Assessing the Productivity of Selective Container Terminals in Africa Using Data Envelopment Analysis (DEA)

By

Barend Jacobus Mienie

Submitted in fulfilment of the requirements for the degree of Master of Commerce (Statistics: Research) in the Faculty of Science at Nelson Mandela Metropolitan University

April 2016

Supervisor: Prof. GD Sharp

Co-Supervisor: Mr. WJ Brettenny

Plagiarism Declaration

I, Barend Jacobus Mienie (student number: 214333027), hereby declare that the dissertation for the degree of MCom (Statistics: Research) to be awarded is my own work and that it has not previously been submitted for assessment or completion of any postgraduate qualification to another University or for another qualification.

Barend Jacobus Mienie

Acknowledgements

I gratefully thank:

- Clarissa Morton for her love and support through this busy project. She keeps me going when I can't anymore.
- My parents for all their support.
- My two supervisors for all their insight and help. Prof. Gary Sharp for his knowledge of people, statistics and research. Prof. Sharp gave me the opportunity to do my Masters when most wouldn't and for that I am immensely thankful. I thank Mr Warren Brettenny for his hours of coding help and support. He is a great academic and will achieve great things in the NMMU statistics department.
- The National Research Foundation (NRF) and the South African Statistical Association (SASA) for the "Statistics in Crisis" funding.

Abstract

Data envelopment analysis (DEA) is used to assess the efficiency of 15 container terminals in Africa. The models proposed by Charnes, Cooper and Rhodes (1978) and Banker, Charnes and Cooper (1984) are used to determine and rank the efficiencies of the container terminals for 2013 and 2014. The results show that selected South African container terminals can improve on their operations relative to some of their neighbours to the North. Bootstrapping methods are used to investigate and clarify the results. The Malmquist Productivity Index (MPI) model is used to track and explain changes in efficiency over the period of assessment.

Key words: Data envelopment analysis, efficiency, performance evaluation, shipping industry, bootstrapping.

Table of Contents

List of Abbreviations

PECT	Port Elizabeth Container Terminal
NCT	Ngqura Container Terminal
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
LP	Linear Programming
CCR	Charnes, Cooper and Rhodes
BCC	Banker, Charnes and Cooper
FDH	Full Disposal Hull
SFA	Stochastic Frontier Analysis
TFPC	Total Factor Productivity Change
MPI	Malmquist Productivity Index
TEU	20 foot container equivalent units
TEUBH	20 foot container equivalent units handled per berth hour
TH	Total number of containers, both 20 and 40 foot, handled per year
CONMIX	The mixture of 20 and 40 foot containers
BRLWT	The average delays in commencing stevedoring, calculated as the difference between the berth time and gross working time
TEUCH	The average quay crane productivity, represented by the number of containers lifted per quay crane hour
CRANE	The number of gantry cranes present at the port
FS	The frequency of ship calls (container ships only)
CH	The average government port charges per container
ECM	Efficiency Contribution Measure

List of Abbreviations

TE	Technical Efficiency
PTE	Pure Technical Efficiency
SE	Scale Efficiency
CRS	Constant Returns-to-Scale
VRS	Variable Returns-to-Scale
IRS	Increasing Returns-to-Scale
DRS	Decreasing Returns-to-Scale
NIRS	Non-Increasing Returns-to-Scale
MPSS	Most Productive Scale Size
TTEC	Total Technical Efficiency Change
TC	Technological Change
PTEC	Pure Technical Efficiency Change
SEC	Scale Efficiency Change
NIRSE	Non-Increasing Returns-to-Scale Efficiency

List of Tables

List of Figures

Chapter One: Introduction

1.1 Introduction

Efficiency methods are useful when trying to compare homogenous operations. Several methods to estimate efficiency are available to researchers, and in many cases these methods have been compared. This dissertation used one of these efficiency methods to evaluate the operations of a selection of container terminals on the African continent.

1.2 Problem Identification

The motivation for this study was based on discussions with shipping-line companies (van Tonder, 2014), and the apparent lack of available information with respect to the efficiencies of different ports. These discussions were initiated by my interest in the shipping industry and a desire to use the technical skills learnt over the years to contribute meaningfully to it.

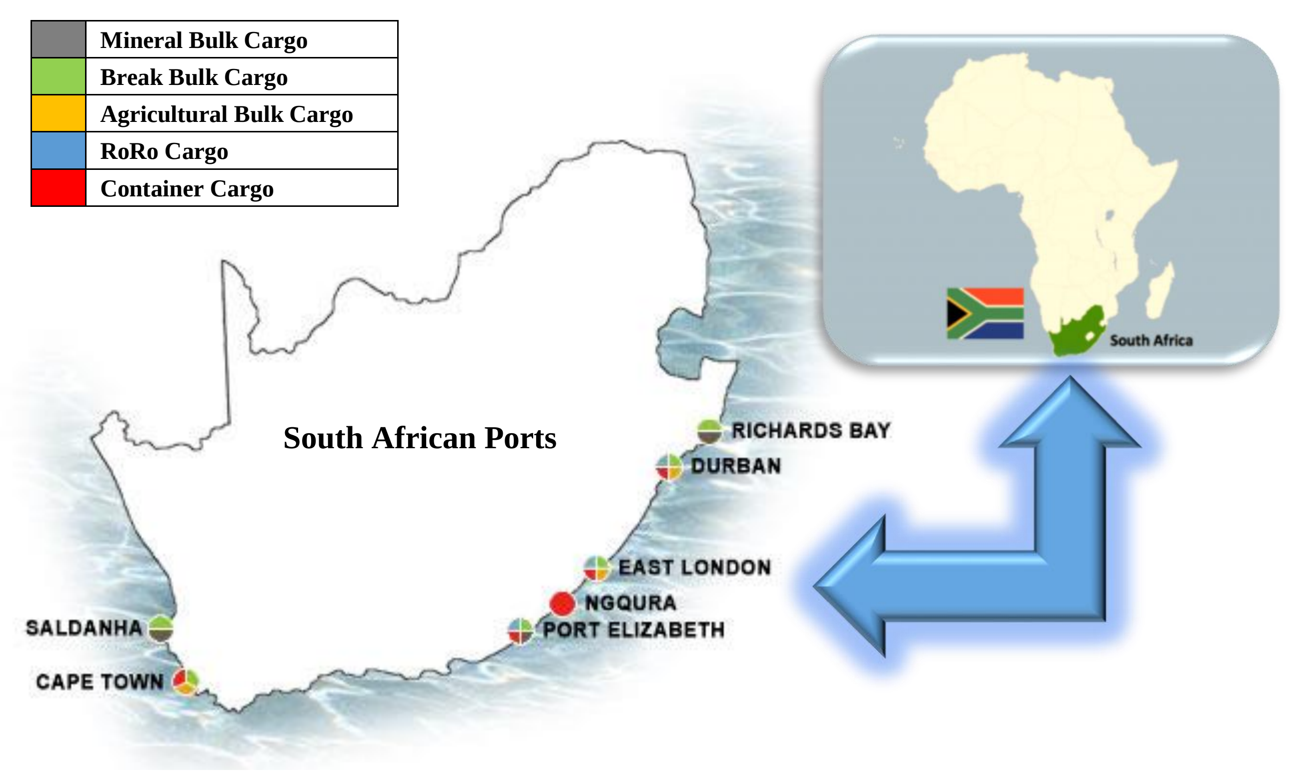
The efficiency of processes used within 15 chosen African containerised terminals was evaluated. These efficiencies were determined by benchmarking African container terminals against each other through the use of data envelopment analysis (DEA). DEA is a linear programming method used to calculate the relative efficiencies of a set of organisations which display homogenous functional traits, but whose efficiency may differ due to internal factors. Such organisations are commonly referred to as decision making units (DMUs) (Charnes, Cooper & Rhodes, 1978).

In addition, there has been very little documented research in the field of efficiency analysis in African container terminals. This lack of knowledge provided an opportunity for this dissertation to make a valuable contribution to the literature.

It was believed that efficiency measurements would allow the shipping-line companies to benchmark container ports against one another. This would enable the companies to make better use of the more productive African container terminals, thereby saving both their time and money. The measurements would also make port authorities aware of inefficiencies within their port processes. Any relevant improvements could then attract more container traffic from shipping-line companies.

Particular focus was given to the efficiency of the Ngqura Container Terminal (NCT) and the Port Elizabeth Container Terminal (PECT), both situated in the Eastern Cape of South Africa. The positioning of NCT and PECT are illustrated in Figure 1.1. This figure shows the positioning of South Africa within Africa as well as the ports that are located in South Africa. This figure also gives a breakdown of the different forms of cargo, in their respective proportions, handled by each South African port.

Figure 1.1: Positioning of NCT and PECT in South Africa



Source: (Transnet Port Terminals, 2013)

1.3 Objectives

The objectives of this study were as follows:

- To provide an overview of global research conducted on port efficiency.

- To use statistical methods to determine the level of efficiency in the provision of container terminal services in South Africa's NCT and PECT, as well as in other selected

African container terminals. DEA and extensions thereof were used to establish relative efficiency scores for all the selected container terminals.

- To use the efficiency results for the period 2013 – 2014 to comment on the efficiencies of the ports as well as any related trends over the sample period. The results were used to make recommendations as to how to improve any efficiency problems faced by the container terminals.

1.4 Structure of the Dissertation

This dissertation was structured as follows: Chapter 2 presents a review of the seaport literature concerning DEA, extensions of DEA, and other operation research techniques used in efficiency analysis. From this review the most suitable techniques were selected for this research. Chapter 3 identifies and justifies the variable selection, sample size and homogeneity of the sample, using the literature, industry objectives and selected statistical tests. Chapter 4 outlines the methodology for several efficiency models. Chapter 5 presents a validation of the code used to produce the DEA results. This was achieved by replicating the results of similar research. The results are presented and discussed in Chapter 6, with conclusions and recommendations in the final chapter, Chapter 7.

Chapter Two: Literature Review

2.1 Introduction

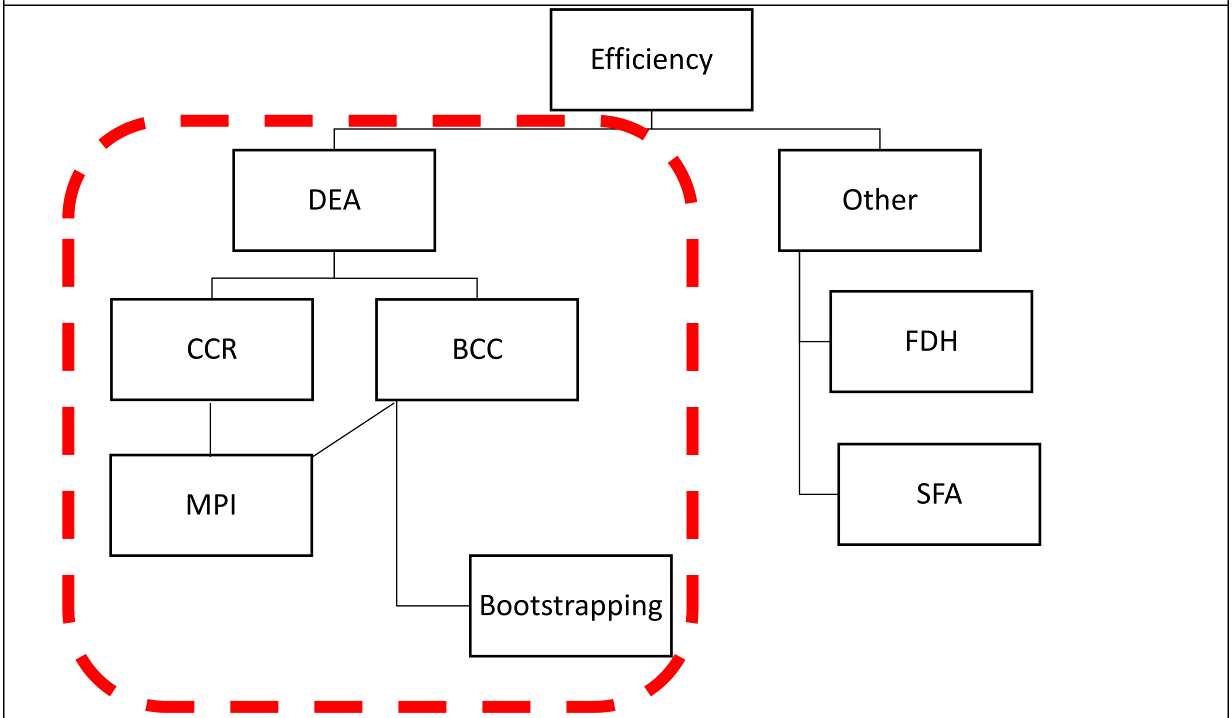
DEA is a mathematical technique which allows for the determination of efficiency measurements in an environment where input operations influence output operations. The volume of literature on DEA research has recently increased with the completion of many international and local studies. This increase can be attributed primarily to the methodological and computational benefits of the DEA technique (Panayides *et al.*, 2009).

The following literature review provides an introduction to the DEA techniques and studies undertaken in the seaport industry. The aim here was to identify the merits and limitations of the DEA method in aiding this study's investigation of container terminal efficiency.

2.2 A Review of the DEA Technique

A flow chart of the breakdown of efficiency analysis techniques is shown in Figure 2.1. The chart shows some of the techniques used directly in this research, as well as a few alternative methods. The DEA techniques that this study used to calculate efficiency are highlighted by a dashed red line.

Figure 2.1: Flowchart of Efficiency Analysis Techniques



DEA is part of a large family of frontier estimation procedures. De Borger, Kerstens and Costa (2002) classify this family of frontier estimation procedures by functional form and measurement error. The functional form relates to a procedure being classified as parametric or non-parametric. The parametric approach assumes that a particular functional form with constant parameters can represent the boundary of the production possibility set. The non-parametric approach imposes minimal regularity axioms on the production possibility set and directly imposes a piecewise approach on the sample. DEA is a non-parametric frontier estimation procedure. The measurement error relates to a procedure being classified as deterministic or stochastic. As a deterministic method, DEA takes all observations as given and implicitly assumes that these observations are exactly measured. Stochastic methods, in contrast, allow for random measurement error.

DEA, as a deterministic non-parametric technique, is used in operations research and econometrics for multivariate frontier estimation and ranking. The source of DEA may be traced to Farrell's 1957 study. These origins stem from a methodology of making evaluations from realised deviations from an idealised production frontier isoquant. Farrell (1957) introduced to this methodology an approach based on developing a piecewise linear, quasi-convex, envelopment of the data in order to determine the frontier. The frontier is then used to measure the relative efficiency. The efficiency values are calculated by comparing the relative performance of the organisation under investigation, to the organisation within the group with the best practice observed. The model produces measures of efficiency reflecting equi-proportional reductions of inputs or outputs onto the best practice frontier, the so-called radial efficiency measures (Farrell, 1957).

The technique referred to as DEA is concerned with the efficiency of individual organisations. The organisation of interest can be defined as the Unit of Assessment (Thanassoulis, 2001) or the Decision Making Unit (DMU) (Charnes, Cooper & Rhodes 1978). This unit is responsible for controlling the process of production and decision making at various levels. These levels include daily operations, short-term tactics and long-term strategies. DEA is best suited to measure the efficiency of multiple DMUs each of which contain several inputs and outputs (Panayides *et al.*, 2009).

DEA efficiency can be classified into two categories. The first of these is called the technical efficiency and is defined as the relative productivity over time, space, or both. The second is the scale efficiency and relates to a possible divergence between the actual and ideal production size (Munisamy & Danxia, 2013; Panayides *et al.*, 2009; Wang, Cullinane & Song, 2005).

The work of Farrell (1957) was expanded by Charnes, Cooper and Rhodes (1978), who introduced a linear programming (LP) methodology, which in turn lead to the DEA Charnes, Cooper and Rhodes (CCR) model. The CCR model is applied to situations in which constant returns-to-scale are applicable. The efficiency generated by the CCR model is a technical efficiency which has both a scale component and a pure technical component, driving the efficiency score. The pure technical efficiency measure is determined by comparing inefficient DMUs to efficient DMUs of the same scale size. In contrast, the technical efficiency measure is determined by comparing each inefficient DMU to efficient DMUs, irrespective of scale size. Therefore, the scale efficiency is the ratio of the technical efficiency and pure technical efficiency.

The CCR model was followed by the introduction of the DEA Banker, Charnes and Cooper (BCC) model by Banker, Charnes and Cooper (1984). The BCC model is applied to situations in which variable returns-to-scale are applicable (Panayides *et al.*, 2009). The efficiency generated by the BCC model is a pure technical efficiency. The difference between the CCR and BCC models is that, while the former provides information on technical and scale efficiency, the latter identifies pure technical efficiency alone. If both models are applied, then pure technical and scale efficiency can be calculated as separate values.

If both the CCR and BCC models have the same efficient DMUs, then DEA super-efficiency and cross-efficiency models are possible. A super-efficiency model, introduced by Andersen and Petersen (1993), enabled researchers to distinguish between units rated as efficient, both within and between, the CCR and BCC models. An alternative to the super-efficiency evaluation is the cross-efficiency evaluation. The cross-efficiency model was pioneered by Doyle and Green in 1994. The model can be used to eliminate unrealistic weight schemes of classical DEA models and to provide further discrimination among efficient DMUs within and between CCR and BCC models. This approach allows for the ranking of the DMUs within each model.

The Free Disposal Hull (FDH) model is different from the CCR and BCC models as it does not operate with a convexity assumption. The FDH model has a discrete nature whereby the efficient reference point for an inefficient DMU is not chosen as a point on a continuous efficiency frontier, but among the existing DMUs (Pachkova, 2005). The results from applying the FDH model, therefore, may be more convincing in practice as counterpart DMUs identified as efficient actually do exist in every case. By the very nature of its underlying logic and step function solution algorithm, however, the FDH model is not very sensitive to comparatively small differences in efficiency. These differences can be better identified by the application of the CCR and/or BCC models (Cullinane, Song & Wang, 2005).

DEA determines efficiency by radially comparing DMUs to the production frontier. The production frontier consists of fully efficient DMUs. Inefficient DMUs are enveloped by the frontier. To correct for inefficiency, the inefficient DMU's are projected to the production frontier. There are three orientations in which such a projection can take place. One of these orientations is called the input orientation and aims to reduce input amounts by as much as possible, whilst keeping the output levels fixed, in order to achieve efficiency. The second orientation is called the output-orientation and maximises output levels under the present input levels to achieve efficiency (Cooper, Seiford & Tone, 2006). The difference between these two orientations lies in how the variables are adjusted to achieve the projection path to the frontier. In the simplest single input-output situation, the input-oriented model's projection path is horizontal, whilst the output-oriented model's projection path is vertical (Panayides *et al.*, 2009). Either the input-orientation or the output-orientation can be used to correct for inefficiency in the CCR and BCC models. The third and final orientation for correcting for inefficiency is used in the Additive model. This model is an alternative DEA model that adjusts input and output levels simultaneously to achieve efficiency (Cooper, Seiford & Tone, 2006). In the simplest single input-output case, the Additive model's projection path is diagonal (Panayides *et al.*, 2009). The BCC and Additive models are identical in terms of their production frontiers. The difference being the different projection paths to the production frontier for the inefficient DMUs. The orientation of a DEA model mainly depends on the nature of the production and the given constraints. The orientation of the selected DEA models is discussed in more detail in the Methodology Section (Chapter 4).

As a deterministic method, DEA does not explicitly model a random error term and the overall deviation from the frontier is interpreted only as inefficiency. However, the DEA method does use a sample for the analysis of efficiency. Thus, differences in estimations may be due to sampling error rather than actual differences in the efficiency levels of the respective DMUs. To overcome this limitation, a bootstrap methodology has been proposed to evaluate the sampling variability of DEA results (Hung, Lu & Wang, 2010). Bootstrapping in this case, is based on the belief that resampling from the original data, creates replicate datasets from which sampling error can be identified and corrected for in the efficiency results (Martınez-Nunez & Perez-Aguiar, 2014).

When analysing cross-sectional data, DEA involves the comparison of one DMU with other DMUs sampled during the same time period. Panel data not only enables a DMU to be compared with other counterparts, but to also assess changes in the efficiency level over a period of time. In so doing, panel data reflect the pattern of efficiency of a DMU, and as such, are often preferred to cross-sectional data if available (Cullinane & Wang, 2010).

When panel data is used, changes in efficiency are compared using the Malmquist Productivity Index (MPI) proposed by Färe *et al.,* (1994). This index produces an efficiency change measure referred to as the Total Factor Productivity Change (TFPC). The TFPC provides an interpretation of the change in efficiency over time and can be divided into three components. These component measurements are; the changes in pure technical efficiency, the changes in scale efficiency and, the final component measures changes in technology. The change in technology is obtained by measuring the shift in the frontier produced by the DEA models from one period to the next. This development has allowed researchers to use DEA techniques in combination with the MPI (Panayides *et al.,* 2009).

One technique that has proven to be a popular alternative to DEA for determining the efficiencies of DMUs, is Stochastic Frontier Analysis (SFA). SFA was introduced simultaneously by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). SFA assumes that a parametric function exists between production inputs and outputs. The analysis not only allows for the calculation of technical efficiency, but also acknowledges the fact that random shocks outside the control of producers can affect output. In SFA, these random shocks are accounted for in an error term composed of two parts. The first is a one-sided component that captures the effects of inefficiency relative to the stochastic frontier. The second is a symmetric component that permits random variation of the frontier across firms, and also captures the effects of measurement error, other statistical noise, and random shocks outside the firms' control (Cullinane *et al.,* 2006).

2.3 DEA Applications to Seaport Efficiency Measurement

This section introduces the chronological sequence of research into seaport efficiency developments and provides the platform for the methodology adopted in this study.

DEA has been used extensively to measure container terminal efficiency. Tables A and B in Appendix One provide a summary of important DEA based studies completed in the last two decades. These tabular summaries list the variables and data sets utilised in each of the 18 studies. Seventeen of the studies were based outside Africa (see Table A). In contrast, only one study to date was conducted on the African continent (see Table B). No study using South African data could be sourced from the literature. Three additional studies, listed in Table B, are included as they are port efficiency studies in Sub-Saharan Africa using methods alternative to those used in this study.

Roll and Hayuth (1993) pioneered the use of DEA to measure port efficiency in the 1990's. Using the CCR model, their analysis evaluated 20 DMUs (ports). The outputs they chose for analysis included: container throughput (including container, general cargo and bulk cargo), the level of service (ratio between handling time and the total time a ship stays in port), users' satisfaction (a score on a linear scale from 1 to 10, as obtained from a satisfaction questionnaire), and ship calls (the number of ship visits to port per year). The inputs selected for analysis included: manpower (the annual average number of labourers (un)loading cargo), capital (total invested capital during the year in the port and all its related facilities), and cargo uniformity (coefficient of variation of the types of cargo). A summary of their results is listed in Appendix One). Roll and Hayuth's application did not use actual data, rather it was the first theoretical attempt to apply the DEA technique to measure port efficiency. The researchers proposed that the derivation of efficiency ratings should be a regular activity for port operators, as it is a useful tool for management control. The researchers concluded that DEA is a promising and easily adaptable approach for obtaining efficiency ratings.

Poitras, Tongzon and Li (1996) used DEA to measure the relative efficiency of 23 Australian and international ports for the year 1991. The empirical results used two output measures; the number of 20 foot container equivalent units (TEUs) handled per berth hour (TEUBH), and the total number of containers, both 20 and 40 foot, handled per year, (TH). The 20 and 40 foot containers were treated equivalently. The input measures used were; the mixture of 20 and 40 foot containers (CONMIX), the average delays in commencing stevedoring, calculated as the difference between the berth time and gross working time (BRLWT), the average quay crane productivity, represented by the number of containers lifted per quay crane hour (TEUCH), the number of gantry cranes present at the port (CRANE), the frequency of ship calls (container ships only) (FS), and the average government port charges per container (CH). These variables are summarised in Appendix One. The researchers applied both the CCR and Additive models to the 1991 data and found that the CCR model identifies more inefficient ports (13) than the Additive model (four). This is due to the CCR model having stricter relative efficiency criteria than the Additive model. The primary contribution of their study is in its methodological developments, such that DEA provides a viable method for evaluating port efficiencies (Panayides *et al.*, 2009).

Martinez-Budria *et al.,* (1999) applied the BCC model to measure the efficiency of 26 Spanish ports. The efficiency for each port was determined yearly during the five year sample period (1993-1997). Their model included two outputs, related to throughput and revenue, while inputs were financially related. The significance of this work is that they introduce two new elements in the application of DEA for the measurement of port efficiency. The first is the use of panel data (between 1993 and 1997) and the second is that the researchers recognise that the 23 ports have major differences in terms of the complexity associated with port size. They separated the 23 ports into three different categories based on complexity level (high, medium, low) and calculated efficiency values for each of these groups. The study concluded that the more complicated ports show higher comparative efficiency levels with a positive trend over time. The medium complexity group illustrate smaller growth in efficiency levels over the five year period, whereas the ports with the lowest complexity show a negative evolution in terms of relative efficiency levels (Panayides *et al.,* 2009).

