total of 65 datasets, which were split into 52 training sets (equivalent to 80%) and 13 test sets (equivalent to 20%). The process was carried out through random sampling of rows consisting of DMUs and their variables without replacement of the data. The resulting training sets are given in Table D.1 and test sets in Table D.2. The training sets obtained were used to specify the parameters of the model using the bootStepAIC package and the test sets were used to evaluate the performance of the model.

After applying the bootStepAIC package starting with all predictor variables in all subsystems of the mining supply systems, the final results are shown in Table 6.1. The results show predictor variables that are significant at a level of 5% ($p < 0.05$), which are used to specify the final model shown by Equation (6.7). The number of mines used for building the models were 52. The coefficient of determination (R^2) suggests that 54.9% of the variability in technical efficiency scores can be explained by the predictors indicated in Table 6.1. The resulting Root Mean Square Error (RMSE) is 0.1177. To explain the final model, consider the following definitions:

A is the ash (%).

R is maximum plant capacity (Mt/yr).

P is the plant employees, this is combined administration and washing employees.

E is export coal tonnage (Mt/yr).

S is the precipitation (mm).

$$\hat{\theta} = 0.7453 + 0.0105 \times A - 0.0209 \times R + 0.0706 \times E - 0.0032 \times P + 0.0001 \times S \qquad (6.7)$$

6.1.2 Evaluation of the technical predictive model

The evaluation of the predictive model for the technical efficiency of mining supply system for local sale and export of thermal coal represented by Equation 6.7 was carried out using the test dataset in Table D.2. The efficiency score of each DMU was re-estimated using the predictive model. The resulting technical efficiency scores estimated were compared to the efficiency scores of the test datasets obtained by solving the CSLE model. The input variables, estimated efficiency scores, and computed efficiency scores from the CSLE model

Table 6.1: Regression Results

	Dependent variable:
	Efficiency
Ash	0.0105^{***}
	(0.0029)
Plant capacity	-0.0209^{***}
	(0.0068)
Export	0.0706^{***}
	(0.0130)
Plant employees	-0.0032^{***}
	(0.0010)
Precipitation	0.0001^{***}
	(0.00002)
Constant	0.7453^{***}
	(0.0641)
Observations	52
R^2	0.5488
Adjusted R^2	0.4997
Residual Std. Error	0.1177 (df = 46)
F Statistic	11.1900^{***} (df = 5; 46)
Note:	*p<0.1; **p<0.05; ***p<0.01

for each DMU are shown in Table 6.2. The predicted efficiency scores were obtained by substituting the input variables given in Table 6.2 into Equation (6.7).

Table 6.2: Comparison of CSLE and predicted efficiency scores for test datasets of given input variables

Mines	Ash %	Plant capacity (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	Number of plant employees	Efficiency scores	Predicted efficiency scores
DMU 3	13.3	7.8	4.80	683	49	0.9120	0.9555
DMU 15	10.1	4.0	2.30	2121	85	0.9745	0.8167
DMU 19	13.3	1.2	0.58	683	15	0.9988	0.9042
DMU 24	13.3	2.4	0.30	630	10	0.9996	0.8715
DMU 25	10.1	4.0	0.58	688	20	0.9694	0.7965
DMU 29	26.5	6.0	0.58	630	28	0.9561	0.8976
DMU 35	6.0	6.0	4.80	683	41	0.9913	0.9418
DMU 47	5.5	15.0	9.90	2809	154	0.9867	0.9056
DMU 51	6.0	6.0	2.30	688	71	0.5666	0.6697
DMU 55	5.5	11.0	6.99	2905	146	0.8179	0.8164
DMU 58	10.1	1.2	0.30	688	13	1.0000	0.8576
DMU 61	5.5	11.0	9.90	2809	154	0.9917	0.9890
DMU 62	6.0	7.8	6.99	2905	118	0.9714	0.9781

A graphical representation of the comparison between the CSLE computed and the predicted efficiency scores is shown in Figure 6.2.

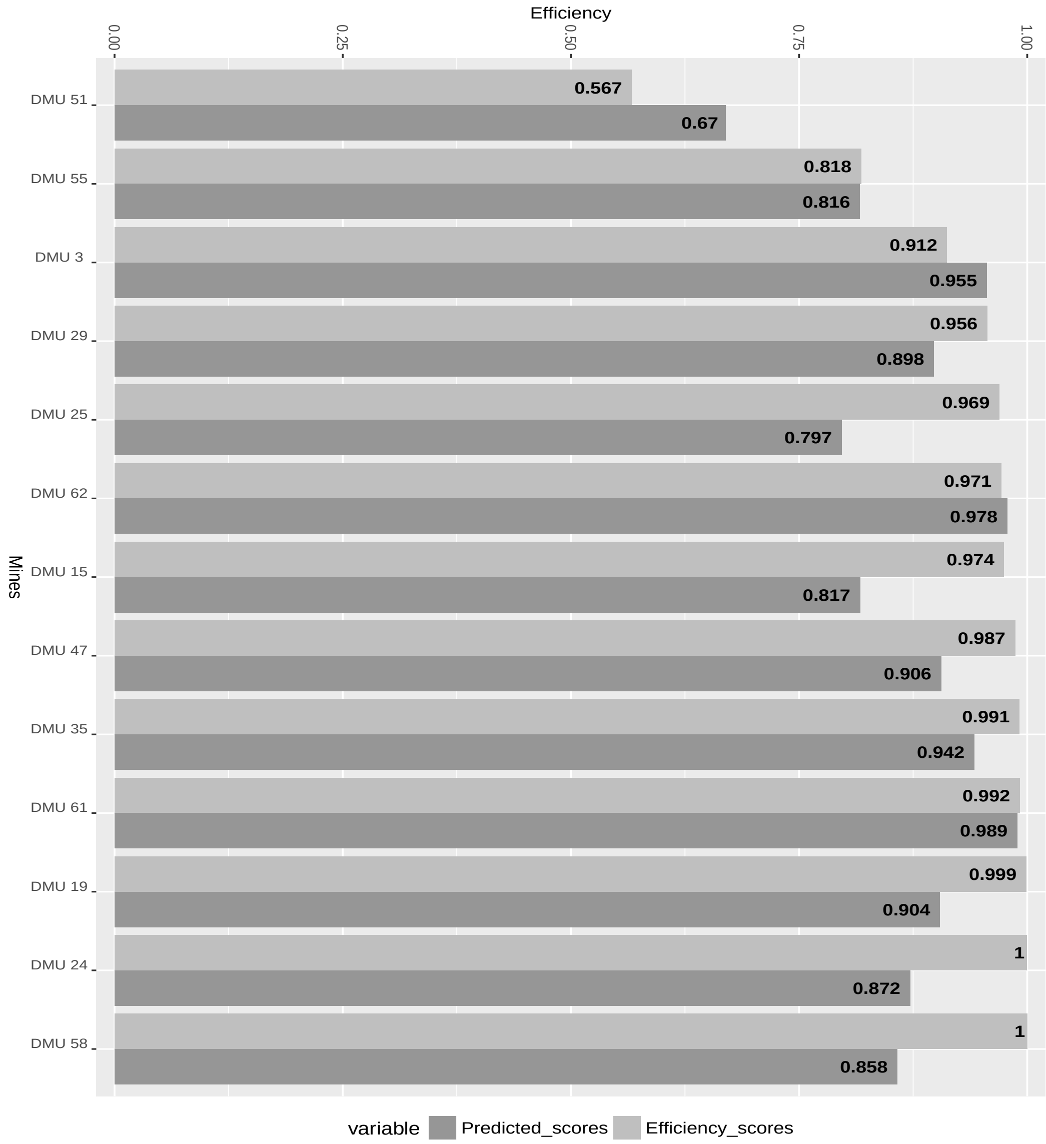

Figure 6.2: Comparison of the predicted and CSLE efficiency scores.

The efficiency scores in Figure 6.2 are presented in increasing order and were obtained from the CSLE model. DMU_{55}, DMU_{61}, and DMU_{62} were estimated very well, whereas DMU_{58}, which appeared to be efficient, was underestimated by 14.3%. Based on the performance of the model on the test set, the largest overestimation was 4.35% for DMU_{3}. The underestimation was 17.29% for DMU_{25}. Therefore, the model can be used to provide an indication of the efficiency score in a range of $\pm 20\%$, which can be helpful in the preliminary stage of a new project.

6.2 Production rate predictive model

To formulate the predictive model for the initial selection of the production rate of a coal mine operation in this research, the Cobb–Douglas function was used. This function shows the relationship between inputs such as capital and labour and the amount of outputs that can be produced using those inputs.

In mining operations, production studies has been reviewed and illustrated by Shahabi et al. (2009). These studies were then used to develop production functions using major factors considering capital and labour force for Iranian mines, including coal, ferrous, lead and zinc, copper, barite, kaolin, and decorative stones. The authors aimed at studying the efficient management of the mining sector in Iran. The study suggested that the Cobb–Douglas function is suitable for estimating the production of the mines, including capital, labour, research and development costs, and investment ratios as inputs. However, the study used all mines together to specify the model and did not generate a model for coal-specific deposits.

In this research, the Cobb–Douglas function was used to develop the predictive model for the production rate in surface coal mines. It is represented by Equation (6.8). The inputs into the extraction subsystem of the CSLE model include Capital Expenditure (CAPEX), number of employees, and Stripping Ratio (SR). The output is the Run-of-Mine (ROM) (Mt/yr). The regression coefficients in Equation (6.8) and the variables that significantly explain the production rate will be specified using the OLS method.

$$\log(Q_j) = \alpha_o + \sum_{f \in \boldsymbol{F}} \alpha_f \log(x_{fj}^m) + \epsilon_j \tag{6.8}$$

where α_f is the parameter $f \in \boldsymbol{F}$ for the model, Q_j is the ROM, which represents the production rate (Mt/yr) of DMU $j \in \boldsymbol{J}$, and x_{fj}^m is the input in the extraction subsystem (m).

6.2.1 Estimation of the parameters of the predictive model for production rate

The datasets used to estimate the regression coefficients in Equation 6.9 are indicated in Table D.1. These are the same training datasets used to specify the predictive model for technical efficiency in subsection 6.1.1. The inputs were used as predictor variables and ROM was used as the response variable. The bootStepAIC package was also applied in this case. The parameter estimation process was carried out and the results are given in Table 6.3. The results suggest that the logarithms of CAPEX, SR, and mining employees are significant at 5%. The coefficient of determination of $R^2 = 0.95$ suggests that 95% of the variation of the logarithm of the production rates of data used can be explained by the logarithms of the CAPEX, SR, and mining employees variables.

Table 6.3: Regression results

	Dependent variable:
	log(ROM)
log(CAPEX)	0.0490^{**}
	(0.0231)
log(SR)	-0.5817^{***}
	(0.0816)
log(Employees)	1.2020^{***}
	(0.0459)
Constant	-4.6022^{***}
	(0.2143)
Observations	52
R^2	0.9492
Adjusted R^2	0.9460
Residual Std. Error	0.2757 (df = 48)
F Statistic	299.0787^{***} (df = 3; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

The number of mining supply systems used for building the models were 52. The coefficient of determination, R^2, suggests that 95% of the variability in technical efficiency scores can be explained by the predictors indicated in Table 6.3. The resulting RMSE is 0.2757. To explain the final model, consider the following definitions:

M is the logarithm of estimated employees in mining.

Q is the logarithm of estimated production rate tonnes/year ROM.

S is the logarithm of stripping ratio SR.

C is the logarithm of estimated CAPEX (US\$M).

$$\log(Q) = -4.6022 + 0.0490 \times \log(C) - 0.5817 \times \log(S) + 1.2020 \times \log(M) \tag{6.9}$$

Equation (6.9) represents the general function, which can help in the initial selection of the production rate for the extraction subsystem of the mining supply system producing coal for local and export markets. The resulting production rate should be optimized iteratively by projecting it using the efficiency scores obtained from the technical efficiency predictive model given by Equation (6.7).

