MEASURING THE EFFICIENCY OF UNIVERSITY LIBRARIES, GHANA- Using Data Envelopment Analysis Method

by

Sampson Owadie

A thesis

presented to the Kwame Nkrumah University of Science and Technology-Ghana

in partial fulfilment of the

requirement for the award of degree of

Master of Science

in

Industrial Mathematics

Kumasi, Ghana, 2011

DECLARATION

I hereby declare that this thesis is my own work in partial fulfilment of the award of Master of Science Degree in Industrial Mathematics, and that to the best of my knowledge it contains no material previously published by another person nor material which has been accepted for award of any other degree of any university except where due acknowledgement / references has been made in the text.

Sampson Owadie **PG3016709**	………………	………………
(Student's Name & ID)	Signature	Date

Certified By

Mr. Charles Sebil	……....………….	………....……
(**Supervisor**)	Signature	Date

Certified by

Dr. S. K. Amponsah	………………….	…………………
(**Head of Department**)	Signature	Date

Certified by

Prof. I. K. Dontwi	………………….	….………………..
(**Dean of Distance Learning**)	Signature	Date

ABSTRACT

Data envelopment analysis is a non- parametric linear programming based technique used for measuring the relative performance of organisational units where the presence of multiple inputs and output makes comparison difficult. The aim of this project is to apply Data Envelopment Analysis in order to measure the relative efficiency of University Libraries. The data for three universities for years 2009 and 2010 has been estimated. We identified one input and two outputs. The input variable is print edition expenses. As outputs variable we identified: number of customers served and number of books borrowed. We found that one library formed the efficiency frontier in 2009 and one university formed the efficiency frontier in 2010.

Acknowledgements

I would like to thank my head of department, Dr. S.K. Amponsah and my supervisor Mr. Charles Sebil of Kwame Nkrumah University of Science and Technology, Department of Mathematics, Kumasi, who have provided insight and have had a significant input to my thesis. Not forgetting my wife who has been waking me up at midnight.

The greatest thanks go to the life giver, JEHOVAH GOD.

TABLE OF CONTENT

LIST OF TABLES:

LIST OF FIGURES:

MEANING OF ABBREVIATIONS/ACCRONYMS

CCR.. **Charnes, Cooper and Rhodes**

DEA.. **Data Envelopment Analysis**

KNUST ....**Kwame Nkrumah University of Science and Technology**

DMU... **Decision Making Unit**

UCC... **University of Cape Coast**

UG ....**University of Ghana**

CHAPTER ONE

1.0 INTRODUCTION

Higher education has come under increasing pressure to demonstrate its value and performance (Council for Aid to Education, 1997; Michalko, 1993). As a result many universities and colleges have begun to look at a wide array of performance and accountability measures. These institutions are not only interested in their own (absolute) accomplishments but also want to know how well they perform when compared to peer institutions. As a sub-unit or department in universities, academic libraries "feel the pressures" (Kyrillidou, 1998, p.4) from the parent institutions to produce evidence on:

1) how well the resources are utilized in terms of producing meaningful outputs, and

2) how well the library compares or competes with other libraries within some peer group.

Efficiency problems in public sector have been at the center of economic and political debates in Ghana for a long time. It has been argued that many Public/Government institutions are using Government funds and resources ineffectively. However, empirical studies addressing the issue are very rare. Therefore, the motivation of this research is to investigate empirically a hotly debated subject and initiate research about efficiency evaluation of public sector institutions using the example of public libraries. It is hoped that our results help policy makers in their decisions in allocating public resources to different services and administrative units and in making the services better.