Tongzon (2001) used DEA to assess the relative efficiency of selected major Australian ports and their international counterparts in the year 1996. The data used was cross-sectional and the selected outputs included throughputs and ship working rate. The chosen inputs were capital, labour, land and delay time. Both the CCR and Additive model were applied, with CCR having identified slightly more inefficient ports. These findings corroborate research by Poitras, Tongzon and Li (1996). It is recognised that the small sample size (16 observations) resulted in more efficient, rather than inefficient ports, leading Tongzon (2001) to recommend a larger sample in order to minimise observational biases.

Valentine and Gray (2001; 2002) attempted to establish whether port performance and ownership structure were related. The researchers did this through the use of DEA. In their 2001 study, 21 container ports sampled were retrieved from the Cargo Systems Journal 1999 list of top 100 container ports. Data was cross-sectional. The outputs used were the number of containers and total throughput in tons. Inputs used were the value of the port assets in US dollars and quayage in metres. The DEA model used was the CCR model. The researchers used public, private and public/private ownership models and simple, divisional and bureaucratic port characteristics, to construct nine categories (Panayides *et al.*, 2009). The study found the most efficient ownership structure to be joint public/private at an average efficiency of 58.5%. This was followed by private ports at 56.78% efficiency, and lastly publicly owned ports at 51.26% efficiency. In 2002, Valentine and Gray carried out a similar study to their 2001 one, with a sample of 19 ports in North America and Europe for the year 1998. The outputs used were number of containers and throughput in tons. The inputs used were total length of berth and container berth length. The DEA model implemented was the CCR model, with cross-sectional data. The researchers found that both geographical regions show similar average efficiencies (Panayides *et al.*, 2009).

Barros (2003) analysed 11 seaports located in a wide geographical area of Portugal. MPI was used to determine the TFPC between 1990 and 2000. Outputs used were number of ships and freight and inputs applied were related to capital and labour. Barros (2003) found that none of the 11 authorities achieved total factor productivity improvements within this period. All ports achieved improvements in technical efficiency but not technological change. The researcher acknowledged the need to benchmark the Portuguese ports with other European ports in order to have a broader perspective of their efficiency (Panayides *et al.*, 2009).

Barros and Athanassiou (2004) used DEA to measure the relative efficiency of two Greek and four Portuguese ports. The study utilised panel data between 1998 and 2000. Outputs included movement of freight, total cargo handled, and containers loaded and unloaded.

Inputs included labour and capital. The values of the outputs and inputs were averaged over the three years for the CCR and BCC efficiencies to be calculated. The study's major finding is that more than half of the selected ports operate at a high level of pure technical efficiency. Barros and Athanassiou (2004) recognised that the dataset was small (the number of DMUs was only six), which could explain why so many ports are purely technically efficient (Panayides *et al.,* 2009). Similar observational bias was observed in the Tongzon (2001) study.

Estache, De La Fé and Trujillo (2004) relied on an MPI to calculate and identify changes in productivity for Mexico's 11 main ports between 1996 and 1999. This is similar to the approach applied by Barros (2003). Merchandise in tons is used as the single output. Labour and capital are used as inputs. The results indicate that TFPC in Mexican ports rose on average by 4.1% per year during 1996–1999. On a year to year basis, this 4.1% average was driven by the first three years of the period immediately after port reforms were initiated. During the last (fourth) year, there was a generalised technological regression. This was an expected result since world trade shrank due to the East Asia crisis—leading to a lack of traffic through the ports and thus a decrease in scale efficiency.

The Cullinane, Song and Wang (2005) study contributed to the extant research in that two non-parametric approaches, the DEA and the FDH model, were evaluated simultaneously in the container terminal industry for the first time. A sample of 57 container ports and terminals was studied during 2001. The output used was container throughput and the inputs consisted of both capital and land factors. Analysis of the efficiency estimates yielded by the two DEA models (CCR and BCC) and the FDH model confirm that the DEA and FDH methodologies tend to give different results. The study found that the results from the FDH model can only identify real life efficient benchmark counterpart(s) for the inefficient DMUs to learn from. These real life efficient benchmarks do not always represent a complete efficient version of the inefficient DMU that the inefficient DMU can learn from. These benchmarks are the closest efficient version of the inefficient DMU. This is because the efficient benchmark will simply be one of the already existing operating DMUs, whereas in DEA, the perfect efficient benchmarks are constructed by weighting fully efficient ports. This produces fully accurate efficient benchmarks for the inefficient DMUs to learn from.

Barros (2006) applied DEA to evaluate the performances of 24 Italian seaports, using data between 2002 and 2003. In the models, operational and financial variables are combined and averaged over the two year period. These values are used to calculate the BCC and CCR efficiencies. In order to discriminate among efficient ports, Barros (2006) used the cross-efficiency and the super-efficiency models, concluding that large ports tend to have higher efficiency scores. This supports Martinez-Budria *et al.,* (1999). It was also reported that containerised ports tend to have higher efficiency scores than less containerised ports. In addition, the ports with a smaller employee to sales ratio are more efficient than those with a higher employee to sales ratio (Panayides *et al.,* 2009).

Cullinane *et al.,* (2006) undertook an empirical study with the aim of comparing DEA and SFA. A 2001 sample of 57 container ports and terminals was used with container throughput as the output and five inputs related to land and equipment. The study found similar estimates of efficiency in terms of the ranking of the ports for the two approaches. Another significant outcome is that the majority of large ports (arbitrarily identified as those with more than one million annual TEU container throughputs) are found to be scale inefficient (Panayides *et al.,* 2009). This contradicted the findings of other similar studies (Barros, 2006; Martinez-Budria *et al.,* 1999) that larger ports are more efficient.

Rios and Maçada (2006) applied DEA to assess and rank the efficiencies of container terminals of the Mercosur (comprising Argentina, Brazil, Paraguay, Uruguay and Venezuela) between 2002 and 2004, using the BCC model. Inputs consisted of land and capital factors. Outputs were the number of containers moved and the rate at which they were moved per hour per ship. The variables were recorded for each year and the efficiencies calculated using the BCC model. Results indicate that 75% of the terminals studied are 100% efficient in 2002. This figure dropped in subsequent years, reaching only 65% in 2004. The researchers conclude that the terminals deemed efficient be considered as benchmarks, and the port managers should take reference of the practices used in the efficient terminals to improve operations (Panayides *et al.,* 2009).

Wang and Cullinane (2006) used DEA to determine the relative efficiency of Europe's leading container terminals. Data for the year 2002 was used, consisting of 69 leading container terminals throughout 24 European countries. The single output used was container throughput and the inputs used were land and capital factors. The data collected was used to estimate individual efficiency scores for each port/terminal. The primary finding of this paper is the significant inefficiency that generally pervades many of the terminals studied. The average efficiency of operations at the container terminals in the study amounted to 42% (assuming constant returns-to-scale or CCR) and 48% (assuming variable returns-to-scale or BCC). Given the large sample used, the efficiency estimates are likely to be more consistent and robust than the results in other studies (Barros & Athanassiou, 2004; Tongzon, 2001). The study reported that large production scale is associated with higher efficiency scores. This is similar to the findings of Barros (2006) and Martinez-Budria *et al.,* (1999).

Munisamy and Danxia (2013) used the smooth homogenous bootstrapped frontier to obtain bias free efficiency estimates of 69 major Asian container ports in 2007. The output used was total throughput in TEUs. The inputs included in the analysis were; berth length (in metres), terminal area (in metres squared), total reefer points (number of points where refrigerated containers can source power), total quayside cranes, and total yard equipment. The BCC model was used to determine the efficiencies. Once bias in the efficiency scores was addressed using the bootstrapping procedure, the ports were ranked in descending order. Munisamy and Danxia (2013) found that efficiency can be improved on average by 37% through the expansion of outputs, while controlling for inputs, in these ports.

The study by Herrera and Pang (2008) determined the efficiency of container ports for the years 2000-2001. The FDH, CCR and BCC models were used to determine the efficiency of 86 ports. Although the models applied resemble those used in Cullinane, Song and Wang (2005), they are only used to determine efficiencies. The study did not compare the FDH and DEA procedures as in the Cullinane, Song and Wang (2005) study. The output used was container throughput. The inputs were land and capital factors. Output and input measures were averaged over the sample period. Results show that the most inefficient ports use inputs in excess of 20 to 40 percent. The results found that privately owned ports are more efficient than those publicly owned. This is similar to the findings of Valentine and Gray (2001). It was also identified that ports in similar geographical regions have similar efficiencies. This corroborates the findings of Valentine and Gray (2002). The researchers also found that larger ports are more efficient, strengthening the findings of Barros (2006), Martinez-Budria *et al.,* (1999), and Wang and Cullinane (2006), but contradicting those of Cullinane *et al.,* (2006). The results showed that scale inefficiency can be remedied by increasing or decreasing the scale of production.

The study by de Oliveira and Cariou (2011) used the CCR and BCC models to assess the efficiency of 122 iron ore and coal ports in 2005. The output used was throughput in tons. The inputs selected were; draft (in metres), berth length (in metres), stockpile capacity (in tons), and (un)loading rates (metric-tons/hour). Efficiency estimates for 54 loading and 68 unloading ports showed that the main source of inefficiency in bulk terminals is related to the scale. This is similar to the finding of Estache, De La Fé and Trujillo (2004) in the fourth and final year of their analysis. The study also found differences between loading and discharging ports.

Limited literature exists on the application of DEA in an African port context. The most recent of these limited applications was conducted by Al-Eraqi *et al.*, (2007). The study determined the efficiency of 22 seaports in Africa and the Middle East. Data was collected during six years (2000-2005) and CCR and BCC models applied. The aim of the study was to compare seaports situated on the maritime trade route between the East and the West. The output used was cargo throughput. The inputs used were; berth length (in metres), distance (in nautical miles), and terminal area (in metres squared). The output and input values used were averaged over the six years to calculate one efficiency value for each port for the sample period. The results showed that the BCC model has more efficient ports than CCR model. The average values were 77% and 69%, respectively, similar to results generated by Poitras, Tongzon and Li (1996). The inefficiency for CCR and BCC models is due to a decline in the numbers of ship calls. Researchers suggest that public and private sector investment can improve the efficiency of the inefficient ports in the region through development and expansion.

2.4 Justifying Selected DEA Techniques

The following section provides a justification for the use of DEA techniques to determine the efficiency scores of the selected African container terminals.

2.4.1 DEA Models, Orientations and Data Sets

CCR and BCC models were used in this study to determine the efficiencies of the selected African container terminals. This decision was based on the ability of the CCR and BCC models to account for constant and variable returns-to-scale. This enabled the author to calculate technical, pure technical and scale efficiency, which provided a thorough overview of efficiency in the ports. Secondary reasons for choosing the CCR and BCC models were the high frequency of their use within the literature. Researchers that have used these models include; Al-Eraqi *et al.*, (2007), Barros (2006), Barros and Athanassiou (2004), Cullinane, Song and Wang (2005), de Oliveira and Cariou (2011), Herrera and Pang (2008) , and Wang and Cullinane (2006).

A sample of DMUs was selected from a wider population of DMUs. Thus, sample bias needs to be accounted for. The bootstrap methodology was used to investigate the sampling variability present within DEA. This study used the Simar and Wilson (2000) method of homogenous bootstrapping to extend the DEA models in order to correct for sampling error. Sampling error was corrected for within the CCR and BCC estimates of efficiency. The removal of sampling error from the efficiency estimates provided a method of distinguishing between fully efficient DMUs. This enabled this study to rank the DMUs. Thus, bootstrapping provided an alternative to the super-efficiency or cross-efficiency methods. Munisamy and Danxia (2013) also applied bootstrapping procedures and successfully identified biases present within efficiency results.

Once efficiencies were calculated using the CCR and BCC models, the MPI was used to track the movements in technical, pure technical, scale and technological efficiency over the sample period. These efficiencies were all sub-components of the TFPC calculated by the MPI. A unique benefit of the MPI was that it accounted for the change in technology in addition to the changes in technical, pure technical and scale efficiency. The change in technology was a result of the shift in the frontier from one sample period to the next. Researchers that have used this technique include Barros (2003) and Estache, De La Fé and Trujillo (2004).

Many researchers have measured the efficiency in each period correctly using CCR and/or BCC models (Martinez-Budria *et al.*, 1999 and Rios & Macada 2006). However, they tracked efficiency changes from one period to the next incorrectly by comparing the DEA efficiency in one period to the efficiency in the next. This does not account for the change in the sample or technology, ignoring the shift in the frontier from one period to the next. Researchers such as Al-Eraqi *et al.*, (2007), Barros (2006), Barros and Anthanassiou (2004), and Herrera and Pang (2008) used panel data but did not track the efficiency over time. Instead these authors averaged the variable values over the sample period to calculate an average efficiency for the sample period. The MPI used in this study was believed to be the best technique for tracking efficiency over time as it considered the shift in technology.

The input-orientation and the output-orientation were used in equal frequencies throughout the literature surveyed. Wang, Song and Cullinane (2002) mentioned that input-oriented models are closely related to operational and managerial issues. However, output-oriented models are more associated with strategy development and evaluation. Given that this dissertation was concerned with strategy development and evaluation of operations in selected African container terminals, the output-orientated approach was selected for the DEA models used.

This study used panel data consisting of 15 selected African container terminals during the period 2013-2014. Panel data were considered more appropriate to facilitate the measurement of efficiency over time. The main implication of cross-sectional data was that one observed efficiency at a certain point in time and not over one time period (Panayides *et al.,* 2009).

2.4.2 Combinations of Operations Research Techniques

No other studies, except for those conducted by Cullinane, Song and Wang (2005), Cullinane *et al*., (2006) and Herrera and Pang (2008), calculated efficiency estimates using DEA and other operations research techniques. Cullinane, Song and Wang (2005) and Herrera and Pang (2008) used DEA and FDH. Cullinane *et al*., (2006) used DEA and SFA. This study produced efficiency estimates using only DEA. Alternative techniques such as FDH and SFA were not considered in this study and are areas for further research and investigation.

2.5 Conclusion

A review of the DEA techniques, as well as 20 years' worth of DEA applications in the seaport industry, highlighted which DEA techniques are best suited to this study. These techniques will be used to calculate the efficiencies of 15 selected African container terminals. Output-orientated CCR and BCC models will be used to calculate the technical, pure technical and scale efficiencies. This decision was based on the fact that the models together provided a thorough overview of port efficiency. Sampling bias within the CCR and BCC efficiencies will be corrected for using the Simar and Wilson (2000) method of homogenous bootstrapping. The MPI will be used to track changes in the CCR and BCC efficiencies and determine not only differences in the technical, pure technical and scale efficiencies, but also changes in technology.

Chapter Three: Data

3.1 Introduction

The selected data was obtained from 15 major container terminals on the African continent. The data was classified as panel data as it was acquired for the years 2013 and 2014. The information was obtained from multiple sources online, with no single source providing the majority of the data. The sources used for the data are listed in Table A and B within Appendix One as well as within the data references section contained in the references.

Homogeneity within the data was necessary for DEA efficiency scores. This requirement was considered by looking at the geographical association between container terminals and the nature of the goods moving through them (Panayides *et al.*, 2009).

A pool of potential input and output variables was selected by considering the objectives of the container terminals in addition to the variables used in the literature. Thereafter, through the use of statistical techniques, a final set of input and output variables was selected from the pool of potential variables.

After the final variables were selected, minimum sample size rules within DEA were considered to ensure that discriminatory power existed when calculating efficiencies (Sakris, 2002).

3.2 Homogeneity of the Sample Data

An important issue in the application of DEA for container terminal efficiency measurement was the choice of the terminals. The rationale for this choice hinged on the principle of competition, as ranking of relative efficiency was more meaningful between competing ports. One had to consider the factors which created a general competitive environment between ports. These included the geographical location of the port and the nature of the goods moving through the terminals (Goss, 1990).

The African container terminals were compared within countries, particularly South Africa, Egypt and Morocco, as well as between countries. Geographical location and the nature of goods needed to be similar for competition to exist, were it within different or identical countries (Panayides *et al.*, 2009).

The African container terminals selected as DMUs are listed below in Table 3.1.

Table 3.1: African Container Terminals Selected as DMUs

Container Terminal	African Country
Alexandria International Container Terminal	Egypt
Cape Town Container Terminal	South Africa
Casablanca Container Terminal	Morocco
Damietta Container Terminal	Egypt
Tanzania International Container Terminal Services (Dar es Salaam)	Tanzania
Doraleh Container Terminal	Djibouti
Durban (Pier 1 and Pier 2)	South Africa
Apapa Container Terminal, Lagos	Nigeria
Luanda Container Terminal (CT2)	Angola
Mombasa Container Terminal	Kenya
Ngqura Container Terminal	South Africa
Port Elizabeth Container Terminal	South Africa
Suez Canal Container Terminal (Port Said)	Egypt
Tanger Med (Terminal 1 and Terminal 2)	Morocco
Tema Port Container Terminal	Ghana

Source: See subsection of “References” entitled “Data References”.

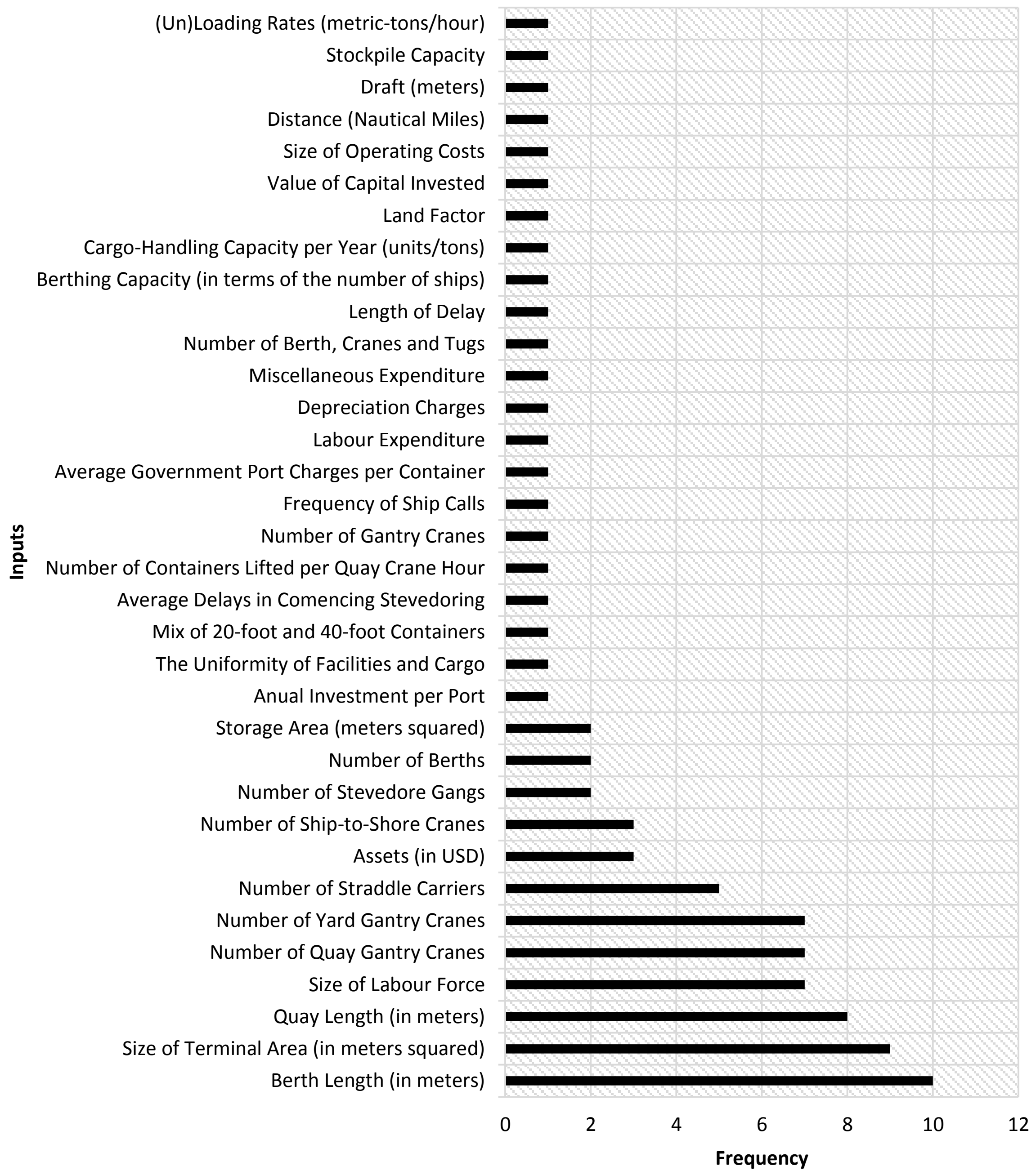
The homogeneity of the above 15 container terminals was believed to be strong as all were located in one geographical region, namely Africa. The nature of goods travelling through the terminals was similar i.e. containers of 20 foot equivalent units (TEU). Some may argue that the geographical region was too large to substantiate competition. It should be noted, however, that container ports find themselves competing more intensively against ports thousands of miles away, in addition to the severe competition experienced from nearby rivals (Talley, 2000). This long distance competition exists as a result of globalisation. It is argued here that due to globalisation it was reasonable to accept that competition, and thus homogeneity, existed between the selected African container terminals.

3.3 Variable Selection

To determine the most frequently used inputs and outputs in the container terminal DEA literature, several papers in the literature were reviewed. The majority of these papers were discussed in Chapter 2.

Focusing particularly on the input variables, Figure 3.1 illustrates the frequency with which the input variables were used in the reviewed papers.

Figure 3.1: Input Frequencies in Literature Survey



It was observed that the most frequently occurring input variables were, in order:

1. Berth Length (in metres) occurring 10 times.
2. Size of Terminal Area (in metres squared) occurring nine times.
3. Quay Length (in metres) occurring eight times.
4. Number of Yard Gantry Cranes/Number of Quay Gantry Cranes/Size of the Labour Force, occurring seven times.
5. Number of Straddle Carriers occurring five times.
6. Assets (in USD)/Number of Ship-to-Shore Cranes occurring three times.

These high frequency inputs above were then cross-referenced with industry objectives, in order to establish which of the inputs should form the basis of the study's input variable pool. During the process of data collection, multiple meetings were held with international shipping-line companies. The purpose of these meetings was to establish which variables were important to determine the efficiency of a container terminal. The discussions revealed that the variables could be categorised into four sections throughout each container terminal. These sections were defined as follows:

<u>Section 1:</u> the size of the Quay i.e. the berthing capacity provided to ships.

<u>Section 2:</u> the equipment available on the Quay to aid the loading and unloading of container vessels.

<u>Section 3:</u> the equipment available in the yard just behind the terminal which is used to store and move containers around in the yard.

<u>Section 4:</u> the labour present in all of Sections 1, 2 and 3.

The information provided by the companies coincided with the inputs used in the literature. The inputs "Berth Length" and "Quay Length" were included in Section 1. The inputs "Number of Ship-to-Shore Cranes" and "Number of Quay Gantry Cranes" were added to Section 2 and Section 3 incorporated the inputs "Number of Yard Gantry Cranes" and "Number of Straddle Carriers". Finally, the input "Size of the Labour Force" was included into Section 4 identified by the shipping-line companies. The inputs "Size of Terminal Area" and "Assets (US$)" could not be integrated into any of the sections and were thus ignored in this study.

These four categories, identified through the meetings with the companies, encompassed all of the most frequently used inputs in the literature. These inputs linked the objectives of academia and industry thereby forming the pool of variables from which the final inputs, to be used in the DEA models, were selected.

When this pool of inputs was related back to the inputs used by the 15 African container terminals, a number of issues became apparent.

The first of these was that information regarding the “Size of the Labour Force” in South African ports was not available to the public. Transnet would not release these statistics. As the labour force could not be determined in South Africa, this study did not pursue acquiring this input in other African countries. Even if this input was acquired for these countries, it would be ignored, as the input would not be present in every DMU. Therefore, this dissertation ignored the labour section, resulting in the pool of input variables becoming smaller by one. The lack of availability of the labour input was found to be a common issue throughout the literature. The exclusion of this variable would therefore not detract from this study’s contribution to the literature.

The second issue was that additional inputs were present in the selected DMUs that were not contained within the pool of inputs. These inputs could not be ignored as they fell into the sections identified as important by the shipping-line companies. One of these sections was Section 2. The additional inputs were “Mobile Cranes” and “Rail Transfer Cranes”. The second section was Section 3. The additional inputs to this section were “Rubber-Tyre Gantries”, “Empty Handlers” and “Reach Stackers”.