6.3 A use case of the technical efficiency predictive model

After the formulation of the predictive models, we address the 'so what?' question by demonstrating a use case of the predictive model for technical efficiency in this section. To help the management of a new mining project make decisions about the operating variables, we use an example of a simulated mine, DMU_{35}. This DMU was not initially used to specify the parameters of the predictive model of the technical efficiency. It was used only as a test set, as indicated in Table 6.2, for which the details are discussed in this section.

Use of the predictive model occurs in cases in which application of the CSLE model is limited because of insufficient available mine data for computation of the relative technical efficiency scores of mining supplying systems, including those of new mine projects. The application of the predictive model to a new project will help the mine position itself competitively based on the variables established by mining management.

The variables of DMU_{35} are shown in Table 6.4. If the variables indicated in the table are the estimates established by the mining management that will be implemented during operation, what will be the mine efficiency for competitiveness?

Table 6.4: Variables of DMU_{35}

Variables	**Amount**
Ash(%)	6.0
Plant capacity(Mt/yr)	6.0
Export(Mt/yr)	4.8
Number of plant employees	41
Precipitation (mm)	683

The efficiency score of DMU_{35} was computed by substituting the variables given in Table 6.4 into Equation (6.7) as follows:

$$\hat{\theta} = 0.7453 + .0105 \times 6 - 0.0209 \times 6 + 0.0706 \times 4.8 - 0.0032 \times 41 + 0.0001 \times 683.$$
$$= 0.9589.$$

The efficiency score of DMU_{35} from the predictive model is 0.9589 and that obtained from the use of the CSLE model for simulated DMUs is 0.9913. The results show that the predictive model for technical efficiency generated an error of 0.0324, which represents a 3.24% underestimation of the efficiency score calculated using the CSLE model.

The predicted efficiency score of 0.9589 obtained from the predictive model suggests an improvement of 4.1% for the mine to be efficient. The mining management can use this efficiency score to reduce the discretionary inputs or increase the value of one or more outputs of the variables that were initially established, together with some of the discretionary variables shown in Table 6.4, to estimate its competitiveness. For example, the washing of coal can be reduced to 5.75% by multiplying the efficiency score by the ash content of 6%.

6.4 Sensitivity analysis of the model parameters

Sensitivity analysis was done in order to understand the uncertainty of the efficiency output from the predictive model developed in Section 6.1.1. In order to identify the model parameters that influences the efficiency, data sets applied for illustration of the use case of the predictive model for technical efficiency in Section 6.3 are revisited for sensitivity analysis. The method used in sensitivity analysis of predictive model is by varying one parameter of the model at a time keeping the other parameters fixed and analysing the result of model estimates (Hamby 1994). Each parameter was varied from -10% to 10% at an increment of 5% while keeping the other parameters of model constants. The results for the variation of the each model at a time are presented in Figure 6.3.

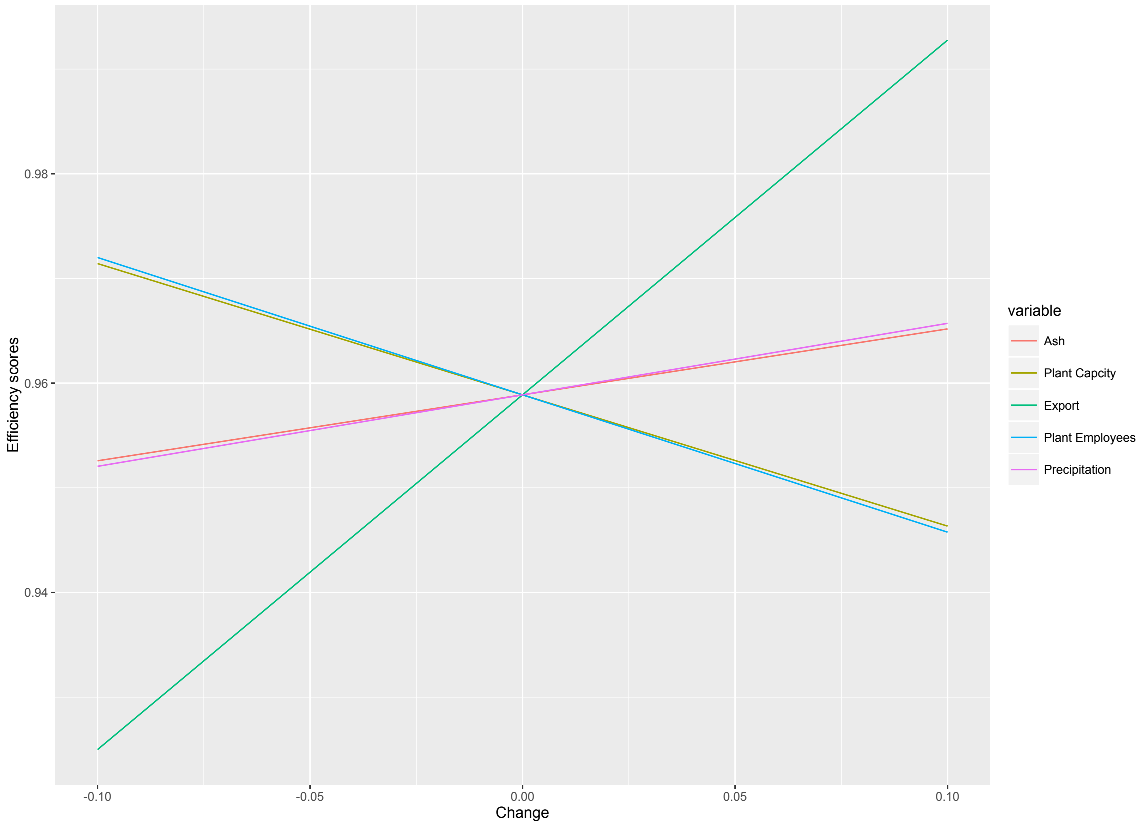

Figure 6.3: Sensitivity analysis of the model parameters.

From Figure 6.3 it can be observed that the most sensitive parameter that influences the results of efficiency scores is coefficient of export variable. For example varying this parameter by -10% or 10% from the base case results to respective change of -3.3% and 3.3% of efficiency score estimates. The predicted efficiency score slightly increase with decrease of the parameters of estimation for either ash or precipitation in the model when varied by -10% each parameter at a time and decrease efficiency with their increase from base case by 10%. On the other hand, the efficiency score slightly decrease with a change of -10% of parameters of estimation for either plant capacity or employees from the base case and vice versa. Therefore, the efficiency is sensitive to all parameters but most sensitive to the parameter of export variable in the model.

6.5 Conclusion

The predictive models formulated can be used to predict technical efficiency scores and to select the initial production rate of a new surface coal mine. These can be used in cases in which there is insufficient data for computing the efficiency scores using the CSLE and special-case models developed in this research.

The performance of the predictive models are constrained within the data generated from a conditional simulation of the characteristics of the collected datasets and used to formulate the models. Simulated data can be used in regression modelling techniques and analysis (Hill and Malone 2004; Lobell and Burke 2010). The resulting model becomes useful provided the data generated by the simulation represent the expected real-world characteristics shown by the samples of the collected dataset.

To use the predictive models, the new mine can estimate their technical inputs first and then predict the technical efficiency in order to assess whether the mine will attain the efficiency of the best practice mines once it starts operating. The mining management can iteratively specify the technical inputs to achieve the desired efficiency.

Chapter 7

Conclusion

Measuring the technical efficiency of coal mining supply systems makes it possible to determine the competitiveness of both new and operating surface mines. The evaluation of the efficiency should be based on a structure that represents the mining supply system in order to show how the mining supply system uses inputs in its subsystems to generate the desired outputs. This will help the management of a mining supply system to identify the source of inefficiency and make improvements to achieve competitiveness in terms of best practices in producing thermal coal using similar inputs that can vary between producers.

Mine efficiency requires the selection and optimization of technical discretionary input variables such as capital that the management of the mining supply system can control when operating under a given set of non-discretionary variables that management cannot control, such as fixed distance from the market. If the mine is inefficient, it implies the use of extra inputs such as capital, referred to as over-capitalization, which will lead to financial loss for the mine project. This risk can be minimized through efficiency measurements and optimization of the inputs before implementation in the production stage.

The research explored the competitiveness of a surface coal mining supply system producing coal for specific markets—export and local, export only, and local only—using the Data Envelopment Analysis (DEA) method. Furthermore, the research formulated predictive models that can help evaluate a new mine when insufficient data is available to include with the new mine. The research provides findings to the questions stated, limitations, and suggests further investigation to improve knowledge of the evaluation of mining supply system performance and competitiveness of both operating mines and new ones.

7.1 Findings with regard to the research questions

The primary research question stated in Chapter 1, which this thesis intended to answer, was *How can a new surface mine producing thermal coal evaluate its competitiveness rela-*

tive to other operating coal mines considering each mine's specific variables those that mine management can control and those it cannot, given a market of thermal coal?

To answer the primary research question, secondary questions were established, which are reviewed here. The findings from the research are also discussed.

> *What model representing the structure of the mine coal supply can be used to measure the relative technical efficiency of a mine considering variables that mine management can control and those it cannot?*

Mathematical models representing the Combined System for Local and Export (CSLE) model and two special-case models for systems that produce and supply thermal coal to the markets were formulated in Chapter 4 using the DEA method. The structure of a surface coal mine production and supply system considered for the CSLE model, consisted of extraction, washing, and port subsystems. These were used to formulate the models for computing the relative technical efficiency representing the CSLE model. The two special-case models of CSLE generated were Export Coal Mine Supply (ECMS) and Local Coal Mine Supply (LCMS), representing an export structure only and a local supply structure only, respectively.

Mines can use these models to evaluate their technical efficiency relative to that of similar producers of thermal coal by taking into consideration the influence of discretionary and non-discretionary variables.

> *How can it be determined which are best practice surface mines producing thermal coal given the unique mine variables?*

To identify the best-practice mines, the CSLE model and its special-case models were solved using computer code developed by the candidate, as presented in Appendix C. The computer code was an implementation of the mathematical model for CSLE. The code used data simulated for each of the mining supply systems, as discussed in Chapter 3, to computer the efficiency scores. The results showed the ability of the models to discriminate between efficient mines, which form the envelope of best practice, and inefficient mines that require improvement. The results and interpretation of the models and their implications are presented in Chapter 5. To solve the special-case models, the code was modified to conform with the mathematical formulations representing them. From the results of the technical efficiency scores, it was found that the best-practice mining supply systems were those that had efficiency scores of 1 relative to the scores of other mines. These are the mines that had no further possibility of gaining technical efficiency by reducing their inputs. The best-practice mines can be used as reference mines by inefficient mines for improvement.

> *What models can be used for predicting the technical efficiency and performance of a surface mine producing thermal coal?*

This question was addressed in Chapter 6, in which the predictive models were developed and explained. The model for predicting the technical efficiency is shown by Equation 6.7, whereas that for the initial selection of production rate is given by Equation 6.9. The technical efficiency predictive model was developed from efficiency scores obtained by solving the CSLE model using both discretionary and non-discretionary simulated input variables, as detailed in Chapter 5. A new surface coal mine in the real world with characteristics similar to those of the simulated mines can thus predict its efficiency for competitiveness using the model in Equation 6.7 in case there is insufficient mine data to relatively position itself competitively. The predictive model in Equation 6.9 can only be used for the initial selection of the production rate, which has to be optimized, together with other discretionary variables, using the efficiency scores obtained for the respective mining supply systems.

7.2 Research contribution

This research contributes the following to scientific knowledge:

- It provides a new perspective on evaluating mine projects as integrated systems considering the unique variables of the project and determining the optimal levels of inputs to ensure the efficiency and cost effectiveness of the mine for competitiveness. Thus, it contributes to the literature on mine projects and performance evaluations.

- Mining companies can use the developed DEA models to evaluate subsidiary mines supplying coal to the markets. In this way, a mining company can determine the resources that may need to be supplied to the individual subsidiary project.