Because of their specific organization, University Libraries present certain difficulties in their efficiency evaluation. One recent approach to the evaluation of library efficiency is Data Envelopment Analysis (DEA). There have been a number of studies that applied DEA technique in order to assess the efficiency of different types of libraries. The most recent and accomplished is the paper of Shim, where a comparison of DEA applications in libraries is put forward. Chen, Vitaliano and Shim examine academic libraries and Hammond, Sharma et al., and Worthington study the efficiency of public libraries. Easun is one of the firsts to apply DEA approach to evaluate school libraries. The aim of the present project work is to apply DEA to measure the efficiency of public University Libraries in Ghana restricting ourselves to the university library at Kwame Nkrumah University of Science and Technology, Kumasi, University of Cape Coast, Cape Coast and University of Ghana, Legon-Accra.

Thus DEA has been found to be particularly suitable in solving the following three basic performance questions that any organization is faced with:

- How well are we doing relative to the others doing the same things as we do?
- What do we need to improve?
- Who are the best performers for benchmarking purposes?

Data Envelopment Analysis (DEA) is a relatively new "data oriented" approach for evaluating the performance of a set of peer entities called Decision Making Units (DMUs) which convert multiple inputs into multiple outputs. The definition of a DMU is generic and flexible. Recent years have seen a great variety of applications of DEA for use in evaluating the performances of many different kinds of entities engaged in many different activities in many different contexts in many different countries. These DEA applications have used DMUs of various forms to evaluate the performance of entities, such as hospitals, universities, cities, courts, business firms, and others, including the performance of countries, regions, etc. Because it requires very few assumptions, DEA has also opened up possibilities for use in cases which have been resistant to other approaches because of the complex (often unknown) nature of the relations between the multiple inputs and multiple outputs involved in DMUs.

As pointed out in Cooper, Seiford and Tone (2000), DEA has also been used to supply new insights into activities (and entities) that have previously been evaluated by other methods.

For instance, studies of benchmarking practices with DEA have identified numerous sources of inefficiency in some of the most profitable firms - firms that had served as benchmarks by reference to this (profitability) criterion – and this has provided a vehicle for identifying better benchmarks in many applied studies. Because of these possibilities, DEA studies of the efficiency of different legal organization firms such as "stock" vs. "mutual" insurance companies have shown that previous studies have fallen short in their attempts to evaluate the potentials of these different forms of organizations. Similarly, a use of DEA has suggested reconsideration of previous studies of the efficiency with which pre- and post-merger activities have been conducted in banks that were studied by DEA.

Since DEA in its present form was first introduced in 1978, researchers in a number of fields have quickly recognized that it is an excellent and easily used methodology for modeling operational processes for performance evaluations. This has been accompanied by other developments. For instance, Zhu (2002) provides a number of DEA spreadsheet models that can be used in performance evaluation and benchmarking. DEA's empirical orientation and the absence of a need for the numerous *a priori* assumptions that accompany other approaches (such as standard forms of statistical regression analysis) have resulted in its use in a number of studies involving efficient frontier estimation in the governmental and nonprofit sector, in the regulated sector, and in the private sector. In their originating study, Charnes, Cooper, and Rhodes (1978) described DEA as a 'mathematical programming model applied to observational data [that] provides a new way of obtaining empirical estimates of relations - such as the production functions and/or efficient production possibility surfaces – that are cornerstones of modern economics'.

Because of this perspective, DEA proves particularly adept at uncovering relationships that remain hidden from other methodologies. For instance, consider what one wants to mean by "efficiency", or more generally, what one wants to mean by saying that one DMU is more efficient than another DMU. This is accomplished in a straightforward manner by DEA without requiring explicitly formulated assumptions and variations with various types of models such as in linear and nonlinear regression models.

Relative efficiency in DEA accords with the following definition, which has the advantage of avoiding the need for assigning a priori measures of relative importance to any input or output,

Definition 1 (Efficiency – Extended Pareto-Koopmans Definition): Full (100%) efficiency is attained by any DMU if and only if none of its inputs or outputs can be improved without worsening some of its other inputs or outputs.

In most management or social science applications the theoretically possible levels of efficiency will not be known. The preceding definition is therefore replaced by emphasizing its uses with only the information that is empirically available as in the following definition:

Definition 2 (Relative Efficiency): A DMU is to be rated as fully (100%) efficient on the basis of available evidence if and only if the performances of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs.