These additional inputs could not be added directly to the pool of inputs as they did not appear consistently in each DMU. To solve this problem, a general input was defined in both Section 2 and Section 3. In Section 2 this general input was called “Number of Terminal Cranes”. This input was equivalent to the sum of; “Mobile Cranes”, “Rail Transfer Cranes”, “Ship-to-Shore (STS) Cranes”, and “Quay Gantry Cranes”. In Section 3 the general input was called “Number of Operating Yard Equipment”. This input was equivalent to the sum of; “Rubber-Tyre Gantries”, “Empty Handlers”, “Reach Stackers”, “Straddle Carriers”, and “Yard Gantry Cranes”. “STS Cranes”, “Quay Gantry Cranes”, “Straddle Carriers” and “Yard Gantry Cranes” appeared consistently under each DMU. Thus, “Number of Terminal Cranes” and “Number of Operating Yard Equipment” constantly appeared in each DMU, irrespective of the proportion in which additional inputs appeared in each DMU.

The inputs present in Section 2 and Section 3, as well as the additional inputs in these sections, were included in; “Number of Terminal Cranes”, and “Number of Operating Yard Equipment” respectively. Consequently, “Number of Terminal Cranes” and “Number of Operating Yard Equipment” replaced their subordinate inputs in Section 2 and Section 3, respectively, and therefore in the potential input pool.

In Section 1, the “Berth Length” was defined as the total length of all the berths. Thus, the variable was the same as the “Quay Length” variable. As a result, the input “Quay Length” was removed from the pool of potential inputs.

To add slightly more detail to the input “Berth Length”, the input variable “Number of Berths” was added to the pool of inputs. This variable was included in Section 1.

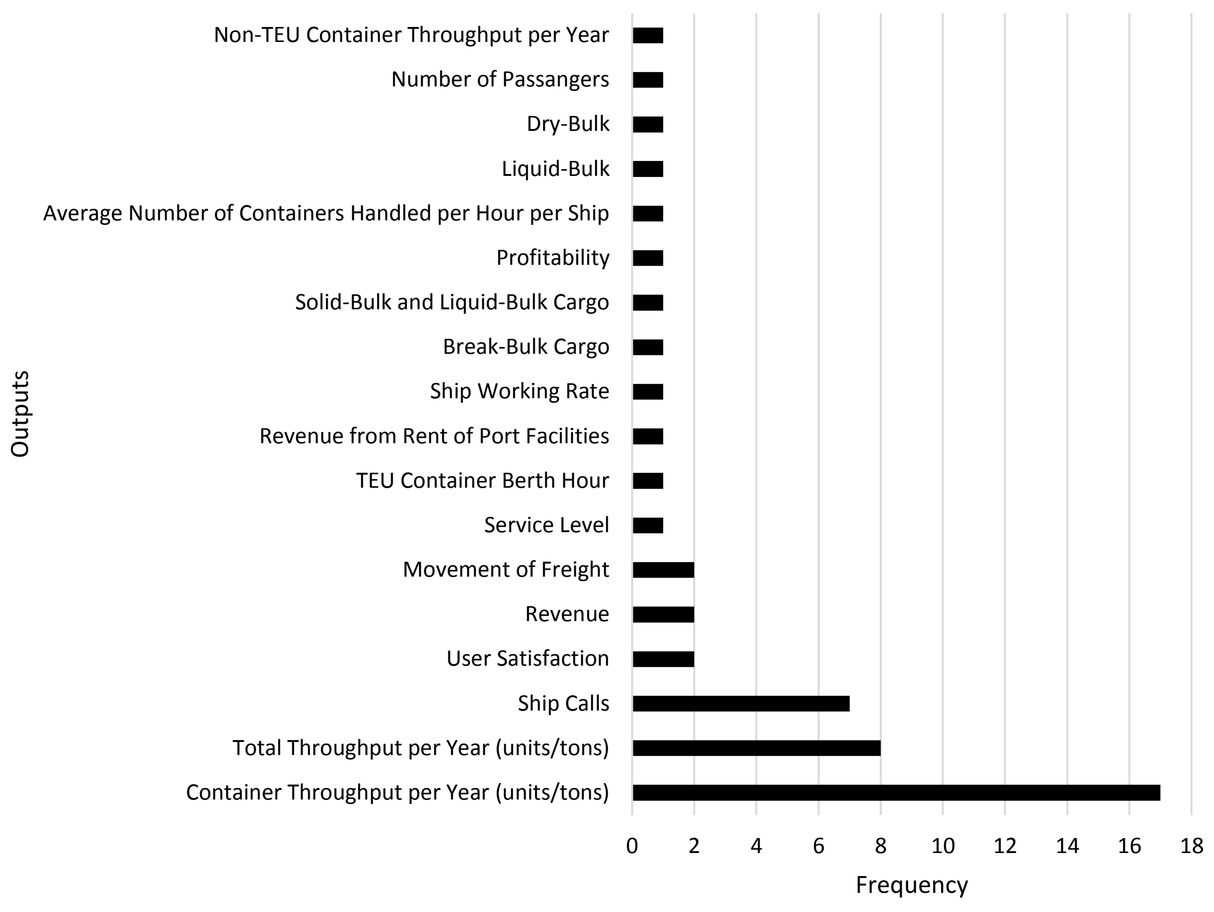
The final pool of potential inputs are summarized by section in Table 3.2. This final pool of potential inputs consisted of; Berth Length (in metres) and Number of Berths in Section 1, Number of Terminal Cranes (including STS Cranes, Quay Cranes, Rail Transfer Cranes and Mobile Cranes) in Section 2, and Number of Operating Yard Equipment (including Straddle Carriers, Gantry Cranes, Rubber-Tyre Gantries, Reach Stackers and Empty Handlers) in Section 3. Section 4 was ignored due to the lack of availability of the labour variable in this section. Please note that in the rest of this study; “Number of Berths” was abbreviated to “Num. Berth”, “Number of Terminal Cranes” abbreviated to “Num. Terminal Cranes”, and “Number of Operating Yard Equipment” abbreviated to “Num. Yard Equipment”. It was not necessary to abbreviate “Berth Length”.

Table 3.2: Allocation of Potential Input Variables to Sections 1-4

Section 1	**Section 2**	**Section 3**	**Section 4**
Berth Length (in metres) Number of Berths	Number of Terminal Cranes	Number of Operating Yard Equipment	Ignored

In terms of the pool of potential output variables, the literature provides overwhelming evidence that "Container Throughput in TEU's (20 Foot Equivalent Units)" is the output to select (see Figure 3.2). This corresponded with the shipping-line companies' feedback that "Container Throughput in TEU's" was the best measure of output in a container terminal. The rationale was due to the relative ease of data collection and it being the primary basis upon which container ports were compared.

Figure 3.2: Output Frequencies in Literature Survey



Thus, the final pool of outputs consisted of a singular output variable "Container Throughput in TEU's". Please note that in the remainder of this study "Container Throughput in TEU's" was abbreviated to "Num. TEU's".

It is highly debatable whether container throughput alone was adequate in assessing the overall efficiency of a container terminal. Researchers feel that this largely ignores other potential outputs, such as market share and customer satisfaction (Panayides *et al.,* 2009). The author agrees with this sentiment, however, due to time constraints, this study did not investigate these issues, and will rather deal with them in future research.

The final pool of potential inputs and outputs for the 15 African container terminals was therefore established. More concrete statistical techniques were required to transform this pool of potential variables into the variables to be used in the selected DEA models.

The first technique involved establishing whether a positive correlation existed between each of the potential inputs and the potential output variable, using the Kendall correlation calculation. This was essential in determining whether the inputs used actually affected the output. This relationship between the variables ensured that efficiency could be calculated. The correlations are listed below in Table 3.3.

Table 3.3: Correlation of Inputs with Output for 2013 and 2014

Input Variables	**2013 Correlation**	**2014 Correlation**
Num. Berths	0.421	0.316
Berth Length	0.295	0.238
Num. Terminal Cranes	0.461	0.473
Num. Yard Equipment	0.490	0.587
Num. TEU's	1.000	1.000

A positive correlation existed between each of the input variables and the output variable for both 2013 and 2014. "Num. Terminal Cranes" and "Num. Yard Equipment" had a stronger correlation with the output variable "Num. TEU's" than the berth variables. "Num. Yard Equipment" exhibited the highest correlation with the output, with values equalling 0.49 and 0.59 in both 2013 and 2014 respectively. "Berth Length", in contrast, had the lowest correlation with the output in both 2013 and 2014, with values equalling 0.30 and 0.24 respectively. This indicated that capital intensity was very important to throughput, more so than the length of the quay in a container terminal. From 2013 to 2014, the berth variable correlations with the output decreased and "Num. Terminal Cranes" and "Num. Yard Equipment" correlations increased.

The second statistical technique applied to identify final variables was called the forward Efficiency Contribution Measure (ECM) (Pastor, Ruiz & Sirvent, 2002). The forward ECM identified the significance of the potential variables, called candidate variables, in terms of their contribution to the efficiency measures of K DMUs (Pastor, Ruiz & Sirvent, 2002). The influence of a candidate variable was measured by the value ϕ_k, where $k = 1, 2, \dots, K$ (Pastor, Ruiz & Sirvent, 2002). The value ϕ_k represents the proportional change to the efficiency of DMU_k, where $k = 1, 2, \dots, K$, when the candidate variable was added to the DEA model (Pastor, Ruiz & Sirvent, 2002). To assist in determining whether the proportional change in the efficiency, ϕ_k, across K DMUs was significant or not the forward ECM procedure defined two parameters externally (Pastor, Ruiz & Sirvent, 2002). These parameters were defined according to what the study believed to be reasonable.

The first parameter was an efficiency score level. This level was represented by $\bar{\rho}_0$, where $\bar{\rho}_0 > 1$. The $\bar{\rho}_0$ was the tolerance level for changes in efficiency scores when a candidate variable was added to the DEA model (Pastor, Ruiz & Sirvent, 2002). The second parameter was the probability level. The probability level was represented by p_0, where $0 < p_0 < 1$. The p_0 was the proportion of DMU's with an efficiency change that exceeded the tolerance level (Pastor, Ruiz & Sirvent, 2002). As an example, $p_0 = 0.20$ and $\bar{\rho}_0 = 1.1$ indicated that efficiency scores of more than 20% of the DMUs would have to increase by more than 10% when a candidate variable was added to the DEA model for the candidate variable to be considered significant (Pastor, Ruiz & Sirvent, 2002). Thus, $\bar{\rho}_0$ and p_0 provided an operative influence statement to be tested by the ECM (Natatraja & Johnson, 2011).

Let $\Omega_1, \Omega_2, \dots, \Omega_K$ denote a sample from the distribution of the random variable Ω, since the values of the random variable are measures of the influence of the candidate variable on K DMUs (Pastor, Ruiz & Sirvent, 2002). The distribution of the random variable Ω was defined on the interval $[0,1)$ (Pastor, Ruiz & Sirvent, 2002). An indicator variable,

$$A_k \begin{cases} 1 & if\ \Omega_k > \bar{\rho}_0 \\ 0 & otherwise \end{cases} \tag{3.1}$$

for $k = 1, 2, \dots, K$, was defined (Pastor, Ruiz & Sirvent, 2002). Let $p = P(\Omega > \bar{\rho}_0)$, then it follows that $A = \sum_{k=1}^{K} A_k$ follows a binomial distribution with parameters K and p (Pastor, Ruiz & Sirvent, 2002). Using A as a test statistics the following was tested:

$$\begin{aligned} H_0 &: p \leq p_0 \\ H_1 &: p > p_0 \end{aligned} \tag{3.2}$$

The rejection of the null hypothesis in this case will indicate that there is sufficient statistical evidence to conclude that the total efficiency scores of more than $p_0 \times 100\%$ changed by more than $\bar{\rho}_0 \times 100\%$ when the candidate variable was included in the model. To calculate p-values for this test it will be considered that $A \sim Binomial\,(K-1, p_0)$ under the null (Pastor, Ruiz & Sirvent, 2002). The distribution considers $K-1$ DMU's as DEA is a benchmarking technique that requires at least one DMU, DMU_k, to be fully efficiency in order to calculate the remaining relative efficiencies. Thus, DMU_k will experience an ϕ_k value equal to zero. The p-value, p_{val}, for the test was calculated as

$$p_{val} = P(A > A_0), \tag{3.3}$$

where A_0 was the observed value of A (Pastor, Ruiz & Sirvent, 2002).

The tolerance level ($\bar{\rho}_0$) used was equivalent to 1.01 and the proportion level (p_0) equivalent to 0.1. The levels were selected to ensure sensitivity to change. The levels selected were also in line with the literature (Pastor, Ruiz & Sirvent, 2002; Natatraja & Johnson, 2011). The p-value, together with a desired significance level, was considered to determine whether the test-statistic, and its associated candidate variable, was significant (Pastor, Ruiz & Sirvent, 2002).

The forward ECM procedure was conducted on both the CCR and the BCC models in 2013 and 2014. The orientation of the models was output-orientated. The significance level used in the forward ECM was 10%. The ECM consisted of three rounds to determine which of the four potential inputs was significant (Pastor, Ruiz & Sirvent, 2002).

Tables C and D in Appendix Two illustrate the forward ECM procedure in detail. The X variables defined in the forward ECM represented the significant input(s) at the start of each step or round. These were the final input(s) to be used in the DEA models selected. Each round needed at least one significant input, or X, so that an efficiency measure could be calculated. As no test was conducted at the start of the forward ECM to prove the significance of the initial X input, this input had to be selected by other, credible, means. Thus, the initial X variable had to be, without a doubt, the most important input amongst the potential pool of input variables. So important that it could immediately be considered as a final input to be used in the selected DEA models. This study identified its initial input by analysing the frequency with which the inputs were used in the literature as well as considering the input's importance in the industry. The initial X input selected was "Berth Length" as it was the most frequently used input in the literature surveyed, as evident in Figure 3.1. "Berth Length" was also identified by shipping-line companies as being important in calculating their industry efficiency measures. As such, "Berth Length" became the first definitive input to be used in the selected DEA models (Pastor, Ruiz & Sirvent, 2002).

The Z variables defined in the forward ECM were the candidate (potential input) variables. These initially included; "Num. Berths", "Num. Terminal Cranes", and "Num. Yard Equipment". In each round, the forward ECM identified the candidate variable that made the most significant (smallest p-value) contribution to the efficiency value. The most significant candidate then became an input variable X in the next round of the forward ECM. This continued until there were no further significant candidates remaining. At this point, all the significant candidates, including the initial X input, were the final input variables to be used in the DEA models selected (Pastor, Ruiz & Sirvent, 2002).

The Y variable defined in the forward ECM represented the final output(s) "Num. TEU's" to be used in the selected DEA models (Pastor, Ruiz & Sirvent, 2002). As there was only one potential output variable, it was not possible to perform this stepwise procedure with the output. Despite that, the output "Container Throughput in TEU's" was selected as the final singular output. The reason for this being that the output was shown to be frequently used in the literature and industry for efficiency measurement. An additional reason for the selection of this output was that it strongly correlated with the potential inputs. The author is therefore confident of its significant contribution towards the efficiency value calculated using DEA (Pastor, Ruiz & Sirvent, 2002).

In Tables C and D of Appendix Two, the colour red indicates the most significant candidate variable in each round. The colour green indicates insignificant candidate variables and the yellow colour highlights the final inputs selected.

In the 2013 CCR model, “Num. Yard Equipment” (1.000e-14), “Num. Terminal Cranes” (1.251e-06) and “Num. Berths” (0.002) was the order of significance. “Num. Yard Equipment” had the biggest effect on the CCR efficiency, followed by “Num. Terminal Cranes” and “Num. Berths” respectively.

By contrast, in the 2013 BCC model, “Num. Terminal Cranes” (2.729e-09), “Num. Yard Equipment” (0.002) and “Num. Berths” (0.009) was the order of significance. “Num. Terminal Cranes” had the biggest effect on the BCC efficiency, followed by “Num. Yard Equipment” and “Num. Berths” respectively.

In the 2014 CCR model, “Num. Yard Equipment” (1.270e-12) followed by “Num. Terminal Cranes” (0.009) was the order of significance. “Num. Yard Equipment” had the biggest effect on the CCR efficiency, followed by “Num. Terminal Cranes”. “Num. Berths” had no significant effect on the efficiency, with a p-value of 0.415.

“Num. Yard Equipment” (1.721e-05) followed by “Num. Terminal Cranes” (0.044) was the order of significance in the 2014 BCC (VRS) model. “Num. Yard Equipment” had the biggest effect on the BCC efficiency, followed by “Num. Terminal Cranes”. Again, “Num. Berths” had no significant effect on the efficiency score, with a p-value of 0.415.

“Num. Berths” made no significant contribution to the DEA efficiency score. Thus, this variable would generally not be considered as a final input. However, it was retained in the CCR and BCC models for 2014 as it had a positive correlation (0.316) to the output “Num. TEU’s”. The variable was also found to be significant in 2013 and was a strong subordinate component of “Berth Length”, which was one of the most important inputs.

Given their positive correlations, significant p-values and general importance, “Berth Length”, “Num. Terminal Cranes”, “Num. Yard Equipment” and “Num. Berths” were chosen as the final inputs. These inputs were used in the respective DEA models in 2013 and 2014 to calculate efficiency.

The descriptive statistics for the final inputs and output are listed in Table 3.4.

Table 3.4: Descriptive Statistics for Input and Output Variables 2013 and 2014			
Variable	**Descriptive Statistic**	**2013**	**2014**
Berth Length (in metres)	Number of Ports	15	15
	Minimum	400	400
	Maximum	2668	2668
	Mean	1109.333	1109.33
	Median	926	926
	SD	667.874	667.874
Number of Berths	Minimum	2	2
	Maximum	11	11
	Mean	4.8	4.8
	Median	4	4
	SD	2.731	2.731
Number of Terminal Cranes	Minimum	4	4
	Maximum	25	27
	Mean	9.867	10.133
	Median	8	8
	SD	6.081	6.424
Number of Operating Yard Equipment	Minimum	13	13
	Maximum	90	90
	Mean	38.2	38.4
	Median	29	29
	SD	20.953	21.260
Container Throughput in TEU’s	Minimum	289963	259917
	Maximum	4100000	4100000
	Mean	1207552	1286135
	Median	825189	860000
	SD	1043799	1085773

The descriptive statistics indicate above that “Berth Length” and “Num. Berths” was unchanged over the sample period. There were slight increases in “Num. of Terminal Cranes”, “Num. Yard Equipment” and “Num. TEUs”. There were large differences between mean and median values in some of the variables. These differences were present in both 2013 and 2014. The distributions of these variables were skewed to the right, believed to be caused by the outlier Port Said. This port was the only African container terminal to be ranked in the top five container terminals in terms of TEU throughput, in the Mediterranean, for both 2013 and 2014 by Containerisation International (2015).

3.4 Rules for Minimum Sample Size in DEA

There are rules within the DEA literature that need be considered to ensure a minimum sample size in order to maintain discriminatory power of the DEA model. If this discriminatory power was not present in the DEA model, efficiency values would be biased upward to a point where DMUs that were inefficient, would be incorrectly identified as efficient. This would occur as a result of the sample of DMUs being too small (Panayides *et al.*, 2009; Sakris, 2002).

Four rules were identified to ensure that discriminatory power existed within the selected DEA models. These rules are listed below in Table 3.5. Each rule was met, ensuring the presence of discriminatory power in the selected DEA models.

Table 3.5: Minimum Sample Size Rules for Discriminatory Power

Rules	**Rule's Minimum Sample Size**	**Study's Minimum Sample Size**	**Discriminatory Power Exist (Yes)/Doesn't Exist (No) in This Study**
Boussofiane, Dyson and Thanassoulis (1991) stipulate that to get good discriminatory power out of the CCR and BCC models the lower bound on the number of DMUs should be the multiple of the number of inputs and the number of outputs.	4	15	Yes
Golany and Roll (1989) establish a rule that the number of units (DMUs) should be at least twice the number of inputs and outputs considered.	10	15	Yes

Bowlin (1998) mentions the need to have three times the number of DMUs as there are input and output variables.	15	15	Yes
Dyson *et al.*,(2001) recommend a total of two times the product of the number of input and output variables.	8	15	Yes
Source: (Sakris, 2002)			

3.5 Conclusion

Homogeneity existed between the terminals given their geographical association and the nature of the goods moved through the terminals (Panayides *et al.*, 2009). The input variables selected for the sample period were; Berth length, Number of Berths, Number of Terminal Cranes, and Number of Operating Yard Equipment. The output variable selected for the sample period was Container Throughput (in TEU's). The variables were selected based on the high frequency with which they occurred in the literature, their importance in industry, as well as their positive correlation and significance in the statistical tests. Four popular minimum sample size rules were met. This ensured that discriminatory power existed within the selected DEA models.

Chapter Four: Methodology

4.1 Introduction

This section introduces the reader to the DEA methodology adopted in this study. Firstly, the technique used to calculate the relative efficiency for the simplest input-output case is illustrated. This is followed by a description of how to calculate efficiency for the single variable input-double variable output case, and concludes with the multiple input-output case. This process provides a basis for the methodologies of the CCR and BCC models used to calculate efficiency in a multiple input-output case. These models will be used to calculate efficiency within the selected African container terminals case study.

The CCR and BCC models calculate technical efficiency (TE) and pure technical efficiency (PTE), respectively. In addition, to account for all forms of efficiency present in the selected African container terminals, the BCC methodology is extended to calculate scale efficiency (SE).

A smooth homogenous bootstrapping procedure and the MPI are also presented in this section. The bootstrapping procedure is used to correct for sampling error bias present in the efficiency results, whilst the MPI is used to track efficiency changes over the sample period.

4.2 Single Input-Output Efficiency Measures

To provide the basics needed to understand the methodology related to DEA a single input-output example was used to produce efficiency measures. This illustration of an example follows from Cooper, Seiford and Tone (2007), and provides a simple approach to those unfamiliar with DEA. The variable "throughput" is the output and the variable "stevedores" is the input. Stevedores are labourers that assist in the uploading and offloading of containers from the vessel. The variables were recorded for eight terminals labelled $T1$ to $T8$. The values of the variables for each terminal, as well as the productivity of each, are listed in Table 4.1.

Table 4.1: Single Input-Output Case								
Terminal	T1	T2	T3	T4	T5	T6	T7	T8
Stevedores	2	3	3	4	5	5	6	8
Throughput	1	3	2	3	4	2	3	5
Throughput/Stevedore (Productivity)	0.5	1	0.67	0.75	0.8	0.4	0.5	0.63
Source: (Cooper , Seiford & Tone, 2007)								

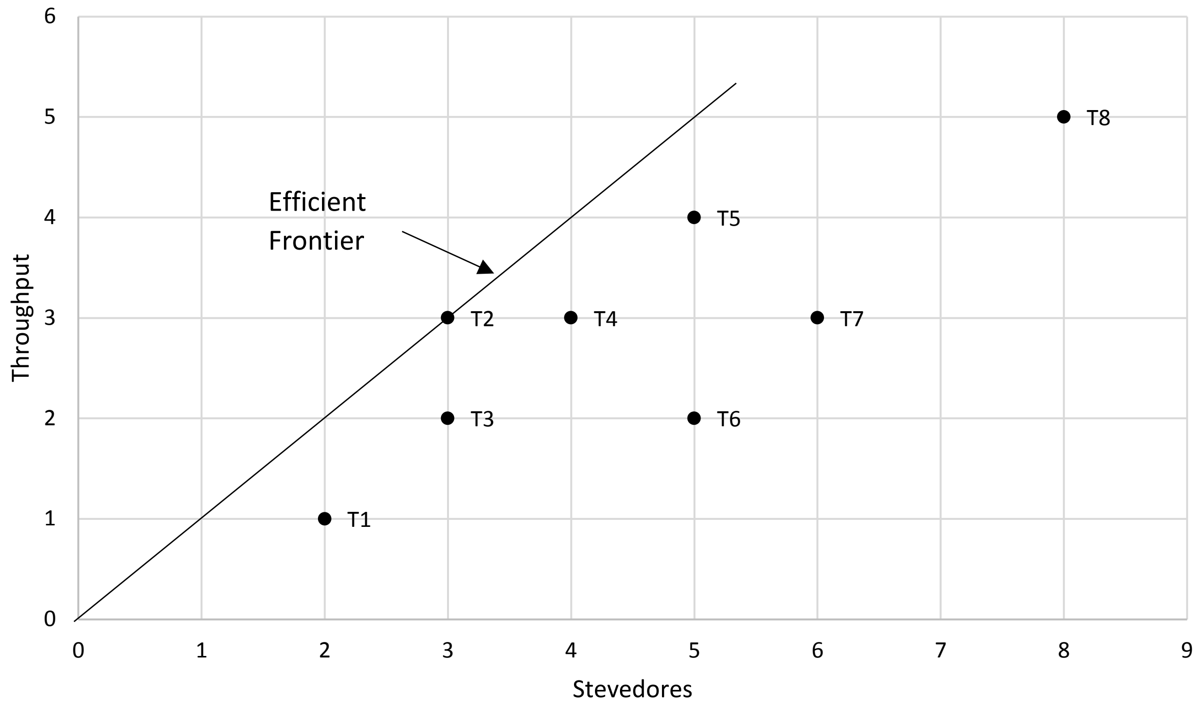
The basis of the DEA efficiency measure was the following productivity ratio

$$\frac{Output}{Input}. \tag{4.1}$$

The above commonly used ratio measured the productivity of the terminals in this study. Based on terminal productivity defined as throughput per stevedore, $T2$ was considered to be the most productive terminal and $T6$ the least productive terminal (Cooper, Seiford & Tone, 2007).