- Management can use the DEA models to identify mines that can be used as benchmarks for an inefficient mine. The models can help choose a good project for investment from a given list of projects, taking into account their technical efficiency over the investment period.

- The research generated computer code that can be used by managers of mining supply systems to evaluate a project's competitiveness using a preferred set of performance indicators, in addition to the variables used in this research.

7.3 Suggestions for further research

The research work was done and compiled in Chapters 1 to 6, in which areas that need further investigation to add to the existing knowledge were identified. These include limitations of the models and some opportunities for improvement.

The models were evaluated for surface coal mining supply systems for thermal coal. An investigation and evaluation of metallurgical coal supply systems are also essential to provide guidelines for the evaluation of their competitiveness performance. Furthermore, the need to develop a hybrid model is important. This can apply to parallel operations and supply systems of both metallurgical and thermal coal. A hybrid DEA model for thermal and metallurgical coal will help evaluate the efficiency of the supply system relative to others.

The capacity of the port infrastructure for mines producing and exporting coal was assumed to be flexible in terms of expansion in this model. This can be a limitation when there are more mines supplying coal to a port whose capacity is small and cannot be expanded. Evaluation of the efficiency considering the capacity of the port could indicate that the efficiency of the port changes with increasing supply of export tonnage. The effect of the capacity of the port should be investigated considering the inventories of coal stockpiled at the port as required.

It is a challenge to evaluate a mine that consists of surface extraction and underground methods operating concurrently. Because the two methods contribute to the mine portfolio of a combined structure, it is suggested that a model should be developed that can consider a combined system running both extraction methods to produce and supply coal to the market. The combined system should consider first the underground method using subsystems such as extraction for stope operation and hoisting using inputs such as number of employees, ventilation, and hoisting capacity needed to deliver the coal to the surface and then add it to the mining subsystems of the surface mine methods discussed in this research. The development of a model that helps determine the efficiency of the parallel mining systems, those using both the surface mine method and underground mine operations, needs to be investigated.

The models formulated in this research can be applied for surface coal mining supply systems having primary washing plants only. Secondary washing plants were not considered in the coal mine supply system. It is suggested that the model be extended to include secondary washing plants of the coal as a subsystem, which will enable evaluation of the efficiency of mines using primary and secondary washing plants.

Moreover, the research is suggested for large mining companies that consist of subsidiary mines, such as those mining gold, iron, copper, platinum, and uranium. The research will help these mining companies assess the efficiency of its subsidiaries and identify those that require improvement. It can also be used to establish a baseline before making any decisions. These also are capital-intensive mines, hence the management needs to ensure that the mines are efficient and cost effective.

In Chapter 6, we formulated predictive models mainly to predict the technical efficiency of a new mine when there is insufficient information about other mines to use to determine the relative efficiency scores. However, both discretionary and non-discretionary variables could only explain 54.5% of the variation in the technical efficiency, whereas 45.5% remains unexplained. This suggests a further study to investigate the influence of qualitative variables that were not investigated in this research, such as the effect of workers' morale, salary disputes between the top management and the workers in the mine, as is happening now in South African mines, and the level of work satisfaction among the members of the production teams. These can influence production output, which in turn impacts the technical efficiency.

The future work should take into consideration of the safety and environmental impact on evaluating competitiveness and performance prediction of mining operations. Communities would like to work with mines that care the safety of their workers and environment surrounding the mines. This will extend the use and the benefits of the models formulated in this research.

The assumption of the model on the virtual inputs and output of value function is linear which gives challenges when dealing with datasets consists of the nonlinearity between some inputs or outputs with their virtual value functions. There is a need of further investigation to account of the non-linearity of some partial virtual value functions and their inputs or outputs in the DEA models developed by this research.

The research has explained the use case of the CSLE model using simulated data from limited number of real mines, this has illustrated the use of the model to support in decision making. It is suggested that the further investigation to verify the application of the models can be done using the identified real mines as Decision Making Units (DMUs) and evaluate the comparative performance of the models and the actual performance of the mines. This needs time for the case studies to be conducted.

7.4 Recommendations

The predictive models for technical efficiency and the initial selection of the production rate should only be used in predictions for mines with characteristics similar to those that were simulated and used to build the models.

Before carrying out optimisation of the value of the mine project, it is recommended that the DEA models developed be used to test the competitiveness strategy, which influences the value of the mine when it is in operation. This will help in choosing parameters such as effective capital for the project. They can then be used in the optimisation algorithms for mine valuation based on Net Present Value (NPV).

Bibliography

Afonso, A. and Aubyn, M. S. (2006). Cross-country efficiency of secondary education provision: A semi-parametric analysis with non-discretionary inputs. *Economic modelling*, 23(3):476–491.

ALLEN, H. (2012). Aspects of evaluating mining projects. http://www.maden.org.tr/resimler/ekler/aad95253aee7437_ek.pdf.

Angloamerican (2012). News letter 15 feb 2013. http://www.angloamerican.com/~/media/Files/A/Anglo-American-PLC-V2/press-release/releases/2013pr/pr2013-02-15.pdf. financial report.

Angloamerican (2013). Focus: operating smarter. http://www.angloamerican.com/~/media/Files/A/Anglo-American-PLC-V2/report-builder-2014/annual-report/aa-ar14-interactive-final.pdf. financial report.

Auguie, B. (2015). *gridExtra: Miscellaneous Functions for "Grid" Graphics*. R package version 2.0.0.

Austin, P. C. and Tu, J. V. (2004). Bootstrap methods for developing predictive models. *The American Statistician*, 58(2):131–137.

Australian Government Department of the Environment (2014). National greenhouse accounts factors. http://www.environment.gov.au/system/files/resources/b24f8db4-e55a-4deb-a0b3-32cf763a5dab/files/national-greenhouse-accounts-factors-2014.pdf. Accessed: 2014-07-24.

Baafi, E. and Mirabediny, H. (1998). Dragline digging methods in Australian strip mines—a survey. In *Proceedings of Coal 1998: First Australasian Coal Operators' Conference*, pages 313–324. University of Wollongong and the AusIMM, Wollongong, Australia.

Banker, R. D. and Natarajan, R. (2008). Evaluating contextual variables affecting productivity using data envelopment analysis. *Operations research*, 56(1):48–58.

Behr, A. (2016). *Production and Efficiency Analysis with R*. Springer International Publishing.

Berkelaar, M. et al. (2015). *Package: lpSolve*. Version:5.6.7.

Bogetoft, P. and Otto, L. (2015). *Benchmarking with DEA and SFA*. R package version 0.26.

Bourne, M., Melnyk, S., Faull, N., Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., Gray, D., et al. (2007). Towards a definition of a business performance measurement system. *International Journal of Operations & Production Management*, 27(8):784–801.

Budeba, M. D., Joubert, J. W., Webber-Youngman, R., and Shafiee, S. (2016). Predicting the efficiency of a surface coal mine for competitiveness. *International Journal of Mining, Reclamation and Environment*. http://dx.doi.org/10.1080/17480930.2015.1135669.

Budeba, M. D., Joubert, J. W., and Webber-Youngman, R. C. W. (2014). Modelling and determining the technical efficiency of a surface coal mine supply chain. *Journal of the Southern African Institute of Mining and Metallurgy*, 114(12):1001–1008.

Budeba, M. D., Joubert, J. W., and Webber-Youngman, R. C. W. (2015). A proposed approach for modelling competitiveness of new surface coal mines. *Journal of the Southern African Institute of Mining and Metallurgy*, 115(11):1057–1064.

Bullock, R. L. (2011). Accuracy of feasibility study evaluations would improve accountability. *Mining Engineering*, 63(4):78–85.

Chan, M. (2008). *Coal supply and cost under technological and environmental uncertainty*. PhD thesis, Carnegie Mellon University. http://search.proquest.com/docview/304667434.

Charnes, A. (1994). *Data Envelopment Analysis: Theory, Methodology, and Applications*. Kluwer Academic Publishers.

Cook, W. D., Zhu, J., Bi, G., and Yang, F. (2010). Network dea: Additive efficiency decomposition. *European Journal of Operational Research*, 207(2):1122 – 1129.

Cooper, W., Seiford, L., and Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Springer, 2nd edition.

CPA Australia (2011). Summary of the australian governments climate change plan. http://brians-satchel.com/test/downloads/CPA%20Summary-of-australian-governments-climate-change-plan.pdf. report.

Curley, R. (2011). *Fossil Fuels*. Energy : past, present, and future. Britannica Educational Pub.

Darling, P. (2011). *Cost estimating for surface coal mines*, pages 281–293. Society for Mining, Metallurgy, and Exploration (SME), 3rd edition.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological methods*, 2(3):292.

Despotis, D. K., Stamati, L. V., and Smirlis, Y. G. (2010). Data envelopment analysis with nonlinear virtual inputs and outputs. *European Journal of Operational Research*, 202(2):604 – 613.

Docker, D. (2011). Coal sale and purchase agreements: The buyer's perspective. In *27th International Coal Supply Contracts & Transport Logistics*.

Efron, B. and Tibshirani, R. (1994). *An Introduction to the Bootstrap*. Taylor & Francis.

Ernst & Young Global Limited (2015). Opportunities to enhance capital productivity, mining and metals megaprojects. reserch report.

Ernst Young (2014a). Business risks facing mining and metals 2014-2015. http://www.ey.com/Publication/vwLUAssets/EY-Business-risks-facing-mining-and-metals-2014%E2%80%932015/$FILE/EY-Business-risks-facing-mining-and-metals-2014%E2%80%932015.pdf. Accessed: 2015-05-05.

Ernst Young (2014b). Productivity in mining a case for transformation. http://www.ey.com/Publication/vwLUAssets/EY-Productivity-in-mining/$FILE/EY-Productivity-in-mining.pdf. Accessed: 2016-07-04.

Fang, H., Wu, J., and Zeng, C. (2009). Comparative study on efficiency performance of listed coal mining companies in China and the US. *Energy Policy*, 37(12):5140–5148.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3):253–290.

Fried, H., Lovell, C., and Schmidt, S. (2008). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press, USA.

Glencore (2016). Mangoola open cut fast facts. http://www.mangoolamine.com.au/EN/Publications/CommunityFactsheet/MangoolaOpenCutFactSheet_20150401.pdf. Accessed: 2016-02-25.

globalCOAL (2015). Scota specifications. https://www.globalcoal.com/Brochureware/standardTradingContract/specifications/.

Gordon, R. L. (1976). The economics of coal supply -the state of the art. *Energy*, 1(3):283 – 289.

Gupta, N. and Grover, S. (2013). Introduction to modeling and simulation. *Int. J. IT, Eng. Appl. Sci. Res*, 2(4):45–50.

Haftendorn, C., Holz, F., and Hirschhausen, C. (2012). The end of cheap coal? a techno-economic analysis until 2030 using the coalmod-world model. *Fuel*, 102(0):305 – 325.

Hamby, D. (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32(2):135–154.

Hartman, H. and Mutmansky, J. (2002). *Introductory Mining Engineering*. Wiley.

Hartman, H. L. (1992). *SME mining engineering handbook*, volume 1. Littleton, Colo. : Society for Mining Metallurgy and Exploration, 2nd edition.

Hesterberg, T. C. (2015). What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *The American Statistician*, 69(4):371–386.

Hill, C. M. and Malone, L. C. (2004). Using simulated data in support of research on regression analysis. In *Proceedings of the 36th conference on Winter simulation*, pages 967–973.

Hlavac, M. (2015). *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Harvard University, Cambridge, USA. R package version 5.2.

Hollingsworth, B. (2003). Non-parametric and parametric applications measuring efficiency in health care. *Health care management science*, 6(4):203–218.

Höök, M., Zittel, W., Schindler, J., and Aleklett, K. (2010). Global coal production outlooks based on a logistic model. *Fuel*, 89(11):3546 – 3558.

Hustrulid, W. and Kuchta, M. (2006). *Open Pit Mine Planning And Design*, pages 11–23, 116–137. John Wiley & Sons.