Notice that this definition avoids the need for recourse to prices or other assumptions of weights which are supposed to reflect the relative importance of the different inputs or outputs. It also avoids the need for explicitly specifying the formal relations that are supposed to exist between inputs and outputs. This basic kind of efficiency, referred to as "technical efficiency" in economics. It can, however, be extended to other kinds of efficiency when data such as prices, unit costs, etc., are available for use in DEA.

1.1. BACKGROUND AND HISTORY of DEA

In an article which represents the inception of DEA, Farrell (1957) was motivated by the need for developing better methods and models for evaluating productivity. He argued that while attempts to solve the problem usually produced careful measurements, they were also very restrictive because they failed to combine the measurements of multiple inputs into any satisfactory overall measure of efficiency. Responding to these inadequacies of separate indices of labor productivity, capital productivity, etc., Farrell proposed an activity analysis approach that could more adequately deal with the problem. His measures were intended to be applicable to any productive organization; in his words, '… from a workshop to a whole economy'. In the process, he extended the concept of "productivity" to the more general concept of "efficiency".

The focus in this thesis is on basic DEA models for measuring the efficiency of a DMU *relative* to similar DMUs in order to estimate a 'best practice' frontier. The initial DEA model, as originally presented in Charnes, Cooper, and Rhodes (CCR) (1978), built on the earlier work of Farrell (1957).

This work by Charnes, Cooper and Rhodes originated in the early 1970s in response to the thesis efforts of Edwardo Rhodes at Carnegie Mellon University's School of Urban & Public Affairs - now the H.J. Heinz III School of Public Policy and Management. Under the supervision of W.W. Cooper, this thesis was to be directed to evaluating educational programs for disadvantaged students (mainly black or Hispanic) in a series of large scale studies undertaken in U.S. public schools with support from the Federal government. Attention was finally centered on Program Follow Through - a huge attempt by the U.S. Office (now Department) of Education to apply principles from the statistical design of experiments to a set of matched schools in a nation-wide study. Rhodes secured access to the data being processed for that study by Abt Associates, a Boston based consulting film, under contract with the US Office of Education. The data base was sufficiently large so that issues of degrees of freedom, etc., were not a serious problem despite the numerous input and output variables used in the study. Nevertheless, unsatisfactory and even absurd results were secured from all of the statistical-econometric approaches that Rhodes attempted to use.

While trying to respond to this situation, Rhodes called Cooper's attention to M.J. Farrell's seminal article "The Measurement of Productive Efficiency," in the 1957 Journal of the Royal_Statistical Society. In this article Farrell used "activity analysis concepts" to correct what he believed were deficiencies in commonly used index number approaches to productivity (and like) measurements.

Cooper had previously worked with A. Charnes in order to give computationally implementable form to Tjalling Koopmans' "activity analysis concepts." So, taking Farrell's statements at face value, Cooper and Rhodes formalized what was involved in the definitions that were given in section 1 of this chapter.

The name of Pareto is assigned to the first of these two definitions for the following reasons. In his Manual of Political Economy (1906) the Swiss-Italian economist, Vilfredo Pareto, established the basis of modern "welfare economics", i.e., the part of economics concerned with evaluating public policies, by noting that a social policy could be justified if it made some persons better off without making others worse off. In this way the need for making comparisons between the value of the gains to some and the losses to others could be avoided. This property, known as the "Pareto criterion" as used in welfare economics, was carried over, or adapted, in Activity Analysis of Production and Allocation,_a book edited by Koopmans (1951). In this context, it was "final goods" which were accorded this property, in that they were all constrained so that no final good was allowed to be improved if this improvement resulted in worsening one or more other final goods. These final goods

(=outputs) were to be satisfied in stipulated amounts while inputs were to be optimally determined in response to the prices and amounts exogenously fixed for each output (=final good). Special attention was then directed by Koopmans to "efficiency prices" which are the prices associated with efficient allocation of resources (=inputs) to satisfy the pre-assigned demands for final goods.