For ease of interpretation, the results were shown graphically. The terminals were plotted as points in terms of their input and output values in Figure 4.1. Stevedores were represented on the horizontal axis and throughput on the vertical axis. The slope of the line connecting each point to the origin corresponded to the productivity of that particular point in terms of Eq. (4.1) (Cooper, Seiford & Tone, 2007).

Figure 4.1: Illustration of the Efficient Frontier



Source: (Cooper, Seiford & Tone, 2007)

The line with the most positive slope in this case was that connecting $T2$ to the origin. This slope indicated the productivity of the terminal and was called the efficient frontier. The efficient frontier enveloped the unproductive points, namely $T1$, $T3$, $T4$, $T5$, $T6$, $T7$ and $T8$, hence the terminology for DEA - data envelopment analysis. The efficient frontier in Figure 4.1 exhibited constant returns-to-scale because for every one unit increase in input there was a corresponding unit increase in output (Cooper, Seiford & Tone, 2007).

The above ratio only considered the productivity of each terminal, and not its efficiency. To compare terminals, efficiency was defined as relative productivity. Within the relative productivity calculation, a benchmark productivity measurement was required to allow for an efficiency calculation for each terminal. Traditionally this benchmark measure was the productivity of the most efficient terminal, in this case, $T2$. The starting point for every efficiency calculation was therefore the determination of the efficient frontier as this frontier represented the performance of the most productive DMU, in this case terminal $T2$. As a result, the efficiency of each of the eight terminals was the productivity of each terminal relative to the productivity of $T2$. Accordingly, DEA identified the most productive DMU, in this case $T2$, to serve as the benchmark to use in the comparisons. Thus, the following computation was applied

$$0 \le \frac{Throughput\ per\ stevedore\ of\ Tn}{Throughput\ per\ stevedore\ of\ T2} \le 1 \qquad (4.2)$$

where $n = 1, 2, \dots, 8$, to determine the efficiency of each terminal. The efficiency results were between zero and unity. The full set of efficiency results obtained by applying this computation is listed in Table 4.2 (Cooper, Seiford & Tone, 2007).

Table 4.2: Efficiency

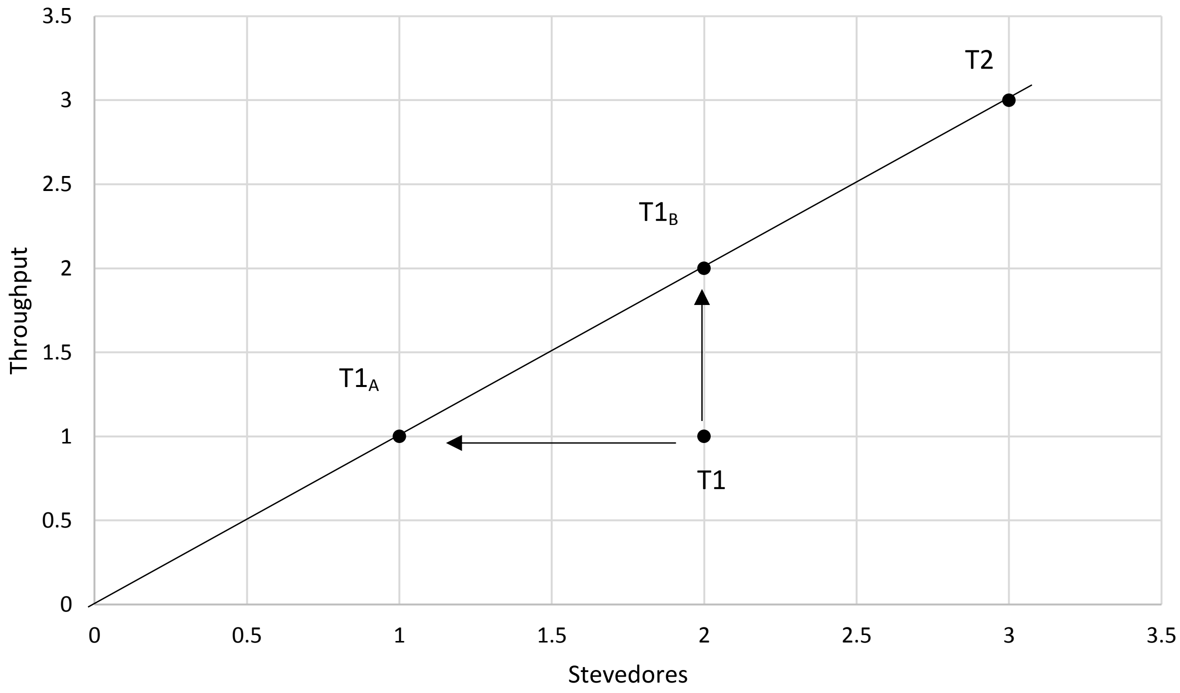
Terminal	T1	T2	T3	T4	T5	T6	T7	T8
Efficiency	0.5	1	0.67	0.75	0.8	0.4	0.5	0.63

Source: (Cooper, Seiford & Tone, 2007)

Defining efficiency as a relative productivity, as seen in Eq. (4.2), was based on the unit invariance property. If this study were to define efficiency and productivity values as equal, the value of efficiency would depend on the unit of measurement. However, defining efficiency as a relative productivity eliminated the effect of the unit of measurement. This was useful when measuring the efficiency of DMUs consisting of multiple inputs-outputs, where all were measured in different units of measurement (Cooper, Seiford & Tone, 2007).

Inefficiency in the DEA model was the distance of an inefficient DMU, from an efficient version of itself, on the efficient frontier. Thus, to correct for inefficiency in a DEA model, an inefficient DMU was projected from its current point to a point on the frontier. To illustrate this movement, one of the inefficient DMU's $T1$ was isolated in Figure 4.2. This movement was achieved by reducing the current input levels (number of stevedores) to move DMU $T1$ to $T1_A$, with coordinates $(1, 1)$ on the efficient frontier. This orientation towards the efficient frontier was referred to as the input-orientation. The input-orientation involved a horizontal shift of DMU $T1$ to the efficient frontier. Another movement for correcting the inefficiency of DMU $T1$ involved raising the throughput up to move $T1$ to $T1_B$ $(2, 2)$ on the efficient frontier. This orientation towards the efficient frontier was referred to as the output-orientation. The output-orientation involved a vertical shift of DMU $T1$ to the efficient frontier. Thus, the orientation selected for the DEA model determined how the model corrected for inefficiency (Cooper, Seiford & Tone, 2007).

Figure 4.2: Correcting for Inefficiency in DEA



Source: (Cooper, Seiford & Tone, 2007)

The approach used to determine the relative efficiency in a single input-output setting would not be appropriate when calculating efficiency in the selected African container terminals. Consequently, this approach needed to be adapted to consider more than a single input and output in order to calculate the efficiency of a multivariable terminal.

The purpose of this simplified case was to introduce the basics of the DEA methodology, including its productivity ratio and the efficient frontier. These two components were essential parts of the benchmarking approach that DEA adopted towards calculating the efficiency of a DMU. This case also explained that DEA calculated relative efficiency to maintain the property of unit invariance. Finally, this case illustrated how the input- and output-orientation work when correcting for inefficiency. These basics are extended to calculate efficiency for the selected African container terminals during the sample period.

4.3 One Input-Two Outputs Case

To make the extension to a multiple input-output case, this section considered the one input-two outputs case. Again this was an illustration of an example from Cooper, Seiford and Tone (2007). Their example provided a simple approach to build on the methodologies used in the single input-output case. The input was "stevedores" and the outputs were "satisfied customers" and "throughput". These variables were recorded for seven terminals labelled $T1$ to $T7$. The variables and terminals are listed in Table 4.3. The input variable was defined as the only x variable. The two output variables were defined as y variables (Cooper, Seiford & Tone, 2007).

Table 4.3: One Input-Two Outputs Case

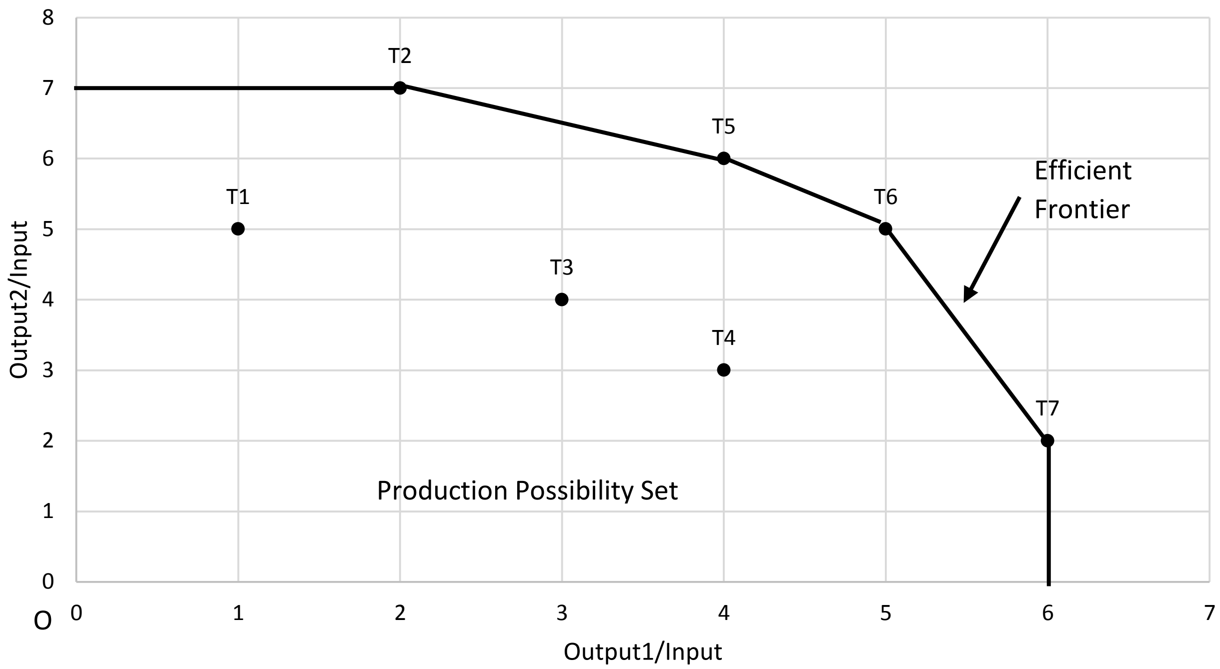
Terminal		T1	T2	T3	T4	T5	T6	T7
Stevedores	x	1	1	1	1	1	1	1
Satisfied Customers	y_1	1	2	3	4	4	5	6
Throughput	y_2	5	7	4	3	6	5	2

Source: (Cooper, Seiford & Tone, 2007)

To calculate the efficiency using DEA in a one input-two outputs case the first step was to divide each output by the number of stevedores as it was considered the only input of interest in this example. This division allowed for a unitised efficient frontier to be constructed.

Figure 4.3 depicts the efficient frontier. This efficient frontier was constructed slightly differently to the efficient frontier in Figure 4.1. The efficient frontier in this multiple variable case was simply the line connecting the terminals that produced the most outputs with their given input. In the single input-double outputs case, these terminals were $T2$, $T5$, $T6$ and $T7$. The trade-offs between these terminals were not discussed here. It was simply noted that none of the terminals on the frontier line could increase one of its outputs without worsening the other. As a result, it made sense to see these terminals as efficient. Figure 4.3 also depicts the production possibility set. This was the region bounded by the axes and the frontier line. Figure 4.3 also depicts the labelled terminals which were plotted in terms of their unitised outputs given the unitisation of the efficient frontier (Cooper, Seiford & Tone, 2007).

Figure 4.3: One Input-Two Outputs Case

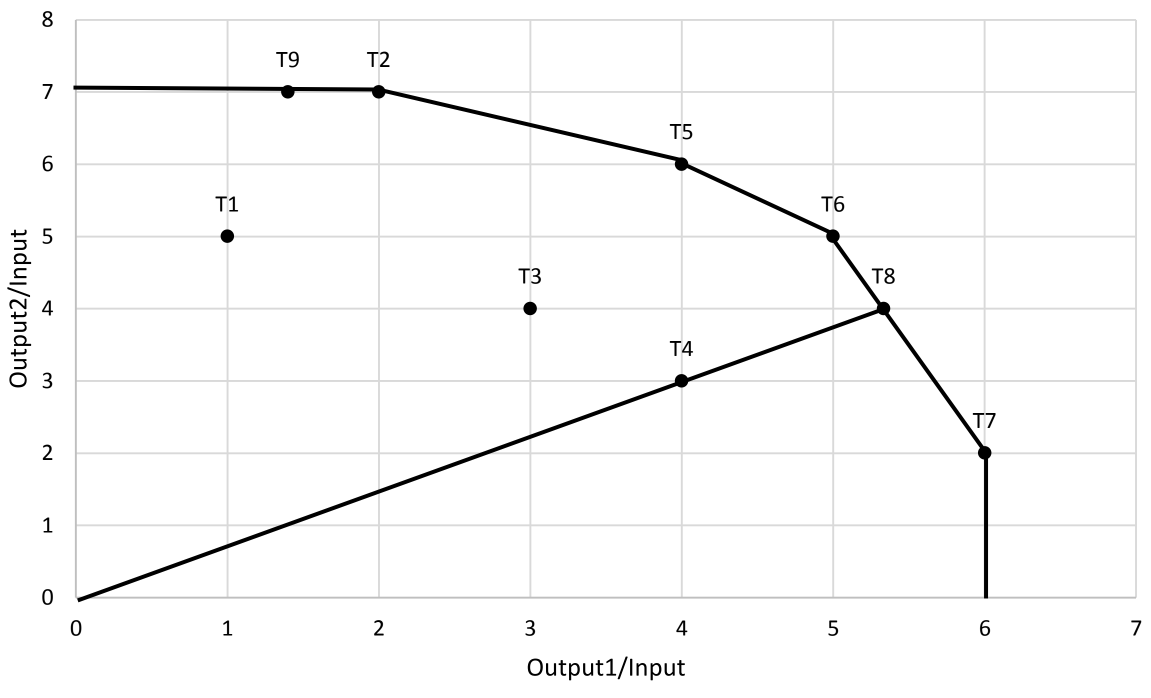


Source: (Cooper, Seiford & Tone, 2007)

Terminals $T1, T3$ and $T4$ were inefficient as they were enveloped by the efficient frontier. Their efficiency could be evaluated by referring to the efficient frontier as a benchmark. To compute the efficiency of a particular DMU in this single input-double outputs variable case, a ratio was constructed and solved. This ratio consisted of the radial distance of a DMU relative to the radial distance of an efficient version of itself on the efficient frontier. This efficient version of itself may have been an existing efficient DMU or just a point on the efficient frontier. This point on the efficient frontier was referred to as the efficient composite of that particular DMU (Cooper, Seiford & Tone, 2007).

To demonstrate the benchmarking approach in this single input-double outputs case, the efficiency of $T4$ was evaluated in Figure 4.4. This efficiency was calculated by the relative distance measure $\frac{d(O,T4)}{d(O,T8)}$. The measures $d(O,T4)$ and $d(O,T8)$ represented the distances from the origin to $T4$ and the origin to $T8$, respectively. The distance ratio used to evaluate the efficiency was referred to as a radial measure. As the radial extension of $T4$ did not coincide with an existing efficient terminal on the efficient frontier, an efficient composite of $T4$ had to be defined. The terminal $T8$ was the efficient composite version of $T4$ and that point represented the radial intersection between the radial extension of $T4$ to the efficient frontier and the efficient frontier. The terminal $T4$ was benchmarked against $T8$ to determine its efficiency. The Euclidian measures were given by $d(O,T4) = \sqrt{4^2+3^2} = 5$ and $d(O,T8) = \sqrt{(\frac{16}{3})^2+4^2} = \frac{20}{3}$. The terms under the radical signs were squares of $\frac{y_1}{x}$ and $\frac{y_2}{x}$ variables of $T4$ and $T8$, respectively. As $T4$ was an existing terminal, its $\frac{y_1}{x}$ and $\frac{y_2}{x}$ values were obtained from Table 4.3. However, T8 was an efficient composite of T4 that needed to be defined. Thus, the $\frac{y_1}{x}$ and $\frac{y_2}{x}$ values needed to be calculated. This point of intersection could be found by solving simultaneously for $\frac{y_2}{x} = \frac{3}{4}\frac{y_1}{x}$ and $\frac{y_2}{x} = 20 - 3\frac{y_1}{x}$. Substituting the distance values into the ratio of distances $\frac{d(O,T4)}{d(O,T8)}$ yielded $5 \div \frac{20}{3} = \frac{15}{20} = 0.75$, or an efficiency of 75% for T4 (Cooper, Seiford & Tone, 2007).

Figure 4.4: Efficiency Illustration in the Single Input-Double Outputs Case



Source: (Cooper, Seiford & Tone, 2007)

The efficiency results generated were between zero and unity. This was due to the ratio being formed relative to the Euclidian distance from the origin over the production possibility set. However, in this case study the output-orientation was needed to correct for inefficiency. Thus, this research used calculations which corrected for outputs, while maintaining inputs, to obtain efficiency. As a result, the interpretation of the efficiency ratio was in terms of its reciprocal $\frac{d(O,T8)}{d(O,T4)} = \frac{20}{3} \div 5 = 1.33$. This result stated that to be efficient, $T4$ would have to increase its outputs by 33% to achieve full efficiency. To confirm that this was the case, this ratio was applied to the y_1 and y_2 values of $T4$ to obtain the co-ordinates $\frac{4}{3}(4,3) = \left(\frac{16}{3}, 4\right)$. The co-ordinates are the values of the efficient composite $T8$. This was the point on the efficient frontier used to evaluate $T4$. Consequently, in these calculations the efficiency value would always be between one and infinity (Cooper, Seiford & Tone, 2007).

In this case study, a higher dimension model is necessary. To demonstrate a higher dimension model and the interpretations, the single input-double outputs case is adapted. The three variable case was used to show that in all multiple variable cases DEA calculated the efficiency of a DMU by constructing an efficient composite of the DMU. The particular DMU was then radially compared to the efficient composite to determine its efficiency. It is important to consider the use of efficient composite DMUs in DEA efficiency calculation in addition to the basics learned in the single input-output case. All these components were used in the CCR and BCC models that estimate efficiency in the container terminals of the case study.

The graphical illustrations such as Figure 4.1 to Figure 4.4 were used to visualise the single input-output and single input-double outputs examples. Unfortunately the analysis in this research used a single output and four inputs that cannot be shown graphically. The number of variables increased the dimensions to a point where the DEA process could no longer be represented two-dimensionally. As a result, greater emphasis was placed on the components identified as important in the single input-output, as well as the single input-double outputs, cases, rather than their graphical depictions. The reason being that these components are more important to the DEA process used to calculate the efficiency of the selected African container terminals.

4.4 The CCR and BCC Models

The CCR and BCC models were used to calculate efficiency in the African container terminals. Each African container terminal represented a multiple input-output environment consisting of more than three variables.

In the CCR and BCC models, virtual inputs and outputs were formed for each DMU using unknown weights λ_i. The λ_i's were referred to as dual weights. DEA used variable weights that were derived directly from the data. The benefit of this was that numerous a priori assumptions and computations involved in fixed weight choices were avoided. The weights in DEA were chosen in a manner that assigned the best set of weights to each terminal. The term "best" was used here to mean that the resulting input-to-output ratio for each terminal was maximised relative to all the other terminals. The weights were determined within the CCR and BCC models (Cooper, Seiford & Tone, 2007).

By weighting the inputs and outputs of respective fully efficient DMUs, an efficient composite DMU was established for each DMU. This efficient composite of the DMU was located on the efficient frontier. The efficiency of each DMU was determined radially, relative to the efficient composite of this DMU. The optimal weights may have (and generally would) vary from one DMU to another, and also between the models. Thus, in addition to the components identified as important in the previous two cases, dual weights were of great importance to efficiency calculations in the CCR and BCC models.

4.4.1 The CCR Model

The CCR (Charnes, Cooper & Rhodes, 1978) model allowed for the construction of an efficient frontier that accounted for the technical efficiency (TE) of the DMUs and assumed constant returns-to-scale (CRS). This CRS assumption was based on a property contained within the CCR production possibility set. This property stated that if (x, y) is a feasible point, then (ax, ay) for any positive a would also be feasible. Thus, the CCR model's efficient frontier, if depicted, would resemble the efficient frontier used in the single input-output case in Figure 4.1 (Cooper, Seiford & Tone, 2007).

An output-orientated CCR model was utilised as this orientation provided for an efficiency assessment of a port's output capacity as recommended by Wu and Goh (2010). This orientation indicated that inefficiency was corrected for by adjusting outputs whilst keeping inputs fixed.

The CCR model used linear programming to calculate the dual weights which then allows for the determination of the TE of a DMU. The output-orientated formulation of the CCR linear programming model evaluated K DMUs using the same m inputs x_t $(t = 1, 2, \dots, m)$ to produce the same n outputs y_s $(s = 1, 2, \dots, n)$ was given as

$$\max_{\eta_{CCR}, \lambda_i} \eta_{CCR} \tag{4.3}$$

subject to

$$x_{tk} - \sum_{i=1}^{K} \lambda_i \, x_{ti} \ \geq 0 \ (t = 1, 2, \dots, m) \tag{4.4}$$

$$\eta_{CCR} y_{sk} - \sum_{i=1}^{K} \lambda_i \, y_{si} \leq 0 \ (s = 1, 2, \dots, n) \tag{4.5}$$

$$\lambda_i \geq 0 \text{ for all } i \tag{4.6}$$

where x_{tk} represented the quantity of input t used by DMU k, and y_{sk} denoted the quantity of output s produced by DMU k (Brettenny & Sharp, 2016).

The CCR model constructed a composite unit for DMU k, that outperformed DMU k, using the dual weights λ_i assigned to DMU k by the linear programme. The efficient composite DMU of inefficient DMU k was constructed by weighting and summing similar inputs and outputs of homogenous fully efficient DMUs. These efficient DMUs had non-zero dual weights λ_i and comprised the reference set for inefficient DMU k. This reference set served as a basis for computing the efficiency score of DMU k, through the construction of an efficient composite of DMU k (El-Mahgary & Lahdelma, 1995). It should be noted that the DMUs of the reference set were the efficient composite versions of themselves. They had dual weights equal to unity and were therefore seen as fully efficient (El-Mahgary & Lahdelma, 1995).

The composite unit consumed inputs $\sum_{i=1}^{K} \lambda_i \, x_{ti}$, where $t = 1, 2, \dots, m$. The efficient composite of DMU k had inputs that were at most equal to the corresponding inputs of unit k

identified as x_{tk}, where $t = 1, 2, \dots, m$. The efficient composite of DMU k produced at least a proportion η_{CCR} of the outputs of DMU k (El-Mahgary & Lahdelma, 1995). The η_{CCR} would represent the TE had the data represented the population. However, given that a sample is being used, solving the linear programme provided an estimate of this proportion, ${\eta_{CCR}}^*$, for each of the K DMUs. The ${\eta_{CCR}}^*$ value for DMU k was referred to as its technical (Farrell) efficiency (TE). The Farrell efficiency measure stated that $[(1 - {\eta_{CCR}}^*) \times 100]$ was equivalent to the percentage by which DMU k must increase its outputs, while maintaining inputs, to become relatively efficient. DMU k was deemed CCR efficient if the solution is ${\eta_{CCR}}^* = 1$ and all associated slacks were equal to zero (Brettenny & Sharp, 2016). For each assessed DMU, the slacks were described as the excesses of inputs and/or shortfalls in outputs which could be required in addition to the proportional increase in outputs by the factor ${\eta_{CCR}}^*$. The input and output slacks (s^-, s^+) are used in the output-orientated CCR model as $\sum_{i=1}^{K} \lambda_i \, x_{ti} + s^- = x_{tk}$ and $\sum_{i=1}^{K} \lambda_i \, y_{si} - s^+ = \eta_{CCR} y_{sk}$, where $t = 1, 2, \dots, m$ and $s = 1, 2, \dots, n$ to adjust for inputs and output, respectively, when necessary (Cooper, Seiford & Tone, 2007).