International Energy Agency (2011). Coal medium-term market report-2011, market trends and projections to 2016. http://www.worldcoal.org/resources/coal-statistics/.

International Energy Agency (2016). Coal. https://www.iea.org/aboutus/faqs/coal/. Accessed: 2016.

IntierraRMG (2014). Raw materials data coal. www.intierrarmg.com.

Joubert, J. (2010). Data envelopment analysis: an overview. Working paper, 14, Optimisation Group, Industrial and Systems Engineering,University of Pretoria.

Kolesnikov, I. (2015). Coal Prices: Long Term Forecast to 2020 Data and Charts. Available at http://knoema.com/xfakeuc/coal-prices-long-term-forecast-to-2020-data-and-charts.

Kong, S. (2007). *Linear programming algorithms using least-squares method.* PhD thesis, Citeseer.

Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling.* Springer.

Kulshreshtha, M. and Parikh, J. K. (2002). Study of efficiency and productivity growth in opencast and underground coal mining in india: a dea analysis. *Energy Economics*, 24(5):439–453.

Kumar, S. and Gulati, R. (2008). An examination of technical, pure technical, and scale efficiencies in indian public sector banks using data envelopment analysis. *Eurasian Journal of Business and Economics*, 1(2):33–69.

Kumbhakar, S. and Lovel, C. (2003). *Stochastic frontier analysis.* Cambridge University Press.

Kurkiewicz, D., Williams, M., and Grajales, C. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research and Evaluation*, 18(11).

Lei, T. and Ding, R.-j. (2008). Efficiency assessment of coal mine safety input by data envelopment analysis. *Journal of China University of Mining and Technology*, 18(1):88–92.

Leinart, J. B. and Schumacher, O. L. (2010). The role of cost estimating in mine planning and equipment selection. In *Mine Planning and Equipment Selection - MPES 2010*, pages 69–80. The Australasian Institute of Mining and Metallurgy (The AusIMM).

Lobell, D. B. and Burke, M. B. (2010). On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, 150(11):1443–1452.

Long, K. R. (2009). A test and re-estimation of taylor's empirical capacity–reserve relationship. *Natural resources research*, 18(1):57–63.

Long, K. R. (2011). Statistical methods of estimating mining costs. In *SME Annual Meeting and Exhibit and CMA 113th National Western Mining Conference 2011*, pages 147–151.

Lotfi, F. H., Jahanshahloo, G., and Esmaeili, M. (2007). Non-discretionary factors and imprecise data in dea. *International Journal of Math Analysis*, 1(5):237–246.

Machiwal, D. and Jha, M. (2012). *Hydrologic Time Series Analysis: Theory and Practice*, chapter 3, pages 44–45. Springer Netherlands.

Maria, A. (1997). Introduction to modeling and simulation. In *Proceedings of the 29th conference on Winter simulation*, pages 7–13. IEEE Computer Society.

Markovits-Somogyi, R. (2012). *Complex technological and economic efficiency assessment methods in freight transport and logistics with special emphasis on data envelopment analysis*. PhD thesis, Budapest University of Technology and Economics.

McCarthy, P. (2013). Why feasibility studies fail.

Merrow, E. (2011). *Industrial Megaprojects: Concepts, Strategies, and Practices for Success*. John Wiley & Sons.

Mitra, R. and Saydam, S. (2012). *Surface coal mining methods in Australia*. INTECH Open Access Publisher.

Mohnot, J. K., Singh, U. K., and Dube, A. K. (2001). Formulation of a model for determining the optimum investment, operating cost and mine life to achieve planned profitability. *Mining Technology*, 110(2):129–132.

Morgan, D. (2014). Global commodities beyond the boom. https://www.ausimm.com.au/content/docs/branch/2014/sydney_2014_07_presentation.pdf. Accessed: 2015-06-08.

Murillo-Zamorano, L. R. and Vega-Cervera, J. A. (2001). The use of parametric and non-parametric frontier methods to measure the productive efficiency in the industrial sector: A comparative study. *International Journal of Production Economics*, 69(3):265–275.

Murrell, P. (2002). *The Grid Graphics Package*. Version: 0.7-4.

Myanmar, A. M. (2016). Geology and nonrenewable mineral resources. http://www.slideshare.net/aung3/chapter14-37590121. Accessed: 2016-02-25.

Nedelea, I. C. and Fannin, J. M. (2013). Technical efficiency of critical access hospitals: an application of the two-stage approach with double bootstrap. *Health care management science*, 16(1):27–36.

Noort, D. and Adams, C. (2006). Effective mining project management systems. In *Proceedings of the International Mine Management Conference 2006*, pages 87–96.

Oraee, K., Bakhtavar, E., and Shahriar, K. (2008). An approach towards ascertaining open-pit to underground transition depth. *Journal of Applied Sciences*, 8(23):4445–4449.

Oraee, K., Lashgari, A., Sayadi, A. R., and Oraee, B. (2011). Estimation of capital and operation costs of backhoe loaders. In *2011 SME Annual Meeting & Exhibit and CMA 113th National Western Mining Conference" Shaping a Strong Future Through Mining"*. SME.

Pessanha, J. F. M., Marinho, A., da Costa Laurencel, L., and do Amaral, M. R. d. S. (2013). Implementing dea models in the R program. In *11th International Conference on Data Envelopment Analysis*, Samsun, Turkey. Conference Presentation.

Porcelli, F. (2009). Measurement of technical efficiency. a brief survey on parametric and non-parametric techniques. *University of Warwick*, 11.

Prentice, A. (2005). Quantification of opencast potential within the waikato coalfields using pit optimisation software. In *Coal Operators' Conference, University of Wollongong*.

PwC (2012). Managing large-scale capital projects. http://www.pwc.com/gx/en/mining/school-of-mines-presentations-2012.jhtml.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Ramanathan, R. (2003). *An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement*. SAGE Publications.

Reddy, G. T., Sudhakar, K., and Krishna, S. J. (2013). Bench marking of coal mines using data envelopment analysis. In *International Journal of Advanced Trends in Computer Science and Engineering*, volume 2, pages 159–164, Lords Institute of Engineering and Technology, Hyderabad. Citeseer.

Rizopoulos, D. (2009).

Runge Pincock Minarco (2015). Minimum engineering study requirement. http://www.rpmglobal.com/wp-content/uploads/2015/08/Issue128-UPDATE-Minimum-Engineering-Study-Requirements.pdf. Accessed: 2016-07-01.

Ruscio, J. and Kaczetow, W. (2008). Simulating multivariate nonnormal data using an iterative algorithm. *Multivariate Behavioral Research*, 43(3):355–381.

Sayadi, A. R., Lashgari, A., Fouladgar, M. M., and Skibniewski, M. J. (2012a). Estimating capital and operational costs of backhoe shovels. *Journal of Civil Engineering and Management*, 18(3):378–385.

Sayadi, A. R., Lashgari, A., and Paraszczak, J. J. (2012b). Hard-rock lhd cost estimation using single and multiple regressions based on principal component analysis. *Tunnelling and Underground Space Technology*, 27(1):133 – 141.

Schernikau, L. (2010). *Economics of the International Coal Trade: The Renaissance of Steam Coal*. Springer Science & Business Media.

Schneid, R. J. and Torries, T. F. (1991). Competitive costs of foreign and u.s. coal in north atlantic markets. *Mining Science and Technology*, 13(1):89 – 104.

Scott, B., Ranjith, P., Choi, S.-K., and Khandelwal, M. (2010). A review on existing opencast coal mining methods within australia. *Journal of mining science*, 46(3):280–297.

Shafiee, S., Nehring, M., and Topal, E. (2009). Estimating average total cost of open pit coal mines in australia. volume 134, pages 134–145.

Shafiee, S. and Topal, E. (2012). New approach for estimating total mining costs in surface coal mines. *Mining Technology*, 121(3):109–116.

Shahabi, R. S., Kakaie, R., Ramazani, R., and Agheli, L. (2009). Estimation of production function for mines in iran. *Journal of Geology and Mining Research*, 1(1):019–024.

Shu-Ming, W. (2011). Evaluation of safety input-output efficiency of coal mine based on dea model. *Procedia Engineering*, 26:2270–2277.

Simar, L. and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, 136(1):31–64.

Stojanovic, L. B., Tomic, P. J., and Stevic, M. Z. (2014). The selection of optimal contour of open pit" delici" near ugljevik. http://doisrpska.nub.rs/index.php/arhivzatehnickenauke/article/download/1580/1509.

Talluri, S. (2000). Data envelopment analysis: models and extensions. *Decision Line*, 31(3):8–11.

US. Bureau of Labor Statistics (2016). Export (end use): Oil drilling, mining, and construction machinery, and equipment [iq210], retrieved from fred, federal reserve bank of st. louis. https://research.stlouisfed.org/fred2/series/IQ210. Accessed: 2016-3-2.

Van Aswegen, G. D. and Koster, M. (2008). From feasibility to reality: A predicament for the mining industry in south africa. In *Management of Engineering & Technology, 2008. PICMET 2008. Portland International Conference on*, pages 1351–1362. IEEE.

Wellmer, F.-W., Dalheimer, M., and Wagner, M. (2007). *Economic evaluations in exploration*. Springer Science & Business Media.

Westcott, P. (2004). Dragline or truck/shovel? some technical and business considerations-unsw/mitsubishi lecture.

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12):1–20.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

wikimapia (2016). Example of open pit coal mine. http://wikimapia.org/18173231/Telford-Open-Cut-Coal-Mine/photo/1434080#/photo/1434080. Accessed: 2016-02-25.

Wikipedia, the free encyclopedia (2016). Open pit coal mine. http://www.wyomingmining.org/wp-content/uploads/2013/05/Dragline-Shovel-Haultruck.jpg. Accessed: 2016-02-11.

Xue, M., Harker, P. T., et al. (1999). Overcoming the inherent dependency of dea efficiency scores: a bootstrap approach. *Unpublished Working Paper, Wharton Financial Institutions Center, University of Pennsylvania*.

Zhu, J. (2008). *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets*. Springer.

Appendices

Appendix A

Source of data declaration proof

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Raw Materials Group

Mr. Mussa Budeba
Department of Mining Engineering
University of Pretoria
Mineral Sciences Building, Room 2-6.1
Pretoria, South Africa

Stockholm, 9 May 2014

To whom it may concern,

Mr. Mussa Budeba, PhD Candidate in the Department of Mining Engineering, University of Pretoria, have during the period 31 May 2013 until 31 May 2014, subscribed to the RMD (Raw Materials Data) database, a database consisting of inter alia minerals and mining entities. For this annual subscription of RMD we have received payment of 1000 EUR. The RMD subscription is provided by Raw Materials Group RMG AB, from January 2014 a subsidiary of SNL Metals & Mining.

Mr. Budeba has used the RMD database to search for information for his thesis on coal mines.