Pareto and Koopmans were concerned with analyses of entire economies. In such a context it is reasonable to allow input prices and quantities to be determined by reference to their ability to satisfy final demands. Farrell, however, extended the Pareto-Koopmans property to inputs as well as outputs and explicitly eschewed any use of prices and/or related "exchange mechanisms." Even more importantly, he used the performance of other DMUs to evaluate the behavior of each DMU relative to the outputs and the inputs they all used. This made it possible to proceed empirically to determine their relative efficiencies.

The resulting measure which is referred to as the "Farrell measure of efficiency," was regarded by Farrell as restricted to meaning "technical efficiency" or the amount of "waste" that can be eliminated without worsening any input or output. This was then distinguished by Farrell from "allocative" and "scale" efficiencies as adapted from the literature of economics. Here we want to note that Farrell's approach to efficiency evaluations, as embodied in the "Farrell measure," carries with it an assumption of equal access to inputs by all DMUs.

This does not mean that all DMUs use the same input amounts, however, and, indeed, part of their efficiency evaluations will depend on the input amounts used by each DMU as well as the outputs which they produce.

This "equal access assumption" is a mild one; at least as far as data availability is concerned. It is less demanding than the data and other requirements needed to deal with aspects of performance such as "allocative" or "scope" and "scale efficiencies." Furthermore, as discussed below, this assumption can now be relaxed. For instance, one can introduce "non-discretionary variables and constraints" to deal with conditions beyond the control of a DMU's management--in the form of "exogenously" fixed resources which may differ for each DMU. One can also introduce "categorical variables" to insure that evaluations are effected by reference to DMUs which have similar characteristics.

To be sure, the definition of efficiency that we have referred to as "Extended Pareto-Koopmans Efficiency" and "Relative Efficiency" were formalized by Charnes, Cooper and Rhodes rather than Farrell. However, these definitions conform both to Farrell's models and the way Farrell used them. In any case, these were the definitions that Charnes, Cooper and Rhodes used to guide the developments that we next describe.

The Program Follow Through data with which Rhodes was concerned in his thesis recorded "outputs" like "increased self esteem in a disadvantaged child" and "inputs" like "time spent by a mother in reading with her child," as measured by psychological tests and prescribed record keeping and reporting practices. Farrell's elimination of the need for information on prices proved attractive for dealing with outputs and inputs like these as reported for each of the schools included in the Program Follow Through experiment.

Farrell's empirical work had been confined to single-output cases and his sketch of extensions to multiple outputs did not supply what was required for applications to large data sets like those involved in Program Follow Through. To obtain what was needed in computationally implementable form, Charnes, Cooper and Rhodes developed the dual pair of linear programming problems.

This approach has been successfully applied where conventional methods did not perform well. DEA has also been acclaimed as a leading edge method that supports benchmarking, continuous improvement and strategic analysis.

Thus DEA has been found to be particularly suitable in solving the following three basic performance questions that any organization is faced with:

· How well are we doing relative to the others doing the same things as we do?

· What do we need to improve?

· Who are the best- in-class performers for benchmarking purposes?

DEA is a special application of linear programming. In recent years it has become an increasingly valuable tool for making provider comparison.

DEA provides considerable flexibility in data selection. The inputs and outputs can be continuous, ordinal or categorical variables. DEA also allows the inputs and outputs to be measured in different units e.g. in Ghana Cedis, kilometres, etc. The term *output* in DEA can be broadly interpreted to mean not only output performance measure but also quality performance or any outcome performance measure. Likewise efficiency can be broadly interpreted to mean not only an assessment of efficiency but also an assessment of quality and effectiveness (outcome). Consequently, DEA can make efficiency assessment, quality assessment, effectiveness assessment and any of this combination.