4.4.2 The BCC Model

Various extensions of the CCR model have been proposed, one of which was the BCC (Banker-Charnes-Cooper, 1984) model. The BCC model was used to construct an efficient frontier that accounted for the pure technical efficiency (PTE) of the DMUs. The BCC model's efficient frontier assumed variable returns-to-scale (VRS). The BCC model had its efficient frontier spanned by the convex hull of efficient DMUs. The frontier had piecewise linear and concave characteristics which lead to VRS characterisation (Cooper, Seiford & Tone, 2007). The VRS assumption of the BCC model resulted in an efficient frontier which can exhibit CRS, increasing returns-to-scale and decreasing returns-to-scale.

The BCC model used linear programming to calculate both the dual weights which then allowed for the determination of the PTE of a DMU. The output-orientated formulation of the BCC linear programming model evaluated K DMUs using the same m inputs x_t $(t = 1, 2, \dots, m)$ to produce the same n outputs y_s $(s = 1, 2, \dots, n)$ was given as

$$\max_{\eta_{BCC}, \lambda_i} \eta_{BCC} \tag{4.7}$$

subject to

$$x_{tk} - \sum_{i=1}^{K} \lambda_i \, x_{ti} \; \geq 0 \; (t = 1, 2, \dots, m) \qquad (4.8)$$

$$\eta_{BCC} y_{sk} - \sum_{i=1}^{K} \lambda_i \, y_{si} \leq 0 \; (s = 1, 2, \dots, n) \qquad (4.9)$$

$$\sum_{i=1}^{K} \lambda_i = 1 \qquad (4.10)$$

$$\lambda_i \geq 0 \text{ for all } i \qquad (4.11)$$

where x_{tk} represented the quantity of input t used by DMU k, and y_{sk} denoted the quantity of output s produced by DMU k (Brettenny & Sharp, 2016).

The efficiency of DMU k was calculated in the BCC model by the construction of an efficient composite of DMU k using dual weights λ_i and the reference set identified in the linear programme. Then DMU k was radially compared to the efficient composite of DMU k, located on the efficient frontier, to determine its efficiency value (El-Mahgary & Lahdelma, 1995). Thus, the BCC model calculated efficiency in the same manner as the CCR model.

Therefore, solving the linear programme provided an estimate of the proportional increase, ${\eta_{BCC}}^*$, in outputs for each of the K DMUs. The ${\eta_{BCC}}^*$ value for DMU k was referred to as its pure technical (Farrell) efficiency (PTE). The Farrell efficiency measure stated that $[(1 - {\eta_{BCC}}^*) \times 100]$ was equivalent to the percentage by which DMU k must increase its outputs, while maintaining inputs, to become relatively efficient. A DMU was deemed BCC efficient if the solution is ${\eta_{BCC}}^* = 1$ and all associated slacks (s^-, s^+) were equal to zero (Brettenny & Sharp, 2016). For each assessed DMU, the slacks were described as the excesses of inputs and/or shortfalls in outputs which may be required in addition to the proportional increase in outputs by the factor ${\eta_{BCC}}^*$. The slacks were applied, when necessary, in the same manner as in the CCR model (Cooper, Seiford & Tone, 2007).

The only difference between the CCR and BCC model was the adjunction of the condition $\sum_{i=1}^{K} \lambda_i = 1$ on the dual weights. Together with the condition $\lambda_i \geq 0$, for all i, this imposed a convexity condition on allowable ways in which efficient DMUs could be combined to generate the efficient composite DMU k. The dual weight restriction ensured that the reference set was selected in such a way as to ensure that the efficient composite unit of DMU k was of the same scale size (Cooper, Seiford & Tone, 2007). No such restriction existed when calculating TE in the CCR model.

Despite this restriction on the dual weights, there would still be instances where $TE = PTE$. However, due to the VRS brought about by the convex hull for an efficient frontier of the BCC model, the PTE would generally be higher than TE. The reason for this was that the efficient frontier with its convex hull shape restricted the data points more in certain areas of the production possibility set than the CCR's CRS efficient frontier. This meant that DMUs were closer to their efficient composites on the efficient frontier and thus had higher efficiencies (Cooper, Seiford & Tone, 2007).

4.5 Scale Efficiency and Returns-to-Scale

Based on the CCR and BCC scores, the study defined the scale efficiency (SE) of DMU k as the ratio of the CCR efficiency of DMU k over the BCC efficiency of DMU k. This ratio was defined as

$$SE = \frac{{\eta_{CCR}}^*}{{\eta_{BCC}}^*} \tag{4.12}$$

This SE would exceed one when using CCR and BCC output-orientated models to calculate TE and PTE. The SE assumed a value between unity and infinity, with unity indicating full SE. This efficiency indicated the estimated proportion by which the scale of operations must be adjusted to achieve the optimal scale of operations and thus full SE for DMU k (Cooper, Seiford & Tone, 2007).

The SE, together with the PTE, formed the drivers of the TE. Using the relationship of $TE = PTE \times SE$, the decomposition of TE identified all sources of efficiency. Determining the SE, in addition to TE and PTE, was required in order to provide a comprehensive analysis of the efficiencies of the DMUs being assessed (Cooper, Seiford & Tone, 2007).

To correct for scale inefficiency, the determination of returns-to-scale of a DMU was necessary. If DMU k was scale inefficient and DMU k experienced increasing returns-to-scale (IRS), then further investment in the scale of the operations was required. The decimal amount by which inefficient DMU k's scale efficiency exceeded one multiplied by 100 constituted the percentage by which DMU k must increase the scale of operations to experience full SE. If DMU k was scale inefficient and DMU k experienced decreased returns-to-scale (DRS), then DMU k should reduce its operations to experience full SE. Scaling back on operations involved a decrease in investment within the DMU's operations (Cooper, Seiford & Tone, 2007).

To determine the nature of the returns-to-scale of a DMU, the method proposed by Färe *et al.*, (1994) was employed. This method required comparing the efficiencies of three DEA models. These were the CCR, the BCC and the non-increasing returns-to-scale (NIRS) model.

The NIRS model used linear programming to calculate the dual weights which then allowed for the determination of the efficiency of a DMU. The output-orientated formulation of the NIRS linear programming model evaluated K DMUs using the same m inputs x_t $(t = 1, 2, \dots, m)$, to produce the same n outputs y_s $(s = 1, 2, \dots, n)$ was given as

$$\max_{\eta_{NIRS}, \lambda_i} \eta_{NIRS} \tag{4.13}$$

subject to

$$x_{tk} - \sum_{i=1}^{K} \lambda_i x_{ti} \geq 0 \; (t = 1, 2, \dots, m) \tag{4.14}$$

$$\eta_{NIRS} y_{sk} - \sum_{i=1}^{K} \lambda_i y_{si} \leq 0 \; (s = 1, 2, \dots, n) \tag{4.15}$$

$$0 \leq \sum_{i=1}^{K} \lambda_i \leq 1 \tag{4.16}$$

$$\lambda_i \geq 0 \text{ for all } i \tag{4.17}$$

where x_{tk} represented the quantity of input t used by DMU k, and y_{sk} denoted the quantity of output s, produced by DMU k (Brettenny & Sharp, 2016).

Efficiency of DMU k was calculated in the NIRS model by the construction of an efficient composite of DMU k using dual weights λ_i and the reference set identified in the linear programme. DMU k was then radially compared to its efficient composite, on the efficient frontier, to determine its efficiency (El-Mahgary & Lahdelma, 1995). Thus, the NIRS model calculated efficiency estimates in the same manner as the CCR and BCC models.

Therefore, solving the linear programme above provided an estimated proportion $\eta_{NIRS}{}^*$, for which outputs should be increased for each of the K DMUs. The $\eta_{NIRS}{}^*$ value for DMU k was referred to as its non-increasing returns-to-scale (Farrell) efficiency (NIRSE). The Farrell efficiency measure stated that $[(1 - \eta_{NIRS}{}^*) \times 100]$ was equivalent to the percentage by which DMU k must increase its outputs, while maintaining inputs, to become relatively efficient. A DMU was deemed NIRS efficient if the solution was $\eta_{NIRS}{}^* = 1$ and all associated slacks (s^-, s^+) were equal to zero (Brettenny & Sharp, 2016). For each assessed DMU, the slacks were described as the excesses of inputs and/or shortfalls in outputs which could be required in addition to the proportional increase in outputs by the factor $\eta_{NIRS}{}^*$. The slacks were applied, when necessary, in the same manner as in the CCR and BCC model (Cooper, Seiford & Tone, 2007).

The difference between the NIRS and BCC models was that the NIRS efficiencies were calculated by extending the BCC model. This was done by relaxing the convexity assumption to $0 \leq \sum_{i=1}^{K} \lambda_i \leq 1$. This created an efficient frontier with less convexity than the efficient frontier in the BCC model, but without the stringent linear frontier of the CCR model. Thus, the NIRS efficient frontier was located between the BCC and CCR frontiers. The NIRS model had its efficient frontier spanned by efficient DMUs exhibiting DRS. This model put emphasis on larger DMUs where returns-to-scale were decreasing (Cooper, Seiford & Tone, 2007).

The relaxation of the convexity assumption in the BCC model to establish the NIRS model could potentially alter the dual weight values λ_i, the reference sets, and thus efficiencies of the NIRS model in comparison to those of the BCC model. How alike or different the BCC and NIRS efficiencies were of importance in the determination of the returns-to-scale.

Scale inefficient DMUs returns-to-scale could be obtained by comparing the efficiency measure derived from the NIRS and BCC models. DMU k experienced IRS when $PTE > NIRSE$ and DRS when $PTE = NIRSE \neq 1$. CRS was experienced when DMU k was at its most productive scale size (MPSS) when, $PTE = TE = SE = 1$. CRS could also be experienced when a scale efficient DMU k was not experiencing the most productive scale size as in the case of $PTE = TE = NIRSE \neq (SE = 1)$ (Camanho & Dyson, 1999).

All DMUs aim to perform at MPSS. When MPSS is achieved DMU k experienced complete efficiency. The TE, PTE and SE are subcomponents of MPSS. Therefore, these efficiencies are an integral part of achieving the aim of all container terminals which is MPSS. This study calculated the TE, PTE and SE to determine whether MPSS was achieved by a terminal. If MPSS was not achieved, these efficiencies were used to suggest corrective procedures to aid in the achievement of MPSS. At MPSS a terminal experienced full TE. The TE represented a global efficiency as it was decomposed into PTE and SE. Thus, PTE and SE were imperative in achieving and correcting for MPSS. The PTE and SE were highlighted when analysing the results in Chapter 6.

4.6 Bootstrapping in DEA

The DEA method used a sample for the analysis of efficiency. However, as a deterministic method, DEA did not explicitly model the random sampling error associated with its efficiency estimates. The DEA method simply interpreted the overall deviation from the frontier as inefficiency only. However, this deviation was driven by both the variability (sampling error) and the location (inefficiency). As a result, the accuracy of the DEA efficiency estimates may have been affected by sampling variation.

In multi-output and multi-input DEA models, the bootstrap methodology is a way to investigate the sampling variability present in the efficiency estimates (Hung, Lu & Wang, 2010). Bootstrapping is based on the idea of resampling from the original values to create replicate datasets from which sampling error can be identified and corrected for (Martınez-Nunez & Perez-Aguiar, 2014).

The Simar and Wilson (2000) method of homogenous bootstrapping was adopted. This method assumes that the bootstrap distribution of efficiencies will imitate the original unknown distribution of efficiencies. To establish the bootstrap distribution of efficiencies, the smooth homogenous bootstrapping approach non-parametrically estimates the densities of the efficiency scores using kernel smoothing methods, combined with a reflection method (Martınez-Nunez & Perez-Aguiar, 2014).

For the DEA approach, the complete bootstrap algorithm, used to determine bias-corrected efficiency estimates, was summarised by the following steps (Hung, Lu & Wang, 2010; Martınez-Nunez & Perez-Aguiar, 2014).

Step 1:
Using the original data set, compute the original efficiency scores η^* (for the respective year) for each of the K DMUs using the CCR, BCC and/or NIRS model(s).

Step 2:
Establish the symmetric set D_{2k} through a reflection method. This was achieved by combining η_i^* (original DEA efficiency scores) values and the $(2-\eta_i^*)$ (a reflection of the original DEA efficiency scores) values, where $i = 1, \dots, K$. The set D_{2k} was thus presented as

$$D_{2K} = \{\eta_1^*, \dots, \eta_K^*, (2 - \eta_1^*), \dots, (2 - \eta_K^*)\} \tag{4.18}$$

Generate a random sample β_i^*, where $i = 1, \dots, K$, by drawing with replacement from the reflected set D_{2k}.

Step 3:
Generate the kernel smoothed efficiencies $\tilde{\theta}_i^*$, for $i = 1, \dots, K$, using

$$\tilde{\theta}_i^* = \begin{cases} \beta_i^* + h\varepsilon_i^* & if\ \beta_i^* + h\varepsilon_i^* \leq 1 \\ 2 - (\beta_i^* + h\varepsilon_i^*) & otherwise \end{cases} \tag{4.19}$$

In this way one obtains the smoothed bootstrap replicates, $\tilde{\theta}_i^*$, which is equivalent to sampling from the kernel smoothed density constructed from the values in the reflected set D_{2K}. The value for h in the kernel density function equation was by rule-of-thumb, as introduced by Silverman (1986), $h = 1.06 s_{\beta_i^*} K^{-1/5}$, where $s_{\beta_i^*}$ represented the sample standard deviation of the observations used to estimate the density, and where ε_i^* was a random variant drawn from the standard normal distribution. The above value for h provided a control parameter which aided the construction of a non-parametric normal kernel density function over the symmetric distribution of D_{2K}.

Step 4:

Compute θ_i^* for $i = 1, \dots, K$, where

$$\theta_i^* = (1/K)\sum_{i=1}^{K}\beta_i^* + \frac{1}{\sqrt{1 + h^2/s_{\beta_i^*}^2}}\left[\tilde{\theta}_i^* - (1/K)\sum_{i=1}^{K}\beta_i^*\right] \tag{4.20}$$

Step 5:

Generate resampled pseudo-efficiencies γ_i^* using

$$\gamma_i^* = \begin{cases} 2 - \theta_i^*, & if\ \theta_i^* < 1 \\ \theta_i^*, & otherwise \end{cases}. \tag{4.21}$$

The pseudo-efficiencies transformed the data back to within the original range of the output-orientated efficiencies between $[1, \infty]$.

Step 6:

Obtain and define the bootstrap sample $\chi_b^* = \{(x_i, y_{ib}^*) | i = 1, \dots, K\}$, where $y_{ib}^* = (\gamma_i^*/\eta_i^*)y_i$. Thus, x_i remained fixed and outputs were shifted by γ_i^*/η_i^* along a ray passing through y_i and the origin. The bootstrap sample χ_b^* was used to construct a new frontier against which the original sample was compared.

Step 7:

Calculate the bootstrapped DEA efficiency score $\eta_{i\,(boot)}^*$ for each of the K DMU's, (x_i, y_i), using the frontier created by the (x_i, y_{ib}^*) data set.

Step 8:

Repeat steps 2 to 7 B times to create a set with B efficiency estimates for each DMU $\eta^*_{i\,(boot),b}; i = 1, \dots, K; b = 1, \dots, B\}$. B was taken to be 2000 and the mean of the bootstrap replicates $\eta^*_{i\,(boot),b}; b = 1, \dots, B$ will be used to approximate the ideal bootstrap estimate of the expected value of η^*_i. This in order to ultimately obtain an estimate of bias.

Step 9:

The next step required the bias-correction of the DEA efficiency estimates using the bootstrapped efficiencies $\eta^*_{i\,(boot),b}$. The bias was defined by Simar and Wilson (2007) as

$$Bias(\eta^*_i) \equiv E(\eta^*_i) - \eta_i, \text{ where } i = 1, \dots, K. \qquad (4.22)$$

However as the true value of η was unknown, it was only possible to determine an estimate of the bias contained in the original DEA efficiency estimate η^*_i. Using the bootstrapped DEA estimates, the bias estimate was determined by Simar and Wilson (2007) as

$$\widehat{Bias}_B(\eta^*_i) = B^{-1} \sum_{b=1}^{B} \eta^*_{i\,(boot),b} - \eta^*_i, \text{ where } i = 1, \dots, K. \qquad (4.23)$$

Step 10:

A bias-corrected DEA efficiency value was then obtained by defining

$$\eta^*_{i,BC} = \eta^*_i - \widehat{Bias}_B(\eta^*_i) \qquad (4.24)$$

$$\Rightarrow \eta^*_{i,BC} = \eta^*_i - (B^{-1} \sum_{b=1}^{B} \eta^*_{i\,(boot),b} - \eta^*_i) \qquad (4.25)$$

$$\Rightarrow \eta^*_{i,BC} = 2\eta^*_i - B^{-1} \sum_{b=1}^{B} \eta^*_{i\,(boot),b}. \qquad (4.26)$$

This method provided bias-corrected DEA estimates for the set of 2000 bootstrap repetitions. Output-orientated DEA estimates between 1 and infinity were subjected to downward bias. This downward bias is a result of the sample error present in the DEA estimate. Since the modulus of the estimated bias was greater than the estimated standard errors in each analysis, the bias-corrected estimates were preferred to the original DEA scores (Munisamy & Danxia, 2013). Ultimately this provided a ranking method for the container terminals.

4.7 Malmquist Productivity Index

The MPI, proposed by Färe *et al.*, (1994), produced an efficiency change measure referred to as the total factor productivity change (TFPC). The TFPC provided an interpretation of the change in efficiency over time and could be decomposed into three components. These component measurements were; the changes in PTE, the changes in SE, and the final component measures changes in technology.

The TFPC measurement was calculated for each of the K DMU's. The TFPC of DMU k was calculated using ratios of distances. Using DEA, the distances making up these ratios were simply radial comparisons of DMU k from period s to the efficient composite of DMU k, located on an efficient frontier, from period t. Thus, these distances represented by $D_t(Y_s, X_s)$ were efficiencies. By placing these distances into the ratio's that made up the TFPC, the TFPC calculated changes in efficiency. In this study the DEA approach was output-orientated, and as a result so were the efficiencies (Estache, De La Fé & Trujillo, 2004). The output-orientated TFPC between the base period (zero) and the reference period (one) was given by

$$MPI = TFPC = \frac{D_0(Y_0, X_0)}{D_1(Y_1, X_1)}\left[\frac{D_1(Y_0, X_0)}{D_0(Y_0, X_0)} \times \frac{D_1(Y_1, X_1)}{D_0(Y_1, X_1)}\right]^{0.5}. \qquad (4.27)$$

The MPI defined by Färe *et al.*, (1994), as seen in Eq. (4.27), defined the geometric mean of two indices, one evaluated with respect to the reference period technology and the second with respect to the base period technology (Estache, De La Fé & Trujillo, 2004).

By comparing DMU k in period s, to the efficient frontier in period t, the TFPC value accounted for changes in TE and shifts in the efficient frontier. The ratio outside the square brackets in Eq. (4.27) measured the change in the output-oriented measure of TE between periods zero and one. This ratio was the total technical efficiency change (TTEC) measure. The bracketed term of the index in Eq. (4.27) was a measure of technological change (TC) which accounted for the shift in technology between the two periods (Estache, De La Fé & Trujillo, 2004). Thus, the MPI, and the TFPC that it calculated, was decomposed into

$$MPI = TFPC = TTEC \times TC. \tag{4.28}$$

To measure TFPC, both PTE changes and SE changes needed to also be accounted for. Färe *et al.*, (1994) used CRS distance functions to calculate the index in Eq. (4.28). The VRS distance functions was required for further decomposition of Eq. (4.28). The introduction of VRS decomposed the TTEC measure into a PTE change component and a SE change component. The mixture of CRS and VRS distance functions that achieved this decomposition of TTEC was given by Eq. (4.29). In Eq. (4.29), the V superscripts referred to VRS technology and the C superscripts referred to CRS technology.

$$\begin{aligned} MPI = TFPC = \frac{D_0^V(Y_0, X_0)}{D_1^V(Y_1, X_1)} \Bigg\{ & \left[\frac{D_1^V(Y_1, X_1)}{D_0^V(Y_0, X_0)} \times \frac{D_0^C(Y_0, X_0)}{D_1^C(Y_1, X_1)} \right]^{0.5} \\ & \times \left[\frac{D_1^C(Y_0, X_0)}{D_0^C(Y_0, X_0)} \times \frac{D_1^C(Y_1, X_1)}{D_0^C(Y_1, X_1)} \right]^{0.5} \Bigg\}. \end{aligned} \tag{4.29}$$

Eq. (4.29) thus gave a pure technical efficiency change (PTEC) measure, a scale efficiency change (SEC) measure, and maintained the TC measure (Estache, De La Fé & Trujillo, 2004). That is

$$MPI = TFPC = PTEC \times SEC \times TC \tag{4.30}$$

The decomposition in Eq. (4.30) was required for this core study in order to allow for all forms of efficiency changes within the selected container terminals to be analysed.

4.8 Summary

Several important components were identified within the DEA methodology in order to produce efficiency estimates for the container terminals case study. The basic components were identified in the single input-output case. These basic components were expanded upon in the single input-double outputs case and within the multiple variable CCR and BCC models. All of the components would be utilised in the CCR and BCC models when calculating the TE and PTE for the selected African container terminals.

These components included the use of the productivity ratio in Eq. (4.1) which was the basis of all efficiency results in DEA. This productivity ratio was used to construct another important component, namely the efficient frontier. This provided a benchmark against which the relative efficiency of a DMU could be determined. The relative efficiency was another important component of DEA. The DEA allowed one to estimate relative efficiency, as the efficiency contained the property of unit invariance, which was essential when dealing with multiple inputs and outputs. To determine the relative efficiency of a DMU, the CCR and BCC models identified the dual weights and an efficient composite of that DMU. These differing dual weights gave rise to the CRS and VRS properties of the CCR and BCC models. Once an efficient composite had been determined for a particular DMU, the CCR and BCC models identified the proportion by which outputs must be increased for that DMU to resemble its efficient composite. In the CCR and BCC models, these estimated proportions were the TE and PTE, respectively. The output-and input-orientations determined how inefficiency was corrected for radially. This was also an important component, as it determined whether outputs were increased, whilst keeping inputs constant, or vice versa, when correcting for inefficiency.

In order to provide a more meaningful practical interpretation to the efficiencies, the SE of each DMU is required. This in turn requires the determination of the returns-to-scale of each DMU to correct for any scale inefficiency. The BCC efficiencies and NIRS efficiencies had to be obtained to establish these returns-to-scale.

The TE, PTE and SE formed the subcomponents of the MPSS at which all container terminals aimed to operate. These efficiencies were essential to determine whether MPPS existed in a terminal, and if not, were used to make suggestions on how efficiency could be achieved by a terminal. Table 4.4 gives a brief description of MPSS, TE, PTE and SE and how the efficiencies fit into the MPSS.

Table 4.4: MPSS and Efficiency	
Most Productive Scale Size (MPSS)	Where all terminals wanted to be operating. MPSS occurs when $TE = 1.$
TE	How technically efficient the current inputs were in generating outputs when comparing terminals of all scale sizes. TE was a global efficiency and was decomposed into $TE = PTE \times SE.$
PTE	How technically efficient the current inputs were in generating outputs when comparing terminals of equal scale size.
SE	How scale efficient a terminal was in comparison to terminals operating at the optimal scale (CRS).

Once TE, PTE and SE were determined for the DMUs under investigation the sampling variability needed to be accounted for. This was achieved using the smooth homogenous bootstrapping procedure, to produce bias free efficiency results. These provided potential corrections for the efficiency estimates. Finally, to track efficiency changes over the sample period, the MPI was used. The TFPC tracked the efficiency changes for the terms TE, PTE and SE. The TFPC also accounted for the shift in the frontier from one sample period to the next, identified as the technological change (TC).

All of the above methodologies attempted to provide DEA efficiency estimates that can be used to give operational interpretations to the case study data for the selected African container terminals.

Chapter Five: Validation of DEA Code

5.1 Introduction

The analysis for this research was done using R v3.2.3 (R Core Team, 2015). Two R packages, Benchmarking (Bogetoft & Otto, 2011) and Frontier Efficiency Analysis with R FEAR (Wilson, 2008) were used for many of the routines. Considerable coding was necessary to complete this research. To validate the code that was written, this study replicated the results of the El-Mahgary and Lahdelma (1995) paper, entitled "Data envelopment analysis: Visualising the results". This paper was selected for two reasons. Firstly, the results were published in a reputable journal entitled the *European Journal of Operational Research*. Secondly, El-Mahgary and Lahdelma (1995) produced visual displays of the results within the paper. This allowed for ease in comparison between this study's results and those of El-Mahgary and Lahdelma (1995).