Sincerely,

Irene Geuken
Account Manager
SNL Metals & Mining (formerly IntierraRMG) for
Raw Materials Group, RMG AB

Postal address:
PO Box 3127
SE – 169 03 Solna
Sweden

Office:
Råsundavägen 1, 9th floor
SE- 169 67 Solna
Sweden

Phone: +46 8 744 0065
Fax: +46 8 744 0066
E-mail: raw.materials.group@rmg.se

www.rmg.se

Figure A.1: letter of proof of source of data, source: IntierraRMG (2014)

Appendix B

Simulated data for surface coal mining supply systems

Table B.1: Simulated data for CSLE

DMUs	CAPEX (US$M)	Age (yrs)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	CV (MJ/Kg)	Thickness (m)	Dquantity (Mt/yr)	EmployeesP
DMU 1	2.00	3.00	13.20	208	1.24	5.50	15.50	2.40	0.58	2905	26.10	3.50	0.30	31
DMU 2	631.33	17.00	5.00	503	8.89	10.10	9.00	11.00	2.30	683	27.80	5.50	7.70	69
DMU 3	424.99	9.00	2.40	320	8.89	13.30	9.00	7.80	4.80	683	27.60	8.70	2.00	49
DMU 4	8.31	17.00	5.20	615	11.07	13.30	16.00	11.00	5.40	2121	25.80	15.00	2.00	82
DMU 5	2.00	4.00	13.20	117	0.60	5.50	13.50	1.20	0.58	2905	27.60	3.00	0.30	19
DMU 6	15.88	2.00	2.00	35	0.60	25.00	9.00	1.20	0.30	656	25.80	3.50	0.84	10
DMU 7	8.31	9.00	5.20	516	8.89	10.10	13.50	6.00	2.30	688	27.60	5.50	0.84	71
DMU 8	15.88	3.00	13.20	999	8.89	5.50	15.50	7.80	6.99	2905	27.80	3.00	0.55	124
DMU 9	62.18	4.00	7.80	142	1.24	5.50	11.00	2.40	0.58	688	27.80	3.00	0.30	23
DMU 10	502.71	3.00	2.00	435	14.80	26.50	11.00	15.00	4.80	683	25.80	15.00	7.70	64
DMU 11	631.33	17.00	5.00	755	14.80	10.10	10.00	15.00	5.40	656	28.90	8.70	7.70	99
DMU 12	631.33	19.00	7.80	1024	14.80	5.50	13.50	15.00	6.99	688	28.90	5.00	2.00	129
DMU 13	15.88	17.00	7.80	682	8.89	5.50	16.00	7.80	5.40	2121	27.60	5.00	0.84	89
DMU 14	62.18	4.00	10.30	440	4.00	5.50	11.00	2.40	2.30	2121	27.80	3.00	0.30	59
DMU 15	2.00	17.00	7.80	646	8.31	10.10	15.50	4.00	2.30	2121	25.80	5.50	0.55	85
DMU 16	8.31	2.00	2.40	38	0.60	25.00	9.00	2.40	0.30	683	20.00	10.00	0.84	10
DMU 17	8.31	3.00	7.80	361	4.00	5.50	13.50	2.40	2.30	2121	27.60	3.50	0.30	51
DMU 18	15.88	17.00	10.30	992	11.07	5.50	15.50	11.00	9.90	2809	27.80	5.50	1.28	124
DMU 19	62.18	2.00	2.40	67	1.24	13.30	2.90	1.20	0.58	683	26.10	3.50	1.28	15
DMU 20	62.18	8.00	5.20	775	14.80	10.10	11.00	11.00	9.90	688	27.60	10.00	7.70	101
DMU 21	424.99	2.00	2.00	35	0.60	25.00	2.90	1.20	0.30	630	26.10	5.50	1.28	10
DMU 22	62.18	8.00	2.40	38	0.60	13.30	9.00	1.20	0.30	656	26.10	3.50	0.55	10

DMUs	CAPEX (US$M)	Age (yrs)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	CV (MJ/Kg)	Thickness (m)	Dquantity (Mt/yr)	EmployeesP
DMU 23	502.71	3.00	2.00	154	4.00	25.00	2.90	7.80	0.58	630	25.80	15.00	7.70	28
DMU 24	631.33	3.00	2.00	35	0.60	13.30	2.90	2.40	0.30	630	27.80	3.00	1.28	10
DMU 25	2.00	3.00	5.20	108	1.24	10.10	15.50	4.00	0.58	688	20.00	8.70	0.55	20
DMU 26	502.71	9.00	2.00	61	1.24	13.30	2.90	6.00	0.58	656	27.60	5.50	1.81	14
DMU 27	62.18	17.00	10.30	440	4.00	5.50	10.00	2.40	2.30	688	28.90	3.00	0.30	59
DMU 28	424.99	9.00	13.20	946	8.31	5.50	13.50	4.00	5.40	2121	28.90	3.00	0.30	118
DMU 29	502.71	8.00	5.00	266	4.00	10.10	2.90	6.00	0.58	683	27.80	3.50	1.28	40
DMU 30	502.71	9.00	2.00	154	4.00	26.50	9.00	6.00	0.58	630	25.80	15.00	7.70	28
DMU 31	15.88	3.00	2.00	154	4.00	26.50	9.00	7.80	0.30	656	20.00	15.00	7.70	28
DMU 32	631.33	17.00	2.00	435	14.80	25.00	2.90	11.00	6.99	630	27.80	15.00	7.70	64
DMU 33	62.18	4.00	5.20	273	4.00	5.50	11.00	6.00	2.30	688	27.80	3.50	0.55	41
DMU 34	8.31	8.00	13.20	946	8.31	5.50	13.50	4.00	6.99	2905	27.80	3.00	0.30	118
DMU 35	8.31	19.00	7.80	1024	14.80	6.00	16.00	11.00	6.99	2905	27.60	10.00	1.28	129
DMU 36	631.33	3.00	2.40	320	8.89	13.30	2.90	7.80	4.80	630	27.80	5.50	7.70	49
DMU 37	424.99	3.00	5.20	273	4.00	6.00	9.00	6.00	4.80	683	27.80	3.50	0.84	41
DMU 38	631.33	9.00	2.40	381	11.07	10.10	10.00	11.00	2.30	656	28.90	5.50	7.70	57
DMU 39	15.88	3.00	10.30	173	1.24	5.50	11.00	2.40	0.58	2121	27.80	3.00	0.30	27
DMU 40	424.99	2.00	2.00	35	0.60	26.50	2.90	2.40	0.30	630	20.00	10.00	1.81	10
DMU 41	8.31	2.00	7.80	682	8.89	5.50	15.50	6.00	4.80	2809	26.10	5.00	0.55	89
DMU 42	8.31	4.00	5.20	516	8.89	10.10	15.50	11.00	6.99	2121	26.10	10.00	1.28	71
DMU 43	424.99	8.00	2.40	304	8.31	13.30	11.00	6.00	2.30	683	27.60	8.70	1.81	47
DMU 44	62.18	8.00	5.20	108	1.24	5.50	11.00	2.40	2.30	688	27.80	3.00	0.30	20
DMU 45	424.99	4.00	2.40	304	8.31	25.00	11.00	7.80	2.30	656	25.80	10.00	7.70	47

DMUs	CAPEX (US$M)	Age (yrs)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	CV (MJ/Kg)	Thickness (m)	Dquantity (Mt/yr)	EmployeesP
DMU 46	424.99	8.00	2.40	170	4.00	13.30	10.00	6.00	0.30	683	26.10	8.70	1.81	30
DMU 47	15.88	19.00	13.20	1500	14.80	5.50	15.50	11.00	9.90	2905	27.80	10.00	1.81	181
DMU 48	15.88	3.00	5.00	266	4.00	13.30	10.00	4.00	0.58	683	25.80	8.70	1.28	40
DMU 49	62.18	19.00	10.30	1250	14.80	5.50	16.00	15.00	9.90	2809	27.80	15.00	7.70	154
DMU 50	62.18	2.00	5.20	489	8.31	10.10	10.00	6.00	0.58	683	27.60	3.50	1.28	67
DMU 51	62.18	17.00	13.20	1500	14.80	5.50	13.50	11.00	9.90	2905	28.90	3.50	1.28	181
DMU 52	502.71	8.00	2.00	345	11.07	25.00	9.00	11.00	4.80	656	25.80	15.00	7.70	53
DMU 53	424.99	4.00	5.20	516	8.89	6.00	10.00	6.00	2.30	688	27.80	3.50	1.28	71
DMU 54	15.88	17.00	10.30	992	11.07	5.50	11.00	7.80	5.40	2809	27.80	3.50	0.84	124
DMU 55	502.71	17.00	10.30	1250	14.80	5.50	11.00	11.00	9.90	2121	28.90	3.50	1.28	154
DMU 56	502.71	3.00	2.00	35	0.60	13.30	2.90	2.40	0.30	630	27.60	3.50	0.84	10
DMU 57	8.31	9.00	13.20	1190	11.07	5.50	15.50	11.00	6.99	2905	27.60	5.00	0.84	146
DMU 58	8.31	3.00	13.20	117	0.60	5.50	11.00	1.20	0.58	2809	27.80	3.00	0.30	19
DMU 59	2.00	2.00	2.40	67	1.24	13.30	13.50	2.40	0.58	688	20.00	10.00	0.84	15
DMU 60	2.00	2.00	5.00	59	0.60	10.10	9.00	1.20	0.30	688	25.80	3.00	0.30	13

Table B.2: Simulated data for ECMS

DMUs	CAPEX (US$M)	Age (yrs)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	CV (MJ/Kg)	Thickness (m)	EmployeesP
DMU 1	167.70	4.00	7.00	252	2.80	9.50	9.00	5.50	2.80	640	27.85	10.40	38
DMU 2	1355.80	3.00	3.20	356	8.21	11.00	14.50	12.00	8.20	640	26.20	38.00	53
DMU 3	2.70	9.00	7.00	128	1.20	6.00	3.50	1.50	1.00	676	31.00	1.50	22
DMU 4	1355.80	4.00	3.20	356	8.21	6.00	14.50	12.00	8.20	643	29.50	38.00	53
DMU 5	1355.80	4.00	4.20	421	8.21	8.75	14.50	12.00	8.20	673	27.90	38.00	60
DMU 6	90.30	3.00	7.00	193	2.00	14.00	8.00	2.80	1.40	640	26.20	3.20	30
DMU 7	167.70	4.00	7.00	252	2.80	8.75	9.00	2.80	2.80	663	29.50	10.40	38
DMU 8	90.30	3.00	7.00	193	2.00	14.00	9.00	2.80	1.40	640	26.20	3.20	30
DMU 9	1355.80	4.00	4.20	421	8.21	9.50	14.50	12.00	8.20	643	27.90	38.00	60
DMU 10	167.70	9.00	7.00	193	2.00	6.00	3.50	1.80	1.40	640	31.00	3.20	30
DMU 11	1355.80	9.00	3.20	356	8.21	6.00	8.00	12.00	8.20	673	31.00	38.00	53
DMU 12	2.70	9.00	7.00	128	1.20	6.00	9.00	1.50	1.00	663	31.00	1.50	22
DMU 13	90.30	4.00	7.00	193	2.00	9.50	8.00	1.80	1.40	640	27.90	3.20	30
DMU 14	1355.80	6.00	3.20	356	8.21	6.00	11.00	12.00	8.20	673	29.50	38.00	53
DMU 15	1355.80	4.00	4.20	179	2.80	8.75	9.00	12.00	8.20	643	27.90	38.00	30
DMU 16	90.30	9.00	7.00	193	2.00	6.00	3.50	1.50	1.40	676	31.00	3.20	30
DMU 17	167.70	3.00	7.00	252	2.80	14.00	9.00	5.50	2.80	643	26.20	10.40	38
DMU 18	2.70	4.00	7.00	193	2.00	11.00	9.00	1.50	1.00	640	27.85	3.20	30
DMU 19	90.30	4.00	7.00	193	2.00	9.50	9.00	2.80	1.40	643	27.90	3.20	30
DMU 20	2.70	9.00	7.00	128	1.20	6.00	3.50	1.50	1.00	673	31.00	1.50	22
DMU 21	2.70	4.00	7.00	128	1.20	9.50	3.50	1.50	1.00	643	27.90	1.50	22
DMU 22	260.70	3.00	4.20	245	4.15	9.50	14.50	12.00	3.50	640	27.85	11.50	38