In terms of making service provider comparisons, DEA has two advantages of interest for benchmarking purposes. First, it can make simultaneous comparisons of multiple dependant performance measures (output, quality, and outcome) and can provide a measure of best overall practice. Providers (or producers) can be simultaneously measured in an economic sense both for *allocative* and *technical* efficiency (or performance). "Allocative" efficiency is measured through the use of cost data (e.g. cost per output, cost per quality or cost per outcome). "Technical" efficiency is measured by using resource data (e.g. output per fuel consumed etc). The results are efficiency scores in the pareto sense i.e. those providers found to be inefficient are strictly inefficient because at least one other provider can produce output, quality or outcomes for less input. Secondly, DEA calculates the amount of resources that could be saved, or conversely the amount of additional output that could have been produced if the provider in question operated using the techniques of the best practice providers.

Thus DEA does not merely state who operates inefficiently but offers the inefficient providers the opportunity to learn from best practice providers.

The method can successfully be applied to profit and non-profit making organizations, as well. The performance of a unit is evaluated by comparing its performance with the best performing units of the sample. Best performing units form the efficiency frontier. If the unit is not on the efficiency frontier it is considered to be inefficient**.** Hence, DEA is called frontier analysis. The aim of DEA is to quantify the distance to the efficient frontier for every DMU. The measure of performance is expressed in the form of efficiency score. After the evaluation of the relative efficiency of the present set of units, the analysis shows how inputs and outputs have to be changed in order to maximize the efficiency of the target DMU.

Charnes, Cooper and Rhodes proposed that each unit should be allowed to adopt the most favorable set of weights. The linear program solution technique will attempt to make the efficiency of the unit as large as possible. This search procedure will terminate when some of the efficiencies hit 1.

DEA gives the weights of inputs and outputs leading to the calculated efficiency. The unit is efficient if the efficiency is equal to 1 and inefficient if it is less than 1. If represented graphically, for a given set of units, the efficient DMUs form the frontier that encloses the inefficient ones (the whole data set). Hence the name of analysis - Data Envelopment Analysis. So, the efficient units use its mix of inputs better than inefficient ones or the efficient units manage to produce more outputs using a given mix of inputs. An input-oriented measure quantifies the input reduction, which is necessary for a DMU to become efficient, holding the output constant. Similarly, an output-oriented measure quantifies the necessary output expansion, holding the input constant. A non-oriented measure quantifies the improvements when both inputs and outputs can be modified simultaneously. DEA suggest the creation of virtual unit B' for the inefficient unit B. B' lies on the efficient frontier and is the best practice for unit B, if it aims to be efficient. The outputs and inputs of such a virtual unit are linear combinations of corresponding outputs and inputs of all other units. Thus DEA gives inputs/outputs targets for inefficient units – a benchmarks. The benchmark represents the peer group for the inefficient DMU.

Since the technique was first proposed much theoretical and empirical work has been done. Many studies have been published dealing with applying DEA in real-world situations. The most important task is to determine the proper set of inputs and outputs for the observed units. Having reviewed literature on economics of hospitals, it concluded that the authors use tree categories of inputs: labor, supplies and capital. Labor is number of physicians, surgeons, nurses, technical staff; the suppliers are pharmaceutical and others; capital includes equipment, vehicles and building space. There are four types of outputs: inpatient days, outpatient visits, surgical operations, and live births. When DEA is undertook to evaluate bank branch efficiency inputs are: staff, interest costs, non-interest costs – expenses for rent, electricity, printing, advertising, post and telephone, repair and maintenance, etc. and the outputs are: number of transactions – deposits, loans, advances, mortgages etc.

One of the strengths of DEA is the fact that inputs and outputs can be measured in different units for example dollars, square meters, number of staff, etc. The analysis can be run using one input and several outputs or vice versa estimating one output produced by multiple inputs. DEA can be run with a very small data set, as is the case in this project work.

The first and probably most difficult step in efficiency evaluation is to decide which inputs and outputs data should be included.