Specifically, the dual weights, reference sets and TE results produced by the CCR model, for inefficient DMU's in the El-Mahgary and Lahdelma (1995) paper, were replicated. The dual weights and reference sets are key to establishing the efficient composite of a DMU. This DMU is then radially compared to its efficient composite to determine its efficiency. Thus, replicating results of El-Mahgary and Lahdelma's (1995) paper would justify the accuracy of the code written and lend credibility to the efficiency estimates generated for the African container terminals.

It should be noted that the dual weights, reference sets and efficiencies were replicated for an input-orientated CCR model as used by El-Mahgary and Lahdelma (1995). As opposed to the output-orientation, the input-orientation efficiency would be between zero and unity. This was because the input-orientation efficiency represented the estimated proportion that inputs should be decreased, while keeping outputs unchanged, in order to achieve full efficiency of a particular DMU (Cooper, Seiford & Tone, 2007).

5.2 The Variables and Data of El-Mahgary and Lahdelma (1995)

The El-Mahgary and Lahdelma (1995) paper collected a cross-sectional data set from 20 major Finnish Universities. The input and output variables, their symbols and units of measurement identified as important are listed in Table 5.1.

The inputs used in the study are expenditure and admission. Expenditure is the amount of Finnish Markka spent by each university on education. Admission is the inverse of the acceptance rate of a Finnish university. The reason for using the inverse of the acceptance rate is to make sure inputs were not too large. This prevented the input value from exceeding the output value. This correction was therefore essential for the efficiency calculations (El-Mahgary & Lahdelma, 1995).

The outputs used are; graduates, post-graduates, graduation speed, and completion. Graduates and post-graduates indicated the number of graduate and post-graduate degrees granted. Graduation speed is measured by the number of years spent acquiring a graduate degree. El-Mahgary and Lahdelma (1995) note in their research that in Finnish universities there was no fixed time span for acquiring the graduate degree, resulting in, many students taking breaks, for various reasons, from their studies. As a result, the inverse of the median (being more robust to fluctuations than the mean), time taken to complete a degree, is used. Again, the inverse is used to keep the input values small, for efficiency calculation purposes. Completion indicated the number of students who finished their graduate degrees. This output is measured using the inverted drop-out rate for a period of six years, which was the typical time taken to complete a graduate degree (El-Mahgary & Lahdelma, 1995).

Table 5.1: Variables in the Sample Analysis

Factor	Symbol	Type	Units
Expenditure	x_1	Input	Millions FIM (Finnish markka)
Admission	x_2	Input	Scalar
Graduates	y_1	Output	Quantity
Post-graduates	y_2	Output	Quantity
Graduation speed	y_3	Output	$1/_{years}$
Completion	y_4	Output	Scalar

Source: Direct extract from (El-Mahgary & Lahdelma, 1995)

Table E in Appendix Three of this study lists each of the 20 major Finnish universities and the values of their respective input and output variables.

5.3 Replication of El-Mahgary and Lahdelma (1995) Dual Weights, Reference Sets and Technical Efficiency

The dual weights were assigned to the inputs and outputs of the reference set DMUs by the CCR model. Weighting and summing the inputs and outputs of a particular DMU's reference set constructed an efficient composite of that DMU. That DMU was then radially compared to its efficient composite by the linear programme to determine the estimated proportion by which the inputs should be changed, while at the same time controlling for outputs. This estimated proportion was the TE of the input-orientated CCR model.

These input orientated CCR model dual weights, reference sets and technical efficiencies were generated for inefficient Finish universities by El-Mahgary and Lahdelma (1995). These estimates are listed in Table 5.2. Each DMU also had a letter assigned to it, as can be seen in Table E of Appendix Three. There was no need to showcase the dual weights, reference sets and TE of the fully efficient Finnish universities, as fully efficient DMUs were the efficient composite DMUs of themselves. As a result, their dual weights were one, they were their own reference set and they all had TE equal to one (El-Mahgary & Lahdelma, 1995).

Table 5.2: Reference Sets, TE and Dual Weights of El-Mahgary and Lahdelma (1995)

DMU	Reference Set	TE	Dual Weights
E	{M, Q}	0.800	{0.67, 0.297}
J	{M, Q}	0.710	{0.0016, 1.1}
K	{Q}	0.770	{0.82}
O	{D, Q}	0.710	{0.315, 0.65}
R	{Q, T}	0.460	{0.71, 0.071}
S	{**H**, P, Q}	0.360	{0.04, 0.398, 0.301}

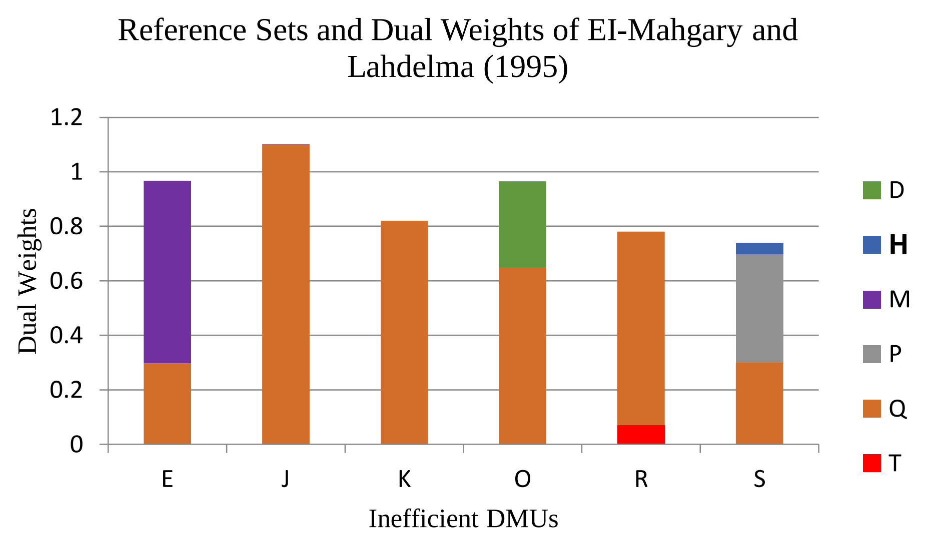
Source: Direct extract from (El-Mahgary & Lahdelma, 1995)

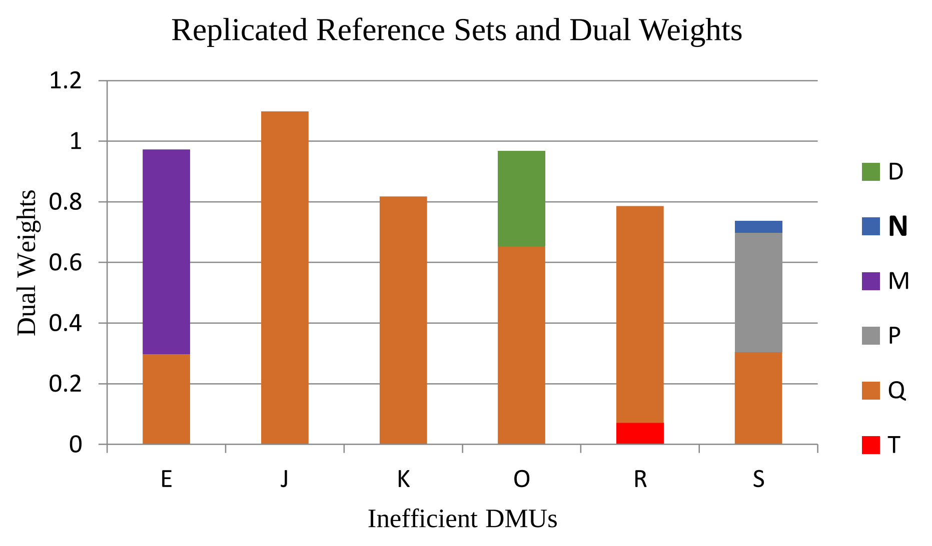
Table 5.3 details this study's replication of the dual weights, reference sets and technical efficiencies generated by El-Mahgary and Lahdelma (1995).

Table 5.3: Replicated Reference Sets, TE and Dual Weights			
DMU	**Reference Set**	**TE**	**Dual Weights**
E	{M, Q}	0.799	{0.676, 0.297}
J	{M, Q}	0.707	{0.00178, 1.097}
K	{Q}	0.772	{0.818}
O	{D, Q}	0.705	{0.315, 0.654}
R	{Q, T}	0.455	{0.714, 0.0716}
S	{**N**, P, Q}	0.361	{0.04, 0.392, 0.305}

A graphical representation of the dual weights and reference sets of Table 5.2 and Table 5.3 is presented in Figure 5.1. This allowed for a visual comparison of the El-Mahgary and Lahdelma (1995) results and their replications. The y-axes in the figures indicate dual weight values. As indicated by the x-axes, each bar in these figures represented an inefficient Finnish university identified. Each bar in these figures were sub-compartmentalised by their reference sets. The particular Finnish universities that made up these reference sets were identified by the key on the left of the figures. The proportion of the total bar that a reference set DMU would constitute was associated with the weight assigned to that reference set DMU by the CCR model.

Figure 5.1: Reference Sets and Dual Weights from El-Mahgary & Lahdelma (1995) vs. Replicated Analysis





Source: Direct extract from (El-Mahgary & Lahdelma, 1995)

The replication of the dual weights, reference sets and technical efficiencies were accurate for all inefficient Finnish university except for the Sibelius Academy, labelled S. Within S, the technical efficiency and dual weight values were replicated by this study accurately. The difference between the El-Mahgary and Lahdelma (1995) dual weights, reference sets and technical efficiencies and those replicated was the allocation of reference set DMUs to S. This study allocated the fully efficient N, P and Q DMUs to the reference set of S. El-Mahgary and Lahdelma (1995) allocated the fully efficient DMUs H, P and Q to the reference set of S. Thus, the difference was one reference set DMU, namely the allocation of N instead of H by this study, to S. This difference was due to a discrepancy present within the El-Mahgary and Lahdelma (1995) paper. The reasoning for this was that the numerical efficiency value of S was replicated by this study. Additionally, despite this study using N instead of H as a reference set DMU, the study still managed to allocate the same numerical dual weight of H to N.

It is believed that this small discrepancy might have arisen from the AskDEA package used by El-Mahgary and Lahdelma (1995). The researchers state that the AskDEA package is used as an experimental tool within their paper. The “Benchmarking” (Bogetoft & Otto, 2011) and “FEAR” (Wilson, 2008) packages used to replicate the El-Mahgary and Lahdelma (1995) results in R v3.2.3 (R Core Team, 2015) are more robust. The reason being that these packages are used more frequently throughout the DEA literature and thus have been validated by others. It is also possible that El-Mahgary and Lahdelma (1995) may have mislabelled the reference set DMU H instead of N. This maybe a typographical error as H and N are adjacent keys on the keyboard. This could probably be the more likely reason.

This section managed to replicate the dual weights, reference sets and TE results produced by El-Mahgary and Lahdelma (1995). As a result, this study has confidence in the efficiency results produced by its code.

5.4 Summary

The technical efficiencies, dual weights and reference sets of the inefficient Finnish universities in the El-Mahgary and Lahdelma (1995) were replicated. The only discrepancy was the allocation of reference sets for Sibelius Academy. This study allocated P and Q as reference set DMUs of S. The same allocation was made by El-Mahgary and Lahdelma

(1995). However, this study allocated N (in)correctly to the reference set of S as opposed to the allocation of H by El-Mahgary and Lahdelma (1995).

This discrepancy is believed to be a typographical error from the El-Mahgary and Lahdelma (1995) study. The reasons for this conclusion is that technical efficiencies, dual weights and all the reference set allocation generated by El-Mahgary and Lahdelma (1995) were replicated. The discrepancy may also have originated either from the experimental AskDEA package used to generate the results by El-Mahgary and Lahdelma (1995). The replication of the results in the El-Mahgary and Lahdelma (1995) study provided sufficient evidence to conclude that the code developed for this study was valid.

Chapter Six: Results and Analysis

6.1 Introduction

The results of the CCR and BCC models were presented in terms of the TE and PTE estimates, respectively. In addition to these estimates, the reference sets and dual weights associated with each DMU were shown. These results were presented for each one of the 15 selected container terminals over the two year sample period.

The technical (global) efficiency estimates of the 15 selected container terminals were identified. The TE estimates were divided into two components, namely PTE and SE. The study believed that PTE was the leading component in achieving full TE, and thus MPSS. The reason being that PTE estimates could be corrected without costly capital investment. The PTE indicated how technically efficient the current inputs were in generating outputs when comparing terminals of equal scale size. Thus, the terminal had only to make sure that existing inputs produced outputs as efficiently as possible. This was not the case when correcting for scale efficiency, which required capital investment. As not all the terminals may have had the capacity for capital investment, PTE was seen as the most important contributor to the global efficiency, or TE, of the port. Scale efficiency was considered as a secondary contributor to TE of a terminal.

As a result, the bootstrapped pure technical efficiency estimates were used to establish a descending efficiency ranking of the selected African container terminals for each of the sample periods. The focus of this analysis was on how the highest ranked ports, some globally efficient and some not, and the lowest ranked ports could improve their current inefficiencies. This was done by suggesting corrections be made to either the PTE or SE inefficiencies, driving the low ranking or preventing higher efficiency in some of inefficient high ranking terminals. The primary objective was to achieve MPSS. In addition to the high and low ranking ports, focus was placed on the Eastern Cape ports, which was where this study originated.

The TFPC was also analysed to showcase container terminals that improved or decreased their efficiency scores over the sample period. The components of the TFPC were analysed to determine what was driving these improvements or decreasing their efficiency.

6.2 Efficiency, Reference Set and Dual Weight Estimates

Data from 15 of the major container terminals on the African continent were collected. The information from these container terminals was classified as panel data as it was acquired for the years 2013 and 2014. The data was acquired from multiple sources online and no single source can be identified as providing the majority of the data.

The TE, PTE, reference sets and dual weight (λ) estimates for the selected container terminals are listed in Table 6.1. This table summarises the results of these estimates for each of the 15 container terminals. Table 6.1 used abbreviated names for the container terminals and thus refers the reader back to Table 3.1 for the full names. Each DMU had a number corresponding to it as well as a TE and PTE estimate. Every efficiency estimate had a corresponding dual weight and reference set DMU estimate. The reference set DMUs were identified by the number allocated to the DMU. All these estimates were generated by the "Benchmarking" (Bogetoft & Otto, 2011) package in R v3.2.3 (R Core Team, 2015). The estimates were generated for both 2013 and 2014.

The PECT was isolated to provide a brief interpretation of the results in Table 6.1. In 2013, PECT had a TE and PTE equivalent to 2.747 and 1.000, respectively. The reference set associated with the TE result comprised of the Port Said SCCT with a dual weight of 0.194. The reference set associated with the PTE was PECT itself as it experienced full PTE. In 2014, PECT had a TE and PTE equivalent to 2.924 and 1.000, respectively. Thus, there was a slight decrease in the TE. The reference set associated with the TE result comprised of the Port Said SCCT and Tanger Med with dual weights of 0.184 and 0.003, respectively. The reference set associated with the PTE was again PECT itself as it experienced full PTE in 2014.

No analysis was reported on these estimates at this stage. Table 6.1 was provided purely as a report of the dual weight and reference set results. These components were identified as important in the methodology of the CCR and BCC model and as such seen as important to inform the reader of their outcomes. The ranked African container terminals were analysed in terms of the PTE and SE in section 6.3. Suggestions are also made, within this section, on how inefficiencies could be corrected to achieve MPSS.

Table 6.1: Technical, Pure Technical and Non-Increasing Returns-to-Scale Efficiencies

DMU	Container Terminal	2013				2014			
		TE	Reference Sets (λ)	PTE	Reference Sets (λ)	TE	Reference Sets (λ)	PTE	Reference Sets (λ)
1	Alexandria (AICT)	1.193	9 (0.182), 13 (0.172)	1.000	1 (1.000)	1.297	13 (0.052), 14 (0.225)	1.000	1 (1.000)
2	Cape Town CT	2.123	6 (0.281), 9 (0.214), 13 (0.378)	2.081	3 (0.077), 9 (0.606), 13 (0.317)	2.398	13 (0.067), 14 (0.606)	2.222	1 (0.207), 9 (0.298), 14 (0.495)
3	Casablanca CT	1.263	6 (0.287), 13 (0.202)	1.000	3 (1.000)	1.355	6 (0.043), 14 (0.362)	1.000	3 (1.000)
4	Damietta CT	2.358	13 (0.278), 14 (0.189)	1.814	1 (0.064), 9 (0.827), 13 (0.109)	2.357	13 (0.236), 14 (0.227)	1.972	1 (0.282), 9 (0.500), 14 (0.218)
5	Dar es Salaam (TICTS)	1.890	9 (0.882), 13 (0.059)	1.866	1 (0.091), 9 (0.866), 13 (0.043)	1.825	9 (0.783), 14 (0.117)	1.777	1 (0.130), 9 (0.797), 14 (0.073)
6	Doraleh CT	1.000	6 (1.000)	1.000	6(1.000)	1.000	6 (1.000)	1.000	6 (1.000)
7	Durban (Pier 1 and 2)	1.461	9 (2.285), 13 (0.236), 14 (0.311)	1.269	9 (0.238), 13 (0.762)	1.647	9 (1.889), 14 (0.778)	1.295	13 (0.364), 14 (0.636)
8	Apapa CT	2.551	9 (0.145), 13 (0.195), 14 (0.284)	2.548	6 (0.225), 9 (0.362), 13 (0.067), 14 (0.346)	1.996	6 (0.342), 9 (0.476), 14 (0.378)	1.990	6 (0.228), 9 (0.310), 14 (0.462)
9	Luanda (CT2)	1.000	9 (1.000)	1.000	9 (1.000)	1.000	9 (1.000)	1.000	9 (1.000)
10	Mombasa CT	1.553	9 (0.976), 13 (0.033), 14 (0.142)	1.520	9 (0.833), 13 (0.111), 14 (0.056)	1.601	6 (0.057), 9 (1.032), 14 (0.158)	1.542	9 (0.750), 14 (0.250)
11	Ngqura CT	1.952	9 (0.857), 14 (0.286)	1.916	9 (0.722), 13 (0.074), 14 (0.204)	2.930	9 (0.500), 14 (0.500)	2.930	9 (0.500), 14 (0.500)
12	Port Eliz. CT	2.747	13 (0.194)	1.000	12 (1.000)	2.924	13 (0.184), 14 (0.003)	1.000	12 (1.000)
13	Port Said SCCT	1.000	13 (1.000)	1.000	13 (1.000)	1.000	13 (1.000)	1.000	13 (1.000)
14	Tanger Med (T1 and T2)	1.000	14 (1.000)	1.000	14 (1.000)	1.000	14 (1.000)	1.000	14 (1.000)
15	Tema Port CT	1.000	15 (1.000)	1.000	15 (1.000)	1.089	6 (0.673), 14 (0.131)	1.000	15 (1.000)

6.3 Ranking and Analysis

The selected African container terminals were ranked for both 2013 and 2014 of the sample period using the smooth homogenous bootstrap. The pure technical efficiencies were bootstrapped to establish bias-corrected PTE values. The container terminals were listed in descending order based on these bias-corrected pure technical efficiencies.

The bias-corrected PTE values were better suited to ranking than the pure technical efficiencies. The bootstrapping procedure provided a method for distinguishing between DMU's with equal PTE values. This was achieved by accounting for sampling variation within the bias-corrected PTE.

In addition to being a valuable ranking tool, the bias-corrected efficiencies indicated further potential adjustments in PTE for a ranked DMU to obtain full PTE. These adjustments were in addition to those proposed by the original, pure technical efficiencies. As a result of the downward bias present in the original PTE, the bias-corrected PTE advocated greater adjustments in output to achieve full PTE efficiency where necessary. However, the original PTE and SE values remained the core result, in terms of achieving MPSS and thus technical efficiency for a ranked terminal. The bias-corrected PTE indicated potential corrections in efficiency, above and beyond the PTE and SE corrections needed for a ranked terminal to achieve MPSS.

Table 6.2 lists the rankings of the 15 selected African container terminals in descending order over the sample period. The bias-corrected PTE values were calculated using the "FEAR" (Wilson, 2008) package in R v3.2.3 (R Core Team, 2015). The table also lists the PTE and SE values for both 2013 and 2014 as well as recording the NIRSE and returns-to-scale results.

Table 6.2: Efficiency Rankings, Scale Efficiency & Returns-to-Scale for 2013 & 2014.

	2013						2014					
Container Terminal	**Rank**	**Bias-corrected PTE**	**PTE**	**NIRSE**	**SE**	**RTS**	**Rank**	**Bias-corrected PTE**	**PTE**	**NIRSE**	**SE**	**RTS**
Tema Port CT	**1**	1.214	1.000	1.000	1.000	CRS	**1**	1.227	1.000	1.089	1.089	IRS
Tanger Med (T1 and T2)	**2**	1.231	1.000	1.000	1.000	CRS	**3**	1.254	1.000	1.000	1.000	CRS
Port Said SCCT	**3**	1.258	1.000	1.000	1.000	CRS	**2**	1.253	1.000	1.000	1.000	CRS
Alexandria (AICT)	**4**	1.263	1.000	1.193	1.193	IRS	**4**	1.289	1.000	1.297	1.297	IRS
Port Eliz. CT	**5**	1.273	1.000	2.747	2.747	IRS	**8**	1.309	1.000	2.924	2.924	IRS
Casablanca CT	**6**	1.275	1.000	1.263	1.263	IRS	**5**	1.299	1.000	1.355	1.355	IRS
Luanda (CT2)	**7**	1.277	1.000	1.000	1.000	CRS	**7**	1.305	1.000	1.000	1.000	CRS
Doraleh CT	**8**	1.278	1.000	1.000	1.000	CRS	**6**	1.304	1.000	1.000	1.000	CRS
Durban (Pier 1 and 2)	**9**	1.471	1.269	1.269	1.151	DRS	**9**	1.526	1.295	1.295	1.272	DRS
Mombasa CT	**10**	1.733	1.520	1.520	1.022	DRS	**10**	1.793	1.542	1.542	1.038	DRS
Damietta CT	**11**	2.029	1.814	2.358	1.300	IRS	**12**	2.234	1.972	2.353	1.195	IRS
Ngqura CT	**12**	2.173	1.916	1.916	1.019	DRS	**15**	3.480	2.930	2.930	1.000	CRS
Dar es Salaam (TICTS)	**13**	2.185	1.866	1.890	1.012	IRS	**11**	2.111	1.777	1.825	1.027	IRS
Cape Town CT	**14**	2.402	2.081	2.123	1.020	IRS	**14**	2.535	2.222	2.398	1.079	IRS
Apapa CT	**15**	2.916	2.548	2.548	1.001	DRS	**13**	2.322	1.990	1.990	1.003	DRS

The results of the analysis showed that Tema was a fully efficient port for both 2013 and 2014, with a PTE of 1.000 in both years. Thus, it appears that best practises were in place within this port's operations. However, the error within the sampling process, which was estimated by the bootstrapping procedure, indicated that there was scope for improvement. The result of the bias-correction estimated this error. The bias-corrected estimates are shown in Table 6.2, under the column labelled as "Bias-corrected PTE". Considering year 2013, the study observed that the bias-corrected estimate increased from 1.000 to 1.214. This increase was an indication that the error within the sampling process was estimated at 21.4%. Given that the study was using an output-orientated analysis, this change indicated that the port operations were potentially inefficient due to sampling error, implying that the outputs could have been increased for port operations to become fully efficient. Thus, Tema could have potentially increased its container throughput by approximately 21% in 2013. The same analysis of sampling error took place in 2014 and thus, potentially Tema could have increased its container throughput by approximately 23%. This port was fully scale efficient and experienced CRS in 2013. In 2014 this port experienced a SE of 1.089. Given the IRS nature of the port, in 2014, there was the potential to increase the scale of operations by 8.9%, in order to achieve the MPSS.

Tema's top ranking and full PTE, for both years, came as a result of the terminal handling 85% of Ghana's trade, with coffee, cotton and fruit the major trade products. The labourers within the terminal experienced extremely high levels of training with the private sector being heavily involved in the labour practises (Ghana Ports & Harbour Authority, 2013). Container cargo constituted 80% of the Tema Port's business and the terminal also served as a gateway for trade to Mali, Niger and Bukina Faso (Ghana Ports & Harbour Authority, 2013).

Apapa container terminal had the lowest PTE of 2.548 in 2013. This terminal needed to increase output levels by 154.8%, to ensure best practises were in place within its operations. In addition, the bias-corrected PTE efficiency value indicated that the port operations were potentially inefficient by a further 36.8%. This estimate was the difference between 2.548 and 2.916, due to sampling error. The SE score of 1.001 indicated there was little potential to improve the scale of the operations. The port exhibiting a very slight DRS nature. Focus needed to be placed on doubling its 2013 level of outputs to achieve full PTE.