DMUs	CAPEX (US$M)	Age (yrs)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	CV (MJ/Kg)	Thickness (m)	EmployeesP
DMU 23	260.70	4.00	4.20	179	2.80	8.75	9.00	5.50	3.50	673	29.50	11.50	30
DMU 24	167.70	4.00	7.00	193	2.00	8.75	9.00	2.80	1.40	663	29.50	10.40	30
DMU 25	2.70	9.00	7.00	128	1.20	8.75	3.50	1.50	1.00	663	29.50	1.50	22
DMU 26	1355.80	3.00	3.20	356	8.21	14.00	14.50	12.00	3.50	640	26.20	38.00	53
DMU 27	90.30	4.00	4.20	137	2.00	11.00	8.00	1.80	1.40	643	27.85	3.20	24
DMU 28	260.70	3.00	7.00	252	2.80	14.00	9.00	5.50	2.80	640	26.20	11.50	38
DMU 29	2.70	4.00	7.00	128	1.20	9.50	3.50	1.50	1.00	643	27.90	3.20	22
DMU 30	167.70	4.00	7.00	252	2.80	9.50	8.00	2.80	1.40	663	27.90	10.40	38
DMU 31	90.30	4.00	7.00	193	2.00	8.75	3.50	1.50	1.00	643	29.50	3.20	30
DMU 32	1355.80	3.00	7.00	344	4.15	9.50	11.00	12.00	3.50	673	27.85	38.00	49
DMU 33	90.30	4.00	7.00	193	2.00	9.50	9.00	1.80	1.40	673	27.90	3.20	30
DMU 34	167.70	6.00	7.00	193	2.00	8.75	8.00	2.80	1.40	643	29.50	3.20	30
DMU 35	1355.80	3.00	7.00	593	8.21	14.00	14.50	12.00	3.50	643	26.20	38.00	79
DMU 36	167.70	4.00	7.00	252	2.80	9.50	8.00	5.50	2.80	643	27.90	10.40	38
DMU 37	260.70	3.00	3.20	207	4.15	11.00	14.50	12.00	3.50	673	26.20	11.50	34
DMU 38	260.70	3.00	7.00	252	2.80	14.00	11.00	5.50	2.80	640	26.20	10.40	38
DMU 39	260.70	4.00	7.00	252	2.80	11.00	9.00	5.50	3.50	643	27.85	11.50	38
DMU 40	2.70	6.00	7.00	128	1.20	8.75	3.50	1.50	1.00	663	29.50	1.50	22
DMU 41	1355.80	9.00	3.20	151	2.80	6.00	9.00	5.50	3.50	663	31.00	11.50	26
DMU 42	2.70	9.00	7.00	128	1.20	6.00	3.50	1.50	1.00	673	31.00	1.50	22
DMU 43	90.30	9.00	7.00	128	1.20	6.00	3.50	1.50	1.40	640	31.00	3.20	22
DMU 44	1355.80	6.00	3.20	207	4.15	6.00	8.00	5.50	3.50	673	29.50	11.50	34
DMU 45	260.70	9.00	7.00	252	2.80	6.00	8.00	2.80	2.80	673	31.00	11.50	38

DMUs	CAPEX (US$M)	Age (yrs)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	Precipitation (mm)	CV (MJ/Kg)	Thickness (m)	EmployeesP
DMU 46	167.70	9.00	4.20	137	2.00	6.00	8.00	1.80	1.40	676	31.00	3.20	24
DMU 47	260.70	3.00	4.20	179	2.80	14.00	11.00	5.50	2.80	643	26.20	11.50	30
DMU 48	2.70	4.00	7.00	128	1.20	11.00	3.50	1.50	1.00	640	27.85	3.20	22
DMU 49	1355.80	3.00	4.20	421	8.21	9.50	14.50	12.00	8.20	643	26.20	38.00	60
DMU 50	1355.80	3.00	4.20	421	8.21	14.00	14.50	12.00	8.20	643	26.20	38.00	60

Table B.3: Simulated data for LCMS

DMUs	CAPEX (US$M)	SR	Employees	ROM (Mt/yr)	Age (yrs)	Precipitation (mm)	Thickness (m)
DMU 1	4.96	3	48	0.72	7.00	693.00	5.00
DMU 2	50.99	4	171	2.80	7.00	689.00	5.00
DMU 3	62.76	10	109	0.72	6.00	623.00	3.00
DMU 4	4.96	3	143	2.80	7.00	693.00	5.60
DMU 5	31.42	10	109	0.72	6.00	623.00	3.00
DMU 6	4.96	3	48	0.72	6.00	689.00	5.00
DMU 7	62.76	3	242	5.40	10.00	693.00	5.60
DMU 8	4.96	4	58	0.72	6.00	634.50	3.00
DMU 9	62.76	10	322	2.80	10.00	634.50	3.00
DMU 10	50.99	10	322	2.80	7.00	634.50	3.00
DMU 11	50.99	4	290	5.40	7.00	693.00	5.00
DMU 12	31.42	4	87	1.20	7.00	693.00	5.00
DMU 13	4.96	3	73	1.20	7.00	693.00	5.00
DMU 14	50.99	10	545	5.40	7.00	634.50	5.00
DMU 15	31.42	3	242	5.40	10.00	693.00	5.60
DMU 16	31.42	10	322	2.80	7.00	634.50	3.00
DMU 17	31.42	4	87	1.20	7.00	634.50	5.00
DMU 18	4.96	4	87	1.20	6.00	689.00	5.00
DMU 19	50.99	4	290	5.40	10.00	689.00	5.60
DMU 20	31.42	10	322	2.80	6.00	623.00	3.00
DMU 21	50.99	10	322	2.80	6.00	689.00	3.00
DMU 22	50.99	3	242	5.40	10.00	693.00	5.60
DMU 23	4.96	10	109	0.72	6.00	623.00	3.00
DMU 24	50.99	4	290	5.40	10.00	693.00	5.60
DMU 25	4.96	4	58	0.72	6.00	689.00	3.00
DMU 26	62.76	4	290	5.40	10.00	693.00	5.60
DMU 27	4.96	4	58	0.72	7.00	634.50	5.00

DMUs	CAPEX (US$M)	SR	Employees	ROM (Mt/yr)	Age (yrs)	Precipitation (mm)	Thickness (m)
DMU 28	50.99	10	545	5.40	7.00	623.00	5.00
DMU 29	62.76	10	163	1.20	6.00	623.00	3.00
DMU 30	4.96	3	48	0.72	6.00	634.50	3.00

Appendix C

R-code to solve the CSLE models

Computer program for implementing CSLE- model for surface coal mines

Mussa Daniel Budeba

09 November 2015

This program represent surface coal mining system supplying thermal coal for both local and export markets. The program was written in R software.

Step 1. Import the file with data sets.

```
#Import file with simulated mine data
ExportVRSLE <-read.csv("F:/Research/CSLE/CSLEsim-model58.csv")
View(ExportVRSLE)
```

Step 2. Reading data from the columns of the dataframe of the data imported in R.

```
inputsm<-data.frame(ExportVRSLE[c(3,5,6)])
outputsm<-data.frame(ExportVRSLE[c(9,7,8)])
outputsm2<-data.frame(ExportVRSLE[18])
inputsb2<-data.frame(ExportVRSLE[c(10,24)])
outputsb<-data.frame(ExportVRSLE[11])
inputspo<-data.frame(ExportVRSLE[14])
outputspo<-data.frame(ExportVRSLE[15])
N1<-data.frame(ExportVRSLE[4])
N2<-data.frame(ExportVRSLE[19])
N3<-data.frame(ExportVRSLE[20])
N4<-data.frame(ExportVRSLE[23])
vrs0<-data.frame(ExportVRSLE[25])
vrs1<-data.frame(ExportVRSLE[26])
vrs2<-data.frame(ExportVRSLE[27])
vrs3<-data.frame(ExportVRSLE[28])
vrs4<-data.frame(ExportVRSLE[29])
vrs5<-data.frame(ExportVRSLE[30])
N<-dim(ExportVRSLE)[1]
s<-dim(inputsm)[2]
g<-dim(outputsm)[2]
k<-dim(inputsb2)[2]
f<-dim(inputspo)[2]
t<-dim(outputsb)[2]
r<-dim(outputspo)[2]
e<-dim(outputsm2)[2]
x1<-dim(N1)[2]
x2<-dim(N2)[2]
x3<-dim(N3)[2]
x4<-dim(N3)[2]
x<-x1+x2+x3+x4
v<-s+g+k+f+t+r+e
w<-s+g+k+f+t+r+e+x+6
```

Step 3. Creates the matrices of the input and outputs for CSLE mode, as indicated in the equations formulated in Chapter 3 of the PhD thesis.

```
aux0 <- cbind(-1 * inputsm, 0 * outputsm, outputsm2, 0 * outputsb, -1 * inputsb2, -1 * inputspo,
    outputspo, -1 * N1, -1 * N2, -1 * N3, -1 * N4, -1 * vrs0, 1 * vrs1, -1 * vrs2, 1 * vrs3,
    -1 * vrs4, 1 * vrs5)
aux1 <- cbind(-1 * inputsm, outputsm, outputsm2, 0 * outputsb, 0 * inputsb2, 0 * inputspo,
    0 * outputspo, 0 * N1, 0 * N2, 0 * N3, 0 * N4, -1 * vrs0, 1 * vrs1, 0 * vrs2, 0 * vrs3,
    0 * vrs4, 0 * vrs5)
aux2 <- cbind(0 * inputsm, -1 * outputsm, 0 * outputsm2, outputsb, -1 * inputsb2, 0 * inputspo,
    0 * outputspo, 0 * N1, 0 * N2, 0 * N3, 0 * N4, 0 * vrs0, 0 * vrs1, -1 * vrs2, 1 * vrs3,
    0 * vrs4, 0 * vrs5)
aux3 <- cbind(0 * inputsm, 0 * outputsm, 0 * outputsm2, -1 * outputsb, 0 * inputsb2, -1 * inputspo,
    outputspo, 0 * N1, 0 * N2, 0 * N3, 0 * N4, 0 * vrs0, 0 * vrs1, 0 * vrs2, 0 * vrs3, -1 *
        vrs4, vrs5)
aux4 <- cbind(inputsm, 0 * outputsm, 0 * outputsm2, 0 * outputsb, 0 * inputsb2, 0 * inputspo,
    0 * outputspo, 0 * N1, 0 * N2, 0 * N3, 0 * N4, 0 * vrs0, 0 * vrs1, 0 * vrs2, 0 * vrs3,
    0 * vrs4, 0 * vrs5)
aux5 <- cbind(0 * inputsm, outputsm, 0 * outputsm2, 0 * outputsb, inputsb2, 0 * inputspo, 0 *
    outputspo, 0 * N1, 0 * N2, 0 * N3, 0 * N4, 0 * vrs0, 0 * vrs1, 0 * vrs2, 0 * vrs3, 0 *
    vrs4, 0 * vrs5)
aux6 <- cbind(0 * inputsm, 0 * outputsm, 0 * outputsm2, 1 * outputsb, 0 * inputsb2, 1 * inputspo,
    0 * outputspo, 0 * N1, 0 * N2, 0 * N3, 0 * N4, 0 * vrs0, 0 * vrs1, 0 * vrs2, 0 * vrs3,
    0 * vrs4, 0 * vrs5)
auxt <- rbind(aux0, aux1, aux2, aux3, aux4, aux5, aux6)
```

Step 4. Creating a square unit matrix eaquals to the number of variables and combining with the rows of the above to form overall matrix for computations purposes.

```
J<-diag(nrow=w)
matc<-as.data.frame(J)
colnames(matc)<-c("CAPEX","SR","Employees","Moisture","ROM","Ash","Drevenue","Export","Cap",
                  "EmployeesP","CO2Emission","Revenue","Age","Dist.port","Precipitation","Thickness",
                  "a0","a1","a2","a3","a4","a5")
gh<-matc
aux8<-rbind(auxt,gh)
```

Step 5. Setting the assignment operators for the equations in the format and size of the matrices in step 3 above.

```
rhs0<-rep(0,N)
dir0<-rep("<=",N)
rhs1<-rep(0,3*N)
dir1 <-rep("<=",3*N)
dir2<-rep(">=",3*N)
rhs2<-rep(0.005,3*N)
dir3 <-rep(">=",v)
rhs3<-rep(0.000001,v)
dir4 <-rep(">=",x)
rhs4<-rep(0,x)
dir5 <-rep(">=",6)
rhs5<-rep(0,6)
f.dir<-c(dir0,dir1,dir2,dir3,dir4,dir5,"=")
f.rhs<-c(rhs0,rhs1,rhs2,rhs3,rhs4,rhs5,1)
```