The bottom ranking and lowest PTE of Apapa in 2013 was possibly due to the lack of rail and road infrastructure leading in and out of the port. Trucks stand in lengthy queue's waiting to off and up load containers as a result of no rail infrastructure in the Apapa container terminal (Rosendahl, 2014).

Ngqura had the lowest PTE of 2.930 in 2014. Therefore, Ngqura needed to increase output levels by 193%, to ensure best practises were in place within the port's operations. The bias-corrected PTE efficiency value indicated that port operations were potentially further inefficient by 55%. Ngqura was fully scale efficient in 2014, as indicated by the CRS of the port. Thus, there was very little potential to increase the scale of operations and the focus needed to be placed on tripling outputs in order to achieve full PTE.

The lowest ranking and PTE results for Ngqura in 2014 were not surprising as this was a new port which only began operations in October 2009. Skills development was still in the early stages and technological expertise still being introduced (Ports & Ships, 2012a). In addition, in 2014, South Africa experienced considerable labour force disputes with widespread striking affecting operations (de Bruyn, 2014).

The Port Elizabeth container terminal (PECT) was the only South African terminal with full PTE of 1.000 in both 2013 and 2014. This terminal therefore had the best practices in place in terms of port operations within South Africa over the study period. However, as indicated by the bias-corrected PTE efficiencies of 1.273 (2013) and 1.309 (2014), there was potential for a 27.3% increase and 30.9% increase in outputs, respectively, due to sampling error. PECT had the lowest scale efficiencies of 2.747 and 2.924 in 2013 and 2014, respectively. Given the IRS nature of the port in both years, this implied there was the potential to increase the scale of operations by 174.7% and 192.4% in 2013 and 2014, respectively, in order to achieve MPSS.

PECT had the highest ranking, relative to the other South African container terminals, in this study for both 2013 and 2014. The achieved ranking and full PTE, in both years, came as a result of high skills levels and technological expertise built up over many years of automotive, citrus and manganese exports. The automotive exports were from Volkswagen, Ford, Mercedes-Benz (Daimler Chrysler) and General Motors. The citrus exports were derived from the seasonal markets during May to October. These forms of trade have been operational since the container and automotive terminal was opened in 1993 (Ports & Ships, 2012b). PECT also played the major role in the export of Manganese ore. The primary destination was the Far East, and demand had climbed from 2.1 million tons per annum in 2005 to 7.5 million tons per annum in 2012. The modest forecasted growth showed this export was expected to approach 16 million tons per annum by 2018 (van Tonder, 2014).

The results of this study rank the Durban container terminal 9th out of the 15 African terminals in both 2013 and 2014. In the South African context, Durban would be ranked 2nd of the four terminals in both years. These findings dispute the 2014 rankings by Maersk (Hutson, 2014a) which claim that Durban is the most efficient container terminal in Africa.

The difference in rankings may be a result of the Maersk study relying on a single key performance indicator (KPI), whilst this current study utilised four input and one output variable. The Maersk KPI for efficiency measurement is "crane moves per hour" which is limited to only one part of a terminal's operations, namely the crane operations on the berth. This measurement did not consider the yard where the containers were organised and stacked, nor had the measurement considered the berthing capacity of the terminal. The efficiency measurements reported in this study have accounted for more sectors of Durban's terminal operations than the Maersk KPI. For this reason the efficiency results here were argued to be a more realistic indicator of actual port operations.

Durban had a PTE of 1.269 and 1.295 in 2013 and 2014, respectively. Thus, Durban would have had to increase output levels by 26.9% in 2013 and 29.5% in 2014 to ensure best practises were in place within the port's operations. The bias-corrected PTE efficiencies of 1.471 in 2013 and 1.526 in 2014 account for the sampling error present in DEA. These values indicated that there was potential for a further 20.2% increase in outputs in 2013 and 23.1% increase in outputs in 2014.

This surprisingly low ranking and low PTE of Durban, in both years, is possibly a result of the port looking to expand the input infrastructure of its terminals with long-term benefits in mind. State-owned logistics utility Transnet recently invested R2bn to enlarge the Durban port (Barradas, 2013; Mkhize, 2014). Thus, capitalisation or investment in inputs has taken preference over operating at the optimal scale in the interim. Durban had an SE of 1.151 in 2013 and 1.272 in 2014. There was potential, therefore, as indicated by the DRS nature of the port, to decrease the scale of operations to achieve MPSS. However, the nature of the returns-to-scale should be reassessed once the new infrastructure is operational.

The North African ports appeared to be more efficient than the other ports. There were four and five in the top seven ranked container terminals, in 2013 and 2014, respectively. The North African ports were all located on the major Asia - Europe trade route and formed the gateway to Africa, resulting in more frequent trading with Europe and other parts of the world (Fourie, 2011). This may have resulted in their high relative efficiency scores.

6.4 Changes in Efficiency over the Sample Period

The results of the TFPC, as well as its sub-components, are listed in Table 6.3. A value of one or more for the TFPC or any of its components indicated an improvement in that source of efficiency over the analysed time period. A value lower than one indicated deterioration over the analysed time period. As an example, a value of 1.025 corresponded to a 2.5% increase and a value of 0.95 corresponded to a 5% decrease over the sample period.

Table 6.3: Malmquist Productivity Index Results

African Container Terminal	TC(1)	PTEC(3)	SEC(4)	TTEC(2)= (3) x (4)	TFPC(5)= (1) x (2)	% Increase/ (decrease) Efficiency
North African Terminals						
Alexandria (AICT)	1.044	1.000	0.919	0.919	0.960	(4%)
Casablanca CT	1.100	1.000	0.933	0.933	1.026	2.6%
Damietta CT	1.027	0.920	1.088	1.001	1.028	2.8%
Doraleh CT	1.067	1.000	1.000	1.000	1.067	6.7%
Port Said SCCT	0.974	1.000	1.000	1.000	0.974	(2.6%)
Tanger Med (T1 and T2)	1.203	1.000	1.000	1.000	1.203	20.3%
West African Terminals						
Apapa CT	1.167	1.281	0.998	1.279	1.492	49.2%
Tema Port CT	1.112	1.000	0.918	0.918	1.021	2.1%
East African Terminals						
Dar es Salaam (TICTS)	1.133	1.050	0.986	1.036	1.173	17.3%
Mombasa CT	1.166	0.985	0.985	0.971	1.132	13.2%
Southern African Terminals						
Ngqura CT	1.185	0.654	1.020	0.667	0.790	(21%)
Port Eliz. CT	0.957	1.000	0.937	0.937	0.896	(10.4%)
Durban (Pier 1 and 2)	1.140	0.980	0.905	0.887	1.012	1.2%

Luanda (CT2)	1.156	1.000	1.000	1.000	1.156	15.6%
Cape Town CT	1.094	0.937	0.946	0.886	0.970	(3%)
Geometric Average	**1.099**	**0.979**	**0.975**	**0.955**	**1.049**	**4.9%**

The results indicated that the port of Tanger Med increased total factor productivity by 20.3% from 2013 to 2014. This result was not surprising given the technological improvements made in 2014 (Navisworld 2015; 2015). The improvements included upgrading the terminal planning system as well as the radio communication system within the terminal. These upgrades were targeted at improving performance and service levels to customers, in order to ultimately improve productivity. Hence the results lend support to the justification for the capital expenditure invested in Tanger Med.

Apapa container terminal was the most improved port during the period 2013 to 2014, with a TFPC increase equivalent to 49.2%. The increase in efficiency over the sample period was expected, given the large investment in rail and road infrastructure, as can be seen by the 16.7% increase in TC. This led to a greater throughput of containers within the Apapa port and thus a greater PTEC, as indicated by the 28.1% increase. The large increase in PTEC filtered through to the TTEC due to a small decrease in the SEC. Thus, strong investment in infrastructure indirectly lead to a greater throughput of containers and a noticeable increase in the total factor productivity of the port.

East Africa made noticeable gains in total factor productivity over the sample period. Mombasa container terminal in East Africa showed a 13.2% increase in total factor productivity. This gain in efficiency was largely attributed to a 16.6% increase in TC. According to managing director Gichiri Ndua, the port invested heavily in technology to increase the capacity at the port, improve marketing and improve coordination between the port authority and the general port community (Huston, 2014b). This resulted in a more efficient port, leading to greater quantities of containers being handled, as observed with the increase in TC and TFPC. Dar es Salaam container terminal had an increase in its total factor productivity, from 2013 to 2014, of 17.3%. This, according to new CEO of the terminal, Paul Wallace, was due to an upgrade in infrastructure to deliver significantly higher levels of operational productivity and service level reliability, an objective supported by the 13.3% increase in TC (The Report Company, 2014). This improvement in efficiency has led to greater quantities of containers being processed.

The largest decrease in total factor productivity over the sample period, 2013 to 2014, was experienced at Ngqura port with a TFPC of 0.79, equivalent to a 21% decrease in total factor productivity. This decrease in total factor productivity over the sample period was due to a decrease in PTEC of 34.6%. This decrease in PTEC resulted in an almost equivalent decrease in TTEC, despite Ngqura being scale efficient in 2014 with a 2% increase in its SEC. The large reduction in PTEC was believed to be a result of declines in container throughput due to labour strikes experienced in South Africa in 2014, and further compounded by a lack of technological knowhow within the port (de Bruyn, 2014). Ngqura invested in technology over the sample period, resulting in an 18.5% increase in the TC of the port, which no doubt contributed to the SE of the port in 2014.

The container terminal based in Port Elizabeth was one of the more efficient ports in Africa, and the most efficient in South Africa. PECT experienced full PTE in both 2013 and 2014. The technical ability of the terminal has not changed, with a PTEC score of 1.000. However, some scale issues were apparent. PECT is heavily developed in terms of infrastructure. It is located in the popular waterfront location of Humewood in Port Elizabeth and is surrounded by housing and industry, limiting the port's ability to expand both internally and externally in order to cope with a potential increase in traffic. This restriction was supported by decreases in SEC (6.3%) and TC (4.3%), from 2013 to 2014, ultimately leading to a decrease in total factor productivity of 10.4%. This is an issue PECT must address in order to avoid losing competitiveness to other African container terminals.

The Luanda container terminal experienced a 15.6% increase in total factor productivity from 2013 to 2014. This was due to infrastructural developments over the past 5 years, shown by the 15.6% increase in TC. This translated into increased container throughput (Portalangop (ANGOP), 2014).

The full sample suggested that the adoption of better technologies by operators led to dramatic improvement, in terms of the average efficiency growth of the 15 African container terminals. This was evident in the TC average which increased by 9.9%. This increase in technology resulted in an increase in the average amount of total factor productivity among the 15 container terminals, as observed by the 4.9% increase in average TFPC.

6.5 Conclusion

This study applied DEA methods to rank container terminals on the African continent for both the years 2013 and 2014. Bias-corrected bootstrapping methods were used to rank efficiency measures and provide additional insight. PTE and SE were analysed to provide corrective measures to assist selective terminals achieve MPSS. The MPI enabled the study to determine changes in the efficiencies of the terminals over the sample period. All the DEA results analysed provided numerical support of what had been reported in annual reports.

The rankings revealed that the port of Tema had the highest ranking, coupled with a relatively high PTE and SE. The North African container terminals close proximity to the Asia-Euro trade route resulted in these terminals clustering at the top of the relative efficiency rankings for 2013 and 2014. In 2013, Apapa container terminal had the lowest ranking and PTE, coupled with an almost perfect SE. Apapa needed to invest in rail and road infrastructure in 2013 to improve container turnover and thus container throughput. These infrastructure improvements were justified by the 2014 results which showed noticeable improvements.

Focusing the analysis on South African container terminals, Ngqura port had the lowest ranking and PTE, coupled with full SE in 2014. The port of Ngqura needed to triple its container throughput in 2014 to achieve full PTE. The PECT port had the highest ranking and PTE in relation to the other South Africa ports, coupled with the lowest SE in relation to the other African ports, for both 2013 and 2014. MPSS could be achieved in PECT through further operational expansion. The relatively low ranking of Durban, in both 2013 and 2014, was surprising given that an independent report listed Durban as the most efficient container terminal in 2014. This low ranking may be due to recent capital investment planning for long-term growth by Transnet. This meant Durban's throughput at the time of this study had not reached its full potential, given the increased infrastructure.

The sampled African ports on average experienced a 4.9% increase in total factor productivity over the sample period, as estimated by the MPI. This increase in efficiency may be a result of investment in technology within the ports of Tanger Med, Apapa, Dar es Salaam, Mombasa and Luanda. The investments within the ports were justified by the greater total factor productivity score within the ports, with an average increase of 9.9% in TC. This average increase in TC was largely responsible for the increase in the average total factor productivity of the sampled African ports.

Regarding the analysis of South African ports, there were some points of concern, particularly in the Eastern Cape, where NCT and PECT experienced decreases in total factor productivity of 21% and 10.4% over the sample period, respectively. These were the largest decreases in efficiency experienced among the sampled African ports. The reasons for these decreases were arguably related to the labour strikes at Ngqura combined with a lack of technical expertise. At PECT, the decrease in efficiency was most likely due to a lack of infrastructure and space to manage expanding container traffic.

Chapter Seven: Conclusions and Recommendations

7.1 Conclusion

The three main objectives of this study were stated in the introductory chapter. The first of these was to provide a thorough review of global port efficiency research. The second objective was to determine the level of efficiency within each of the African container terminals, using the statistical technique, DEA. The third and final objective was to comment on the efficiency results as well as the trend of the results over the sample period. Within these comments, this study also aimed to provide corrective measures for any inefficiency experienced over the sample period.

All the defined objectives were met. In terms of the first objective, the study reviewed 20 years' worth of DEA applications in the seaport industry. During this review, DEA techniques were found to be suitable for this case study research. These techniques included; the output-orientated CCR and BCC models, Simar and Wilson's (2000) method of homogenous bootstrapping, and the MPI. These techniques were used to calculate the efficiencies of the 15 selected African container terminals. In addition to identifying DEA techniques, three frequently occurring results within the seaport literature were also identified. The first of these was that the smaller the sample size, the less discriminatory power existed within the DEA model. The second outcome was that the larger the port, the higher its efficiency and the final result was that privatisation led to higher efficiency.

To ensure discriminatory power, identified to be important from the literature review, carefully selected variables were used in the DEA models. The selected variables were based on; the high frequency with which they occurred in the literature, their importance in industry, as well as their positive correlation and significance in the statistical tests. In addition, four popular minimum sample size rules were satisfied in order to ensure discriminatory power.

The second objective was met by determining the TE, the PTE and the SE of the selected African container terminals using the CCR and BCC models. These three efficiencies provided a comprehensive measurement of the efficiency present within each of the 15 African container terminals.

The third objective was met by focusing on the PTE and SE components of TE. The bias-corrected pure technical efficiencies were used to rank the container terminals in descending order for each of the sample periods. Once ranking was established, high and low ranking terminals, as well as PECT and NCT, were discussed. Reference to their PTE and SE were made as suggestions for improving areas of inefficiency to achieve MPSS. Changes in efficiency over the sample period were recorded and analysed using the MPI.

Tema had the highest ranking, operating at MPSS in 2013 and very close to MPSS in 2014. A large portion of Tema's port operations are privately run, corroborating the literature review finding that privately run ports were more efficient.

The North African container terminals close proximity to the Asia-Euro trade route resulted in these terminals clustering at the top of the relative efficiency rankings for 2013 and 2014. Half these terminals operated at the MPSS. The North African container terminals are some of the largest in Africa, supporting the literature reviews' finding that larger ports are more efficient.

PECT appeared to be a slight contradiction to the above finding that larger is better in terms of efficiency. PECT had the highest ranking and PTE in relation to the other South Africa ports. However in contrast, PECT had the lowest SE in relation to the other African ports, for both 2013 and 2014. To achieve MPSS, PECT would have to expand its operations given its low SE. Thus, ultimately larger was better in terms of efficiency.

In 2013, Apapa container terminal had the lowest ranking and PTE coupled with an almost perfect SE. Apapa needed to invest in rail and road infrastructure leading into the terminal in 2013 to improve container turnover and therefore container throughput. Ngqura port had the lowest ranking and PTE coupled with full SE in 2014. Ngqura port needed to triple its container throughput in 2014 to achieve full PTE.

The MPI indicated that investments within the ports lead to an average increase of roughly 5% in total factor productivity within the ports and ultimately an average increase of about 10% in TC.

By achieving the three objectives, this study addressed the problem of a lack of information with respect to efficiencies of different ports, identified through discussions with shipping-line companies. This study provided the shipping-line companies with efficiency information that could be used to benchmark the container ports against each other, thereby enabling the companies to make more use of the more efficient container terminals. The study also made port authorities aware of inefficiencies in their processes. These inefficiencies could then be addressed to attract more container traffic from shipping-line companies. In achieving these objectives this study maintained its focus on NCT and PECT in the Eastern Cape.

7.2 Recommendations for Future Research

Two limitations were identified, the first of which was the difficulty in sourcing data. The second shortcoming was the lack of comparison of the DEA efficiency results with alternative models' efficiency results (e.g. FDH, SFA etc.).

If a researcher were able to acquire data for more than two years, unlike this study, compiling a greater number of DMUs and variables, the results should be more robust. The sample would be much larger and thus provide more discriminatory power when calculating the efficiency results. Thus, the use of a larger data set when calculating African container terminal efficiencies would be a possible future research option.

Calculating the African container terminal efficiencies using DEA and an alternative model, for comparison purposes, could provide a more valuable overall analysis of efficiency. As an example, using both DEA and SFA to calculate efficiency would provide both a non-parametric and parametric approach to calculating efficiency, respectively. Thus, allowing for random error to be accounted for in the efficiency result. Therefore, the use of a greater scope of efficiency models when calculating efficiency would be a possible future research option.

However, despite these future recommendations it is believed that a useful contribution to the DEA literature has been made, particularly given the limited amount of DEA applications in African seaports.

References

[1] AIGNER, D., LOVELL, C. & SCHMIDT, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, **6(1)**, pp. 21–37.

[2] AL-ERAQI, A.S., BARROS, C.P., MUSTAFFA, A. & KHADER, A.T., 2007. Evaluating the location efficiency of Arabian and African seaports using data envelopment analysis (DEA).*Working Paper, 019/2007/DE/UECE*

[3] ANDERSEN, P. & PETERSEN, N.C., 1993. A procedure for ranking units in data envelopment analysis. *Management Science,* **39(10)**, pp.1261–1264.

[4] BANKER, R.D., CHARNES, A. & COOPER, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science,* **30(9)**, pp.1078–1092.

[5] BARRADAS, S., 2013. *Engineeringnews: Durban port upgrade and expansion project.* [online] Available at: <http://www.engineeringnews.co.za/article/durban-port-upgrade-and-expansion-project-south-africa-2015-05-22/article_comments:1> [Accessed 14 May, 2015].

[6] BARROS, C.P., 2003. The measurement of efficiency of Portuguese sea port authorities with dea. *International Journal of Transport Economics,* **30(3)**, pp.335–354.

[7] BARROS, C.P., 2006. A benchmark analysis of Italian seaports using data envelopment analysis. *Maritime Economics & Logistics,* **8(4)**, pp.347–365.

[8] BARROS, C.P. & ATHANASSIOU, M., 2004. Efficiency in European seaports with DEA: evidence from Greece and Portugal. *Maritime Economics & Logistics,* **6(2)**, pp.122–140.

[9] BOGETOFT, P. & OTTO, L., 2011. *Benchmarking with DEA, SFA and RSFA*, Springer 2011, R package version 0.26.

[10] BOUSSOFIANE, A., DYSON, R.G. & THANASSOULIS, E., 1991. Applied data envelopment analysis. *European Journal of Operational Research*, **52(1)**, pp.1–15.

[11] BOWLIN, W.F., 1998. Measuring performance: an introduction to data envelopment analysis (DEA). *Journal of Cost Analysis*, **15(2)**, pp.3-27.

[12] BRETTENNY, W.J. & SHARP, G.D., 2016. Efficiency Evaluation of Urban and Rural Municipal Water Service Authorities in South Africa: A Data Envelopment Analysis Approach. *Water SA*, **42(1)**, pp.11-19.

[13] CAMANHO, A.S. & DYSON, R.G., 1999. Efficiency, size, benchmarks and targets for bank branches: an application of data envelopment analysis. *Journal of the Operational Research Society*, **50(9)**, pp.903–915.

[14] CHARNES, A., COOPER, W.W. & RHODES, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, **2(6)**, pp.429–444.

[15] *Containerisation International: March 2015 Online Issue.* 2015. [online] Available at: <www.containershipping.com> [Accessed 14 May, 2015].

[16] COOPER, W.W., SEIFORD, L.M. & TONE, K., 2007. *Data envelopment analysis: A comprehensive text with models, applications, references and dea-solver software*. 2nd ed. New York City: Springer. , pp.1-176; 283-307; 328-333.

[17] CULLINANE, K.P.B., SONG, D.W. & WANG, T.F., 2005. The application of mathematical programming approaches to estimating container port production efficiency. *Journal of Productivity Analysis*, **24(1)**, pp.73-92.

[18] CULLINANE, K.P.B., WANG, T.F., SONG, D.W. & JI, P., 2006. The technical efficiency of container ports: comparing data envelopment analysis and stochastic frontier analysis. *Transportation Research A*, **40(4)**, pp.354–374.

[19] CULLINANE, K.P.B. & WANG, T.F., 2010, The efficiency analysis of container port production using DEA panel data approaches, *Operations Research Spectrum*, **32(3)**, pp.717–738.

[20] DE BORGER, B., KERSTENS, K. & COSTA, A., 2002. Public transit performance: what does one learn from frontier studies? *Transport Reviews*, **22(1)**, pp.1–38.

[21] DE BRUYN, C., 2014. *Engineering news: Numsa strike at Ngqura declared unprotected.* [online] Available at: <http://www.engineeringnews.co.za/article/numsa-strike-at-ngqura-declared-unprotected-2014-07-29> [Accessed 15 May, 2015].

[22] DE OLIVEIRA, G.F. & CARIOU, P., 2011. A dea study of the efficiency of 122 iron ore and coal ports and of 15/17 countries in 2005. *Maritime Policy & Management*, **38(7)**, pp.727–743.

[23] DOYLE, J. & GREEN, R., 1994. Efficiency and cross-efficiency in DEA: derivations, meanings and uses. *Journal of the Operational Research Society*, **45(5)**, pp.567–578.

[24] DYSON, R.G., ALLEN, R., CAMANHO, A.S., PODINOVSKI, V.V., SARRICO, C.S. & SHALE, E.A., 2001. Pitfalls and Protocols in DEA. *European Journal of Operational Research*, **132(2)**, pp.245-259.

[25] EL-MAHGARY, S. & LAHDELMA, R., 1995. Data envelopment analysis: visualizing the results. *European Journal of Operational Research*, **83(3)**, pp.700–710.

[26] ESTACHE, A., DE LA FÉ, B.T. & TRUJILLO, L., 2004. Sources of efficiency gains in port reform: a DEA decomposition of a malmquist tfp index for Mexico. *Utilities Policy*, **12(4)**, pp.221–230.

[27] Färe, R., Grosskopf, S., Norris, M. & Zhang, Z., 1994. Productivity growth, technical progress and efficiency change in industrialized countries. *American Economic Review*, **84(1)**, pp.66–83.

[28] Farrell, M.J., 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society A*, **120(3)**, pp.253–281.

[29] Fourie, Y., 2011. Part 3: Transhipments. *NDoT SA Maritime Transport Sector Study*. [online] Available at: <http://www.transport.gov.za/Portals/0/Maritime/Part%203_Final_280711yf.pdf> [Accessed 12 May, 2015].

[30] *Ghana Ports and Harbour Authority*. 2013. [online] Available at: <http://ghanaports.gov.gh/tm/default> [Accessed 12 May, 2015].

[31] Golany, B. & Roll, Y., 1993. Some Extensions of Techniques to Handle Non-Discretionary Factors in Data Envelopment Analysis. *The Journal of Productivity Analysis*, **4(4)**, pp.419-432.

[32] Goss, R.O., 1990. Economic policies and seaports – part 4: strategies for port authorities. *Maritime Policy & Management*, **17(3)**, pp.273–287.

[33] Herrera, S. & Pang, G., 2008. Efficiency of infrastructure : the case of container ports. *Revista Economia*, **9(1)**, pp.165–194.

[34] Hung, S.W., Lu, W.M. & Wang, T.P., 2010. Benchmarking the operating efficiency of Asia container ports. *European Journal of Operational Research*, **203(3)**, pp.706–713.