Step 6. Load the package (lpSolve) and formulate the system of equations to implement the compuation of efficiency for each DMU with index "i" for the overall system. The systems is composed of the components shown in step 1-step 4 above.

```
library("lpSolve")
for (i in 1:N) {
  obj<-c(0*inputsm[i,],outputsm[i,],outputsm2[i,],outputsb[i,],0*inputsb2[i,],0*inputspo[i,],
         outputspo[i,],-1*N1[i,],-1*N2[i,],-1*N3[i,],-1*N4[i,],-1*vrs0[i,],1*vrs1[i,],
         -1*vrs2[i,],1*vrs3[i,],-1*vrs4[i,],1*vrs5[i,])
  H<-c(inputsm[i,],outputsm[i,],0*outputsm2[i,],outputsb[i,],inputsb2[i,],inputspo[i,],
       0*outputspo[i,],0*N1[i,],0*N2[i,],0*N3[i,],0*N4[i,],0*vrs0[i,],0*vrs1[i,],0*vrs2[i,],
       0*vrs3[i,],0*vrs4[i,],0*vrs5[i,])
  names(H) <- names(aux8)
  f.con <-rbind(aux8,H)
  results <-lp("max",obj,f.con,f.dir,f.rhs,scale=1,compute.sens=TRUE)
  options(digits=4)
  multipliers <-results$solution
  options(digits=4)
  efficiency <-results$objval
  if (i==1) {
    weights <-multipliers
    effcrs <-efficiency

  } else {
    weights <-rbind(weights,multipliers)
    effcrs <-rbind(effcrs, efficiency)
  }
}
```

Step 7. Extract results for the efficiency scores and multipliers and Writting the optimal weightings in the external file.

```
spreadsheet<- cbind(effcrs,weights)
rownames(spreadsheet)<- ExportVRSLE[,1]
colnames(spreadsheet)<- c('efficiency',names(inputsm),names(outputsm),names(outputsm2),
                          names(outputsb),names(inputsb2),names(inputspo),names(outputspo),
                          names(N1),names(N2),names(N3),names(N4),names(vrs0),names(vrs1),
                          names(vrs2),names(vrs3),names(vrs4),names(vrs5))
write.csv(spreadsheet,"resultscrs.csv")
```

Step 8. Solving the efficiency scores of mining subsystem for each DMU.

```
Xin<--1*aux1[,1:3]
names(Xin) <- NULL
inputsmx<-as.matrix(Xin)
Yout<-aux1[,c(4,5,6,7,17,18)]
names(Yout) <- NULL
outputsmY<-as.matrix(Yout)
X<-0
Y<-0
EFmgvrs<-0
for (i in 1:N){
  X[i]<-t(inputsmx[i,])%*%weights[i,1:3]
  Y[i]<-t(outputsmY[i,])%*%weights[i,c(4,5,6,7,17,18)]
  EFmgvrs[i]<-Y[i]/X[i]}
```

Step 9. Solving the efficiency scores of washing subsystem for each DMU.

```
Xinb<--1*aux2[,c(4,5,6,9,10)]
names(Xinb) <- NULL
inputsmxb<-as.matrix(Xinb)
Youtb<-aux2[,c(8,19,20)]
names(Youtb) <- NULL
outputsmYb<-as.matrix(Youtb)
Xb<-0
Yb<-0
EFbgvrs<-0
for (i in 1:N){
  Xb[i]<-t(inputsmxb[i,])%*%weights[i,c(4,5,6,9,10)]
  Yb[i]<-t(outputsmYb[i,])%*%weights[i,c(8,19,20)]
  EFbgvrs[i]<-Yb[i]/Xb[i]}
```

Step 10. Solving the efficiency of port subsystem for each DMU.

```
Xinp<--1*aux3[,c(8,11)]
names(Xinp) <- NULL
inputsmxp<-as.matrix(Xinp)
Youtp<-aux3[,c(12,21,22)]
names(Youtp) <- NULL
outputsmYp<-as.matrix(Youtp)
Xp<-0
Yp<-0
EFpgvrs<-0
for (i in 1:N){
  Xp[i]<-t(inputsmxp[i,])%*%weights[i,c(8,11)]
  Yp[i]<-t(outputsmYp[i,])%*%weights[i,c(12,21,22)]
  EFpgvrs[i]<-Yp[i]/Xp[i]}
```

Step 11. Generating plots of the efficiency scores for CSLE system and its subsystems.

```
library("ggplot2")
library("gridExtra")
library("grid")
EFF_F <- ggplot(resultsct, aes(x = Mines, y = Efficiency)) + geom_bar(width = 0.6,
    fill = "skyblue", stat = "identity") + geom_text(aes(label = round(Efficiency,
    3)), vjust = 0, size = 2.8, angle = 90) + theme(axis.text.x = element_text(angle = 90,
    vjust = -0.1)) + ylab("Overall Efficiency") + scale_y_continuous(expand = c(0,
    0.02))
CSLE1 <- ggplot(EFM_CSLE1, aes(x = Mines, y = EFmgvrs)) + geom_bar(width = 0.6,
    fill = "skyblue", stat = "identity") + geom_text(aes(label = round(EFmgvrs,
    3)), vjust = 0, size = 2.8, angle = 90) + theme(axis.text.x = element_text(angle = 90,
    vjust = -0.1)) + ylab("Extraction Efficiency") + scale_y_continuous(expand = c(0,
    0.02))
CSLE2 <- ggplot(EFM_CSLE2, aes(x = Mines, y = EFbgvrs)) + geom_bar(width = 0.6,
    fill = "skyblue", stat = "identity") + geom_text(aes(label = round(EFbgvrs,
    3)), vjust = 0, size = 2.8, angle = 90) + theme(axis.text.x = element_text(angle = 90,
    vjust = -0.1)) + ylab("Washing Efficiency") + scale_y_continuous(expand = c(0,
    0.02))
CSLE3 <- ggplot(EFM_CSLE3, aes(x = Mines, y = EFpgvrs)) + geom_bar(width = 0.6,
```

```
    fill = "skyblue", stat = "identity") + geom_text(aes(label = round(EFpgvrs,
    3)), vjust = 0, size = 2.8, angle = 90) + theme(axis.text.x = element_text(angle = 90,
    vjust = -0.1)) + ylab("Port Efficiency") + scale_y_continuous(expand = c(0,
    0.02))
EEoverall1 <- grid.arrange(EFF_F, ncol = 1)
```

```
EEoverall2 <- grid.arrange(CSLE1, ncol = 1)
```

```
EEoverall3 <- grid.arrange(CSLE2, ncol = 1)
```

```
EEoverall4 <- grid.arrange(CSLE3, ncol = 1)
```

Table C.1: Excess inputs for CSLE

	CAPEX (US$M)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	EmployeesP
DMU 1	0.00	0.00	0	1.01	4.46	12.57	1.95	0.01	0
DMU 2	299.17	2.37	238	5.24	5.95	5.31	6.49	0.10	33
DMU 3	90.63	0.51	68	1.46	2.18	1.47	1.28	0.06	10
DMU 4	0.00	0.00	0	6.24	7.49	9.02	6.20	0.08	0
DMU 5	0.00	0.00	0	0.00	0.01	0.01	0.00	0.00	0
DMU 6	0.00	0.00	0	0.00	0.06	0.02	0.00	0.00	0
DMU 7	1.18	0.74	73	6.47	7.35	9.82	4.36	0.10	10
DMU 8	10.16	8.44	639	0.13	0.08	0.22	0.11	0.02	79
DMU 9	37.99	4.77	87	0.56	2.49	4.97	1.08	0.01	14
DMU 10	0.00	0.00	0	5.38	9.63	4.00	5.45	0.14	0
DMU 11	261.40	2.07	313	6.63	4.53	4.48	6.72	0.13	41
DMU 12	183.02	2.26	297	5.28	1.96	4.82	5.35	0.10	37
DMU 13	6.41	3.15	275	2.16	1.34	3.90	1.90	0.05	36

	CAPEX (US$M)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	EmployeesP
DMU 14	49.90	8.27	353	0.00	0.00	0.00	0.00	0.03	47
DMU 15	0.00	0.00	0	3.37	4.10	6.29	1.62	0.09	0
DMU 16	0.00	0.00	0	0.00	0.09	0.03	0.01	0.00	0
DMU 17	5.81	5.45	252	0.00	0.00	0.00	0.00	0.03	36
DMU 18	7.79	5.05	487	0.00	0.00	0.00	0.00	0.00	61
DMU 19	11.18	0.43	12	0.00	0.00	0.00	0.00	0.01	3
DMU 20	0.00	0.00	0	0.00	0.00	0.00	0.00	0.05	0
DMU 21	0.10	0.00	0	0.00	0.00	0.00	0.00	0.00	0
DMU 22	5.63	0.22	3	0.00	0.00	0.00	0.00	0.00	1
DMU 23	5.20	0.02	2	2.10	13.16	1.53	4.10	0.04	0
DMU 24	0.28	0.00	0	0.00	0.00	0.00	0.00	0.00	0
DMU 25	0.00	0.00	0	0.53	4.34	6.66	1.72	0.11	0
DMU 26	0.28	0.00	0	0.23	2.47	0.54	1.11	0.01	0
DMU 27	49.97	8.28	354	0.00	0.00	0.00	0.00	0.03	47
DMU 28	260.18	2.59	138	3.27	8.24	2.37	4.90	0.04	21

	CAPEX (US$M)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	EmployeesP
DMU 29	0.27	0.00	0	2.34	15.51	5.27	3.51	0.04	0
DMU 30	0.00	0.00	0	2.60	17.20	5.84	5.06	0.00	0
DMU 31	0.11	0.00	0	0.00	0.00	0.00	0.00	0.10	0
DMU 32	54.93	4.59	241	1.73	2.38	4.76	2.60	0.03	36
DMU 33	0.00	0.00	0	4.69	1.90	5.07	3.48	0.10	0
DMU 34	132.29	0.50	67	1.46	2.18	0.47	1.28	0.06	10
DMU 35	272.58	3.34	175	0.00	0.00	0.00	0.00	0.00	26
DMU 36	105.36	0.40	64	8.31	7.58	7.50	8.25	0.12	10
DMU 37	10.10	6.55	110	0.59	2.63	5.27	1.15	0.01	17
DMU 38	0.00	0.00	0	0.00	0.07	0.01	0.01	0.00	0
DMU 39	2.56	2.40	210	1.20	0.74	2.09	0.81	0.06	27
DMU 40	0.45	0.28	28	0.41	0.47	0.71	0.51	0.02	4
DMU 41	70.10	0.40	50	3.96	6.33	5.24	2.86	0.09	8
DMU 42	28.37	2.37	49	0.00	0.00	0.00	0.00	0.00	9
DMU 43	69.74	0.39	50	4.10	12.34	5.43	3.85	0.09	8

	CAPEX (US$M)	SR	Employees	ROM (Mt/yr)	Ash (%)	Moisture (%)	Cap (Mt/yr)	Export (Mt/yr)	EmployeesP
DMU 44	70.62	0.40	28	2.70	8.97	6.75	4.05	0.00	5
DMU 45	7.64	6.35	722	0.11	0.04	0.11	0.08	0.05	87
DMU 46	5.07	1.60	85	3.28	10.91	8.20	3.28	0.04	13
DMU 47	0.00	0.00	0	2.19	0.81	2.37	2.22	0.05	0
DMU 48	16.32	1.36	128	7.19	8.73	8.65	5.19	0.10	18
DMU 49	33.16	7.04	800	0.10	0.04	0.09	0.08	0.05	97
DMU 50	0.11	0.00	0	2.55	5.76	2.07	2.54	0.09	0
DMU 51	339.86	4.16	413	4.54	3.07	5.11	3.07	0.10	57
DMU 52	9.12	5.92	570	2.69	1.34	2.67	1.89	0.08	71
DMU 53	310.83	6.37	773	0.00	0.00	0.00	0.00	0.05	95
DMU 54	0.38	0.00	0	0.00	0.00	0.00	0.00	0.00	0
DMU 55	3.74	5.95	536	3.23	1.60	4.52	3.21	0.05	66
DMU 56	4.48	7.12	63	0.00	0.00	0.00	0.00	0.00	10
DMU 57	0.00	0.00	0	0.30	3.19	3.24	0.58	0.11	0
DMU 58	0.00	0.00	0	0.00	0.00	0.00	0.00	0.00	0