[35] Hutson, T., 2014a. *Ports & Ships Maritime News: DCT ranked no.1 amongst Africa ports for berth productivity, says Maersk.* [online] Available at: < http://ports.co.za/news/news_2014_11_25_01.php#one> [Accessed 19 May, 2015].

[36] HUTSON, T., 2014b. *Ports & Ships Maritime News: Strong growth in throughput forecast for port of Mombasa*. [online] Available at: < http://ports.co.za/news/news_2014_11_25_01.php#one> [Accessed 14 May, 2015].

[37] MARTINEZ-BUDRIA, E., DIAZ-ARMAS, R., NAVARRO-IBANEZ, M . & RAVELO-MESA,T., 1999. A study of the efficiency of Spanish port authorities using data envelopment analysis. *International Journal of Transport Economics*, **26(2)**, pp.237–253.

[38] MARTINEZ-NUNEZ, M. & PEREZ-AGUIAR, W.S., 2014. Efficiency analysis of information technology and online social networks management: an integrated DEA-model assessment. *Information & Management*, **51(6)**, pp.712–725.

[39] MEEUSEN, W. & VAN DEN BROECK, J., 1977. Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review*, **18(2)**, pp.435–444.

[40] MKHIZE, N., 2014. *Business day live: Transnet set to start Durban port upgrade.* [online] Available at: <http://www.bdlive.co.za/business/transport/2014/02/06/transnet-set-to-start-durban-port-upgrade> [Accessed 14 May, 2015].

[41] MUNISAMY, S. & DANXIA, W., 2013. Ranking efficiency of Asian container ports : a bootstrapped frontier approach. *International Journal of Shipping and Transport Logistics*, **5(6)**, pp.668-690.

[42] NATARAJA, N.R. & JOHNSON, A.L., 2011. Guidelines for using variable selection techniques in data envelopment analysis. *European Journal of Operational Research*, **215(3)**, pp.662-669.

[43] *Navisworld 2015*. 2015. [online] Available from: <http://navisworld.navis.com/nominees> [Accessed 14 May, 2015].

[44] NOTTEBOOM, T., 2010. From multi-porting to a hub port configuration: the South-African container port system in transition. *International Journal of Shipping and Transport Logistics*, **2(2)**, pp.224-245.

[45] NOTTEBOOM, T., 2011. An application of multi-criteria analysis (MCA) to the location of a container hub port in South Africa. *Maritime Policy and Management*, **38(1)**, pp.51-79.

[46] Ocean Shipping Consultants. 2008. *Beyond the Bottlenecks: Ports in Sub-Saharan Africa*. In: African Infrastructure Country Diagnostic. 2009. Background Paper 8 (Phase 2). Washington, DC (USA): A publication of the World Bank.

[47] PACHKOVA, E.V., 2005. *FDH model* [online] Available at: <http://www.math.ku.dk/~elena/index_files/FDH%20model.pdf> [Accessed 24 August, 2015].

[48] PANAYIDES, P.M., MAXOULIS, C.N., WANG, T.F. & NG, K.Y.A., 2009. A critical analysis of dea applications to seaport economic efficiency measurement. *Transport Reviews*, **29(2)**, pp.183–206.

[49] *Portalangop (ANGOP)*. 2014. *Luanda port distinguished best regional container terminal*. [online] Available at: <http://www.portalangop.co.ao/angola/en_us/noticias/economia/2014/2/12/Luanda-Port-distinguished-best-regional-container-terminal,aaea1048-2ebe-4d5d-a6b3-88b79e9e2636.html> [Accessed 14 May, 2015].

[50] PASTOR, J.T., RUIZ, J.L. & SIRVENT, I., 2002. A statistical test for nested radial dea models. *Operations Research*, **50(4)**, pp.728–735.

[51] POITRAS, G., TONGZON, J. & LI, H., 1996. Measuring port efficiency: an application of data envelopment analysis. *Working Paper*, Singapore: National University of Singapore.

[52] *Ports & Ships: Ngqura container terminal.* 2012a. [online] Available at: <http://ports.co.za/coega.php> [Accessed 12 May, 2015].

[53] *Ports & Ships: Port Elizabeth container terminal*. 2012b. [online] Available at: <http://ports.co.za/port-elizabeth.php> [Accessed 12 May, 2015].

[54] R CORE TEAM, 2015. *R: A language and environment for statistical computing (Version 3.2.3)*, R Foundation for Statistical Computing, Vienna, Austria, Available at: <https://www.R-project.org/> [Accessed 15 January, 2014].

[55] RIOS, L.R. & MAÇADA, A.C.C., 2006. Analysing the relative efficiency of container terminals of Mercosur using dea. *Maritime Economics & Logistics*, **8(4)**, pp.331–346.

[56] ROLL, Y. & HAYUTH, Y., 1993. Port performance comparison applying data envelopment analysis. *Maritime Policy & Management*, **20(2)**, pp.153–161.

[57] ROSENDAHL, A., 2014. *Maersk: Apapa looks to the hinterland*. [online] Available at: <http://www.maersk.com/en/the-maersk-group/about-us/maersk-post/2014-5/apapa-looks-to-the-hinterland> [Accessed 12 May 2015].

[58] SAKRIS, J., 2002. Preparing your data for dea. In: Avkiran, N.K, 2006. *Productivity Analysis in the Service Sector with Data Envelopment Analysis*. 3rd ed. Queensland (Aus): University of Queensland Business School.

[59] SIMAR, L. & WILSON, P.W., 2000. A general methodology for bootstrapping in nonparametric frontier models. *Journal of Applied Statistics*, **27(6)**, pp.779–802.

[60] SIMAR, L. & WILSON, P.W., 2007. *Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives*. 2nd ed. Oxford (UK): Oxford University Press.

[61] SILVERMAN, B.W., 1986. *Density estimation for statistics and data analysis*. London (UK): Chapman and Hall.

[62] TALLEY, W.K., 2000. Ocean container shipping: impacts of a technological improvement. *Journal of Economic Issues*, **34(4)**, pp.933–948.

[63] THANASSOULIS, E., 2001. *Introduction to the Theory and Application of Data Envelopment Analysis: a Foundation Text with Integrated Software*. USA: Kluwer Academic Publishers.

[64] *The report company: Q&A Paul Wallace - CEO of Tanzania international container terminal services (TICTS)*. 2014. [online] Available at: <http://www.thereport.net/tanzania/nov2014/1192-interview-paul-wallace-ceo-of-tanzania-international-container-terminal-services-ticts> [Accessed 14 May, 2015].

[65] TONGZON, J., 2001. Efficiency measurement of selected Australian and other international ports using data envelopment analysis. *Transportation Research A*, **35(2)**, pp.113–128.

[66] *Transnet Port Terminals*. 2013. [online] Available at: <http://www.transnetportterminals.net/pages/default.aspx> [Accessed 14 May, 2015].

[67] VALENTINE, V.F. & GRAY, R., 2001. *The measurement of port efficiency using data envelopment analysis*. In: Proceedings of the 9th World Conference on Transport Research. Seoul, Republic of Korea, July 22-27, 2001.

[68] VALENTINE, V.F. & GRAY, R., 2002. *An organisational approach to port efficiency*. In: Proceedings of the 10th International Association of Maritime Economists Conference. Panama, 13-15 November, 2002.

[69] VAN TONDER, A., 2014. Operations manager at Eastern Cape (Maersk)., (Personal Comunication, 2014). Contactable at Andre.Van.Tonder@maersk.com

[70] WANG, T.F. & CULLINANE, K.P.B., 2006. The efficiency of European container terminals and implications for supply chain management. *Maritime Economics & Logistics*, **8(1)**, pp.82–99.

[71] WANG, T.F., CULLINANE, K.P.B. & SONG, D.W., 2005. *Container port production and economic efficiency*. Basingstoke: Palgrave-Macmillan.

[72] WANG, T.F., SONG, D.W. & CULLINANE, K.P.B., 2002. *The applicability of data envelopment analysis to efficiency measurement of container ports*. In: Proceedings of the 10th International Association of Maritime Economists Conference. Panama: 13–15 November 2002.

[73] WILSON, P.W., 2008. FEAR 1.0: A Software Package for Frontier Efficiency Analysis with R. *Socio-Economic Planning Sciences*, **42**, pp.247-254.

[74] WU, Y.C.J. & GOH, M., 2010. Container port efficiency in emerging and more advanced markets. *Transportation Research Part E: Logistics and Transportation Review*, **46(6)**, pp. 1030–1042.

Data References

Egypt

[i] *Alexandria Port Authority*. 2015. [online], Available at: <http://www.apa.gov.eg/index.php?option=com_content&view=article&id=67&Itemid=84&lang=en> [Accessed 26 March, 2015].

[ii] *Alexandria Port Authprity Department of Information Centre, Statement of Port Operations During July 2013-July 2014*. 2015. [online] Available at: <http://www.apa.gov.eg> [Accessed 25 March, 2015].

[iii]BARNARD, B., 2015. *joc.com, TEU Volume Jumps 18.8 Percent at Cosco Piraeus Terminal*. [online] Available at:<http://www.joc.com> [Accessed March 25, 2015].

[iv]*Containerisation International: March 2015 Online Issue*. 2015. [online] Available at: <www.containershipping.com> [Accessed 26 March, 2015].

[v] *Damietta Port Authority, Damietta Port 2014 in Figures;Reports and Statistics;Terminals*. 2015. [online] Available at: <http://www.dpa.gov.eg> [Accessed 25 March, 2015].

[vi]SHIEH, V., 2014. *Port Report Africa (Port Said)*. [online], Available at: <http://www.portoverview.com/> [Accessed 26 March, 2015].

Morocco

[i] SHIEH, V., 2014. *Port Report Africa (Casablanca)*. [online] Available at: [Accessed 25 March 2015].

[ii] SHIEH, V., 2014. *Port Report Africa (Tanger Med)*. [online] Available at: .[Accessed 26 March, 2015].

[iii]*The National Ports Agency, Traffic 2014*. 2012. [online] Available at: <http://www.anp.org.ma>[Accessed 26 March, 2015].

South Africa

[i] HUTSON, T., 2015. *SA PORT STATISTICS FOR THE CALENDAR YEAR 2014 ARE AVAILABLE HERE, Durban: Ports & Ships Maritime News*. [online] Available at: <http://ports.co.za/news/news> [Accessed March 25, 2015].

[ii] *Inchcape Shipping Services, Inchcape Shipping Services: Port of Durban*. 2007. [online] Available at: <http://iss-shipping.com/Microsites/Document%20Library/Port%20Of%20Durban.pdf> [Accessed 26 March, 2015].

[iii]*Transnet National Port Authority*. 2010. [online] Available at: <http://www.transnetnationalportsauthority.net/OurPorts/Pages/default.aspx> [Accessed 25 March, 2015].

Luanda

[i] *Portalangop (ANGOP): Luanda port distinguished best regional container terminal*.2014.[online] Available at: <http://www.portalangop.co.ao/angola/en_us/noticias/economia/2014/2/12/Luanda-Port-distinguished-best-regional-container-terminal,aaea1048-2ebe-4d5d-a6b3-88b79e9e2636.html> [Accessed 14 May, 2015].

[ii] *macauhub*. 2015. [online] Available at: <http://www.macauhub.com.mo/en/2015/03/18/port-of-luanda-angola-handles-over-13-million-tons-of-cargo-in-2014/> [Accessed 26, March 2015].

[iii]SHIEH, V., 2014. *Port Report Africa (Luanda)*. [online] Available at: <http://www.portoverview.com> [Accessed 26 March, 2015].

References

Mombasa

[i] Burite, J., 2015. *East African Business Week*. [online] Available at: <http://www.busiweek.com/index1.php?Ctp=2&pI=2873&pLv=3&srI=59&spI=47> [Accessed 26 March, 2015].

[ii] Shieh, V., 2014. *Port Report Africa (Mombasa)*. [online] Available at: <http://www.portoverview.com> [Accessed 26 March, 2015].

Tema

[i] *Container Management*. 2014. [online] Available at: <http://container-mag.com/2014/11/18/tema-port-undertake-us1-5bn-expansion/> [Accessed 26 March, 2015].

[ii] Shieh, V., 2014. *Port Report Africa (Tema)*. [online] Available at: <http://www.portoverview.com/> [Accessed 25 March, 2015].

Dar es Salaam

[i] Elinaza, A., 2015. *Daily News*. [online] Available at: <http://www.dailynews.co.tz/index.php/biz/40102-dar-port-to-unlock-neighbours> [Accessed 26 March, 2015].

[ii] *Port Strategy*. 2014. [online] Available at: <http://www.portstrategy.com/news101/insight-and-opinion/post-script/dar-saga-plays-on> [Accessed 26 March, 2015].

[iii] Shieh, V., 2014. *Port Report Africa (Dar es Salaam)*. [online] Available at: <http://www.portoverview.com/> [Accessed 25 March, 2015].

References

Djibouti

[i] *Port of Djibouti*. 2014. [online] Available at: <http://www.portdedjibouti.com/PDF/2014%20CONTAINERIZED%20CARGO%20TRAFFIC%20DCT.pdf> [Accessed 25 March, 2015].

[ii] Shieh, V., 2014. *Port Report Africa (Djibouti)*. [online] Available at: <http://www.portoverview.com/> [Accessed 25 March, 2015].

Lagos

[i] Hutson, T., 2015. *Ports and Ships*. [online] Available at: <http://ports.co.za/news/news_2015_01_20_01.php> [Accessed 26 March, 2015].

[ii] Rosendahl, A., 2014. *Maersk: Apapa looks to the hinterland*. [online] Available at: <http://www.maersk.com/en/the-maersk-group/about-us/maersk-post/2014-5/apapa-looks-to-the-hinterland> [Accessed 12 May, 2015].

[iii]Shieh, V., 2014. *Port Report Africa (Apapa, Lagos)*. [online] Available at: <http://www.portoverview.com/.> [Accessed 25 March, 2015].

Appendices

Appendix One

Table A: International Literature					
Literature Synthesis Sorted by Date of Publication					
				Model Variables	
Author	**Domain**	**Data**	**DMU's**	**Outputs**	**Inputs**
Roll & Hayuth, (1993)	Entire world	Fictitious and cross-sectional, single period	20 ports	1.Container throughput 2.Service level 3.User satisfaction 4.Ship calls	1.Size of labour force 2.Annual investment per port 3.The uniformity of facilities and cargo
Poitras, Tongzon & Li, (1996)	Australian and international	Cross-sectional	23 ports	1.TEU container berth hour (TEUBH) 2.Total number of containers handled per year (TH)	1.Mix of 20-foot and 40-foot containers (CONMIX) 2.Average delays in commencing stevedoring, difference between the berth time and gross working time (BRLWT) 3.Number of containers lifted per quay crane hour (TEUCH) 4.Number of gantry cranes (CRANE) 5.Frequency of ship calls (FH) 6.Average government port charges per container (CH)

Martinez-Budria, Diax-Armas, Navarro-Ibanez & Ravelo-Mesa, (1999)	Spain	Panel	26 ports in five year span	1.Total cargo moved through the docks (in thousands of tons) 2.Revenue obtained from rent of port facilities (millions of pesetas)	1.Labour expenditures 2.Depreciation charges 3.Miscellaneous expenditures
Tongzon, (2001)	Australia & International	Cross-sectional	16 ports	1.Cargo throughput (containers) 2.Ship working rate (container moves per hour)	1.Capital (number of berths, cranes, tugs) 2.Labour (number of stevedore gangs) 3.Land (size of terminal areas) 4.Length of delay
Valentine & Gray, (2001)	Entire World	Cross-sectional	21 Ports	1.Total tons throughput 2.Number of containers	1.Quay length (in metres) 2.Assets (USD OR $)
Valentine & Gray (2002)	North America and Europe	Cross-sectional	19 Ports	1.Total tons throughput 2.Number of containers	1.Total length of berth (in metres) 2.Container berth length (in metres)
Barros (2003)	Portugal	Panel	11 ports	1.Number of Ships 2.Movement of freight 3.Gross gauge 4.Break-bulk cargo 5.Containerized freight 6.Solid bulk and liquid bulk	1.Labour (number of workers) 2.Capital (book value of the assets)

Barros & Anthanassiou (2004)	Greece and Portugal	Panel	6 ports	1.Movement of freight 2.Total cargo handled 3.Containers loaded and unloaded	1.Number of workers 2.Book value of assets
Estache, De la Fe & Trujillo (2004)	Mexico	Panel	11 ports	1.Cargo (volume in tons of merchandize handled)	1.Labour (number at each port) 2.Capital (length of docks)
Cullinane, Song & Wang (2005)	Worldwide	Cross-sectional	57 ports/container terminals	1.Container throughput (in TEU containers)	1.Terminal length (in metres) 2.Terminal area (in hectors) 3.Number of quayside gantry 4.Number of yard gantry 5.Number of straddle carrier
Barros (2006)	Italy	Panel	24 ports	1.Liquid bulk (oil and other liquid products) 2.Dry bulk (Ro-Ro Cargo and other dry bulk) 3.Number of ships 4.Number of passengers 5.Number of TEU containers 6.Number of non TEU containers 7.Total sales	1.Number of personnel 2.Value of capital invested 3.Size of operating costs
Cullinane, Wang, Song & Ji (2006)	Worldwide	Cross-sectional	57	1.Container throughput (in TEU containers)	1.Terminal length (in metres) 2.Terminal area (in hectares) 3.Number of quayside gantry cranes 4.Number of yard gantry cranes 5.Number of straddle carriers

Rios & Macada (2006)	Brazil, Argentina and Uruguay	Panel	23 terminals	1.TEU containers handled 2.Average number of containers handled per hour per ship	1.Number of cranes 2.Number of berths 3.Number of employees 4.Terminal area (in square metres) 5.Amount of yard equipment
Wang & Cullinane (2006)	Pan European	Cross-sectional	69 terminals	1.Container throughput (in TEU containers)	1.Terminal length (in metres) 2.Terminal area (in hectares) 3.Amount of equipment
Herrera & Pang (2008)	International	Panel	86 ports	1.Throughput (container TEUs)	1.Terminal Area 2.Ship-to-shore gantries 3.Number of quay gantries 4.Number of yard gantries 5.Number of mobile gantries 6.Number of tractors and trailers
de Oliveira & Cariou (2011)	International	Cross-sectional	122 ports	1.Cargo throughput (in tons)	1.Draaght (in metres) (nautical assistance and resources to accommodate the vessel) 2.Berth length (in metres) 3.Stockpile capacity (in tons) 4.(Un)Loading rates (in metric-tons/hour)

Munisamy & Danxia (2013)	Asia	Cross-sectional	69 Ports	1.Total throughput in TEUs	1.Berth length (in metres) 2.Terminal area (in metres squared) 3.Total reefer points (number of points where refrigerated containers can be plugged in to keep them cold) 4.Total quayside cranes 5.Total yard equipment

Table B: Local (African and/or South African) Literature					
Literature Synthesis Sorted by Date of Publication					
				Model Variables	
Author	**Domain**	**Data**	**DMU's**	**Outputs**	**Inputs**
Al-Eraqi, Barros, Mustaffa & Khader (2007)	Middle East and East Africa	Panel	22 ports	1.Number of ship Calls 2.Cargo throughput (in tons)	1.Berth length (in metres) 2.Distance (in nautical miles) 3.Terminal area (in metres squared)
Ocean Shipping Consultants (2008)	Sub-Saharan Africa	Non-DEA application, but is a seaport based application.			
Notteboom (2010)	South Africa	Non-DEA application, but is a seaport based application.			
Notteboom (2011)	South Africa	Non-DEA application, but is a seaport based application.			

Appendix Two

Table C: Forward ECM for 2013 (CCR and BCC) Inputs

Rounds	**Input Variable(s)**		***X*** **(Significant Inputs)**		***Z*** **(Candidate Inputs)**		***Y*** **(Significant Output)**		**Test Statistic for Candidates (*Z*)**		**P-value for Candidates (*Z*)**	
Model	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**
Round 1	Berth Length	Berth Length	Berth Length	Berth Length	Num. Berths	Num. Berths	Num. TEU's	Num. TEU's	9	6	1.2510e-6	1.474e-03
					Num. Terminal Cranes	Num. Terminal Cranes			13	11	1.270e-12	**2.729e-09**
					Num. Yard Equipment	Num. Yard Equipment			14	8	**1.000e-14**	1.721e-05
Round 2	Berth Length	Berth Length	Berth Length	Berth Length	Num. Berths	Num. Berths	Num. TEU's	Num. TEU's	12	6	1.474e-03	0.002
	Num. Yard Equipment	Num. Terminal Cranes	Num. Yard Equipment	Num. Terminal Cranes	Num. Terminal Cranes	Num. Yard Equipment			11	6	**1.251e-06**	**0.002**
Round 3	Berth Length	Berth Length	Berth Length	Berth Length	Num. Berths	Num. Berths	Num. TEU's	Num. TEU's	7	5	**0.0002**	**0.009**
	Num. Yard Equipment	Num. Terminal Cranes	Num. Berths	Num. Terminal Cranes								
	Num. Terminal Cranes	Num. Yard Equipment	Num. Terminal Cranes	Num. Yard Equipment								

Round 4	Berth Length	Berth Length	-	-	-	-	-
	Num. Berths	Num. Berths					
	Num. Terminal Cranes	Num. Terminal Cranes					
	Num. Yard Equipment	Num. Yard Equipment					

Table D: Forward ECM for 2014 (CCR and BCC) Inputs

Rounds	**Input Variable(s)**		***X* (Significant Inputs)**		***Z* (Candidate Inputs)**		***Y* (Significant Output)**		**Test Statistic for Candidates (*Z*)**		**P-value for Candidates (*Z*)**	
Model	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**	**CCR**	**BCC**
Round 1	Berth Length	Berth Length	Berth Length	Berth Length	Num. Berths	Num. Berths	Num. TEU's	Num. TEU's	7	3	1.814e-04	0.158
					Num. Terminal Cranes	Num. Terminal Cranes			12	7	7.498e-11	1.814e-04
					Num. Yard Equipment	Num. Yard Equipment			13	8	**1.270e-12**	**1.721e-05**
Round 2	Berth Length	Berth Length	Berth Length	Berth Length	Num. Berths	Num. Berths	Num. TEU's	Num. TEU's	3	3	0.158	0.158

	Num. Yard Equipment	Num. Yard Equipment	Num. Yard Equipment	Num. Yard Equipment	Num. Terminal Cranes	Num. Terminal Cranes			5	4	**0.009**	**0.044**
Round 3	Berth Length Num. Yard Equipment Num. Terminal Cranes	Berth Length Num. Yard Equipment Num. Terminal Cranes	Berth Length Num. Berths Num. Terminal Cranes	Berth Length Num. Yard Equipment Num. Terminal Cranes	Num. Berths	Num. Berths	Num. TEU’s	Num. TEU’s	2	2	**0.415**	**0.415**
Round 4	Berth Length Num. Berths Num. Terminal Cranes Num. Yard Equipment	Berth Length Num. Berths Num. Terminal Cranes Num. Yard Equipment	-		-		-		-		-	

Appendix Three

Table E: Universities A-T Variable Values used in El-Mahgary & Lahdelma (1995)

Universities	Inputs			Outputs			
	Symbol	Expenditure	Admission Policy	Graduates	Post-graduates	Graduation Speed	Completion
University of Helsinki	A	1204.651	4.542	1707	330	0.143	0.587
University of Jyvaskyla	B	349.531	4.966	776	107	0.167	0.718
University of Oulu	C	504.882	2.983	860	115	0.154	0.662
University of Joensuu	D	179.618	3.445	492	52	0.167	0.717
University of Kuopio	E	196.747	3.66	265	50	0.167	0.593
University of Turku	F	457.718	4.727	881	105	0.154	0.68
University of Tampere	G	338.626	5.28	722	91	0.154	0.537
Abo Academy	H	207.752	1.796	377	51	0.143	0.701
University of Vaasa	I	71.724	3.162	227	11	0.2	0.739
University of Lapland	J	82.839	5.941	225	10	0.2	1.018
College of Veterinary Medicine	K	56.176	7.349	33	2	0.143	0.767
Helsinki University of Technology	L	467.668	2.563	724	156	0.133	0.681
Tampere University of Technology	M	209.132	2.701	364	70	0.167	0.704
Lappeenranta Univer. of Technology	N	105.861	1.718	190	11	0.154	0.629
Helsinki School of Econ. & Bus. Adm.	O	129.407	4.551	293	17	0.167	0.72
Swedish School of Econ. & Bus. Adm.	P	50.129	2.551	140	6	0.182	0.543
Turku School of Econ. & Bus. Adm.	Q	53.018	3.247	211	9	0.182	0.938
University of Industrial Arts	R	90.132	16.429	119	2	0.143	0.696
Sibelius Academy	S	111.031	5.706	88	2	0.133	0.524
Theatre Academy	T	44.482	50.277	19	0	0.182	0.365

Source: Direct extract from (El-Mahgary & Lahdelma, 1995)