Appendix D

Training and test data sets for predictive modelling

Table D.1: Training data sets CSLE

Mine	Efficiency	CAPEX	Age	SR	Employees	ROM	Ash	Moisture	Cap	Export	CO2Emission	Dquantity	Dist.port	Precipitation	CV	Thickness	EmployeesP
DMU 1	0.97	2.00	3	13.20	208	1.24	5.50	15.50	2.40	0.58	0.03	0.30	517.00	2905	26.10	3.50	31
DMU 40	0.97	8.31	4	5.20	516	8.89	10.10	15.50	11.00	6.99	0.38	1.28	517.00	2121	26.10	10.00	71
DMU 18	0.99	15.88	17	10.30	992	11.07	5.50	15.50	11.00	9.90	0.47	1.28	98.00	2809	27.80	5.50	124
DMU 54	1.00	502.71	3	2.00	35	0.60	13.30	2.90	2.40	0.30	0.00	0.84	951.20	630	27.60	3.50	10
DMU 11	0.68	631.33	17	5.00	755	14.80	10.10	10.00	15.00	5.40	0.64	7.70	41.50	656	28.90	8.70	99
DMU 7	0.66	8.31	9	5.20	516	8.89	10.10	13.50	6.00	2.30	0.38	0.84	517.00	688	27.60	5.50	71
DMU 36	0.79	631.33	9	2.40	381	11.07	10.10	10.00	11.00	2.30	0.47	7.70	262.00	656	28.90	5.50	57
DMU 2	0.52	631.33	17	5.00	503	8.89	10.10	9.00	11.00	2.30	0.38	7.70	570.00	683	27.80	5.50	69
DMU 38	1.00	424.99	2	2.00	35	0.60	26.50	2.90	2.40	0.30	0.00	1.81	951.20	630	20.00	10.00	10
DMU 63	1.00	2.00	2	7.80	79	0.60	5.50	13.50	4.00	0.30	0.00	0.30	98.00	2121	27.60	3.50	15
DMU 48	0.46	62.18	2	5.20	489	8.31	10.10	10.00	6.00	0.58	0.35	1.28	275.00	683	27.60	3.50	67
DMU 31	1.00	631.33	17	2.00	435	14.80	25.00	2.90	11.00	6.99	0.64	7.70	98.00	630	27.80	15.00	64
DMU 41	0.81	424.99	8	2.40	304	8.31	13.30	11.00	6.00	2.30	0.35	1.81	98.00	683	27.60	8.70	47
DMU 46	0.48	15.88	3	5.00	266	4.00	13.30	10.00	4.00	0.58	0.15	1.28	517.00	683	25.80	8.70	40
DMU 37	0.57	15.88	3	10.30	173	1.24	5.50	11.00	2.40	0.58	0.03	0.30	98.00	2121	27.80	3.00	27
DMU 32	0.65	62.18	4	5.20	273	4.00	5.50	11.00	6.00	2.30	0.15	0.55	517.00	688	27.80	3.50	41
DMU 27	0.97	62.18	17	10.30	440	4.00	5.50	10.00	2.40	2.30	0.15	0.30	275.00	688	28.90	3.00	59
DMU 39	0.92	8.31	2	7.80	682	8.89	5.50	15.50	6.00	4.80	0.38	0.55	275.00	2809	26.10	5.00	89
DMU 52	0.85	15.88	17	10.30	992	11.07	5.50	11.00	7.80	5.40	0.47	0.84	517.00	2809	27.80	3.50	124
DMU 42	0.99	62.18	8	5.20	108	1.24	5.50	11.00	2.40	2.30	0.03	0.30	517.00	688	27.80	3.00	20
DMU 49	0.98	62.18	17	13.20	1500	14.80	5.50	13.50	11.00	9.90	0.64	1.28	41.50	2905	28.90	3.50	181
DMU 44	0.79	424.99	8	2.40	170	4.00	13.30	10.00	6.00	0.30	0.15	1.81	517.00	683	26.10	8.70	30
DMU 60	0.98	502.71	19	5.00	256	8.89	10.10	11.00	6.00	5.40	0.38	2.00	275.00	656	27.80	5.50	69
DMU 28	0.40	502.71	8	5.00	266	4.00	10.10	2.90	6.00	0.58	0.15	1.28	98.00	683	27.80	3.50	40
DMU 53	0.98	502.71	17	10.30	1250	14.80	5.50	11.00	11.00	9.90	0.64	1.28	41.50	2121	28.90	3.50	154
DMU 13	0.85	15.88	17	7.80	682	8.89	5.50	16.00	7.80	5.40	0.38	0.84	275.00	2121	27.60	5.00	89
DMU 8	0.97	15.88	3	13.20	999	8.89	5.50	15.50	7.80	6.99	0.38	0.55	41.50	2905	27.80	3.00	124
DMU 4	0.83	8.31	17	5.20	615	11.07	13.30	16.00	11.00	5.40	0.47	2.00	275.00	2121	25.80	15.00	82
DMU 20	1.00	62.18	8	5.20	775	14.80	10.10	11.00	11.00	9.90	0.64	7.70	275.00	688	27.60	10.00	101
DMU 34	0.91	631.33	3	2.40	320	8.89	13.30	2.90	7.80	4.80	0.38	7.70	98.00	630	27.80	5.50	49
DMU 16	1.00	8.31	2	2.40	38	0.60	25.00	9.00	2.40	0.30	0.00	0.84	951.20	683	20.00	10.00	10
DMU 33	0.83	8.31	19	7.80	1024	14.80	6.00	16.00	11.00	6.99	0.64	1.28	41.50	2905	27.60	10.00	129
DMU 14	0.97	62.18	4	10.30	440	4.00	5.50	11.00	2.40	2.30	0.15	0.30	275.00	2121	27.80	3.00	59
DMU 64	1.00	8.31	4	2.40	195	1.24	25.00	10.00	1.20	0.58	0.03	1.28	951.20	688	20.00	10.00	15
DMU 45	0.98	15.88	19	13.20	1500	14.80	5.50	15.50	11.00	9.90	0.64	1.81	262.00	2905	27.80	10.00	181

Mine	Efficiency	CAPEX	Age	SR	Employees	ROM	Ash	Moisture	Cap	Export	CO2Emission	Dquantity	Dist.port	Precipitation	CV	Thickness	EmployeesP
DMU 5	1.00	2.00	4	13.20	117	0.60	5.50	13.50	1.20	0.58	0.00	0.30	570.00	2905	27.60	3.00	19
DMU 59	0.98	631.33	8	2.00	497	11.07	26.50	9.00	15.00	4.80	0.47	7.70	262.00	630	26.10	15.00	53
DMU 50	0.98	502.71	8	2.00	345	11.07	25.00	9.00	11.00	4.80	0.47	7.70	98.00	656	25.80	15.00	53
DMU 10	0.96	502.71	3	2.00	435	14.80	26.50	11.00	15.00	4.80	0.64	7.70	275.00	683	25.80	15.00	64
DMU 57	0.99	2.00	2	2.40	67	1.24	13.30	13.50	2.40	0.58	0.03	0.84	570.00	688	20.00	10.00	15
DMU 12	0.77	631.33	19	7.80	1024	14.80	5.50	13.50	15.00	6.99	0.64	2.00	41.50	688	28.90	5.00	129
DMU 30	0.95	15.88	3	2.00	154	4.00	26.50	9.00	7.80	0.30	0.15	7.70	951.20	656	20.00	15.00	28
DMU 26	0.99	502.71	9	2.00	61	1.24	13.30	2.90	6.00	0.58	0.03	1.81	570.00	656	27.60	5.50	14
DMU 9	0.60	62.18	4	7.80	142	1.24	5.50	11.00	2.40	0.58	0.03	0.30	517.00	688	27.80	3.00	23
DMU 22	1.00	62.18	8	2.40	38	0.60	13.30	9.00	1.20	0.30	0.00	0.55	517.00	656	26.10	3.50	10
DMU 21	1.00	424.99	2	2.00	35	0.60	25.00	2.90	1.20	0.30	0.00	1.28	570.00	630	26.10	5.50	10
DMU 23	0.96	502.71	3	2.00	154	4.00	25.00	2.90	7.80	0.58	0.15	7.70	517.00	630	25.80	15.00	28
DMU 17	0.98	8.31	3	7.80	361	4.00	5.50	13.50	2.40	2.30	0.15	0.30	98.00	2121	27.60	3.50	51
DMU 56	0.99	8.31	3	13.20	117	0.60	5.50	11.00	1.20	0.58	0.00	0.30	262.00	2809	27.80	3.00	19
DMU 43	0.80	424.99	4	2.40	304	8.31	25.00	11.00	7.80	2.30	0.35	7.70	262.00	656	25.80	10.00	47
DMU 6	1.00	15.88	2	2.00	35	0.60	25.00	9.00	1.20	0.30	0.00	0.84	951.20	656	25.80	3.50	10
DMU 65	0.99	424.99	3	5.20	222	4.00	13.30	2.90	2.40	2.30	0.15	0.55	570.00	683	27.80	3.00	41

Table D.2: Test data sets CSLE

Mine	Efficiency	CAPEX	Age	SR	Employees	ROM	Ash	Moisture	Cap	Export	CO2Emission	Dquantity(%)	Dist.port (Km)	Precipitation	CV	Thickness	EmployeesP
DMU 3	0.91	424.99	9	2.40	320	8.89	13.30	9.00	7.80	4.80	0.38	2.00	570.00	683	27.60	8.70	49
DMU 15	0.97	2.00	17	7.80	646	8.31	10.10	15.50	4.00	2.30	0.35	0.55	262.00	2121	25.80	5.50	85
DMU 19	1.00	62.18	2	2.40	67	1.24	13.30	2.90	1.20	0.58	0.03	1.28	517.00	683	26.10	3.50	15
DMU 24	1.00	631.33	3	2.00	35	0.60	13.30	2.90	2.40	0.30	0.00	1.28	517.00	630	27.80	3.00	10
DMU 25	0.97	2.00	3	5.20	108	1.24	10.10	15.50	4.00	0.58	0.03	0.55	570.00	688	20.00	8.70	20
DMU 29	0.96	502.71	9	2.00	154	4.00	26.50	9.00	6.00	0.58	0.15	7.70	951.20	630	25.80	15.00	28
DMU 35	0.99	424.99	3	5.20	273	4.00	6.00	9.00	6.00	4.80	0.15	0.84	262.00	683	27.80	3.50	41
DMU 47	0.99	62.18	19	10.30	1250	14.80	5.50	16.00	15.00	9.90	0.64	7.70	41.50	2809	27.80	15.00	154
DMU 51	0.57	424.99	4	5.20	516	8.89	6.00	10.00	6.00	2.30	0.38	1.28	98.00	688	27.80	3.50	71
DMU 55	0.82	8.31	9	13.20	1190	11.07	5.50	15.50	11.00	6.99	0.47	0.84	517.00	2905	27.60	5.00	146
DMU 58	1.00	2.00	2	5.00	59	0.60	10.10	9.00	1.20	0.30	0.00	0.30	951.20	688	25.80	3.00	13
DMU 61	0.99	15.88	17	10.30	550	14.80	5.50	15.50	11.00	9.90	0.64	1.81	41.50	2809	25.80	8.70	154
DMU 62	0.97	62.18	9	13.20	564	8.31	6.00	16.00	7.80	6.99	0.35	0.84	517.00	2905	28.90	5.00	118