The literature on applying the DEA technique to library evaluation shows various schemes of inputs and outputs sets. The inputs usually are library staff (Chen 1997; Sharma, Leung and

Zane 1999), weekly hours (Vitaliano 1998), volumes held (Shim 2000), book collection (Sharma, Leung and Zane 1999), material resources (Easun 1992). The most frequently used outputs are total circulation, reference transactions, library visits, interlibrary lending, online search and provision of information. The inputs or outputs that can be controlled by the DMUs are called "standard" or "discretionary" variables. "Nondiscretionary" variables are beyond the control of library administration, like population density, area size, resident population, nonresidential borrowers, and socioeconomic indices.

1.2 PROBLEM STATEMENT

Researchers recognize two broad aspects of evaluating library performance: "effectiveness" and "efficiency." Effectiveness here means the extent to which library services meet the expectations or goals set by the organization. In the library field, there has been a growing desire to measure effectiveness in terms of impact of library services on their users.

The second aspect of library performance measurement, "efficiency," measures the library's ability to transform its inputs (resources) into production of outputs (services), or to produce a given level of outputs with the minimum amount of inputs. The efficiency aspect of library performance has received less attention in the library literature, but it is an immediate concern for decision-makers at the parent institution.

The success of the library, like that of other organizations, depends on its ability to behave both effectively and efficiently.

We can put these two dimensions of library performance in a 2 by 2 matrix as shown in table 1.2.1

Table 1.2.1: Matrix of Performance

HIGH EFFECTIVENESS

LOW EFFICIENCY	Effective but Excessively Costly	Best All-around Performers	HIGH EFFICIENCY
	Problematic, Underperforming	Efficiently Managed for Insignificant Results	

LOW EFFECTIVENESS

Performance improvement requires constant and careful monitoring and assessment of library activities and operating environments. This, in turn, requires the development of proper measurement tools or devices. This study assesses the technical efficiency of academic research libraries that are members of the Association of Research Libraries using a complex tool called DEA. While the development of effectiveness is equally important, this study is focused solely on measuring library efficiency.

1.3 RESEARCH OBJECTIVE

Different techniques have been applied to study the efficiency of a group of organizations or operating units. Data Envelopment Analysis (DEA) is one of the most widely used approaches because of its sound mathematical basis and non-parametric nature. This research therefore focuses primarily on application of the DEA technique for evaluating the efficiency of library service systems. The research has the following two specific objectives:

- Assess the suitability of the Data Envelopment Analysis (DEA) approach for evaluating the efficiency of library systems using data from the University Libraries;
- Identify the best performers to serve as a benchmark for the non-performers of the University Library systems.

1.4 RESEARCH METHODOLOGY

This research proposes to utilize the Data Envelopment Analysis (DEA) method to quantify the efficiency score of individual University Library Systems

The investigation involves mainly two steps as follows:

- Firstly, data of University Library Systems is analyzed in order to define the inputs and outputs for evaluating the efficiency.
- Secondly, DEA model is applied to calculate the efficiency score, and Linear Programming- Simplex Applet algorithm' developed by Pedro Miguel Silva and Tiago Castro Guise -*Version 1.0 - Lisbon, July 1998, updated on October 1999* is employed for further analysis.

1.5 JUSTIFICATION

Increasingly, scientific libraries are becoming aware of the importance of managerial aspects in library management. University libraries in particular are skilled as modern service providers which aim to meet demands as well as possible. As regards performance assessment, library performance can be assessed by surveying library users' opinions (subjective component) and/or analyzing library performance indicators (objective component). For more than twenty years, librarians, especially in the United States and Great Britain, have used performance indicators to analyze library performance (see e.g. Brophy (1989), King Research Ltd (1990), McDonald et al. (1994), Poll (1993) and Van House et al. (1990)). Comparable studies in Austria, Germany or Switzerland are mostly confined to drawing up simple library statistics (see e.g. Boekhorst (1995), Deutsches Bibliotheksinstitut (1996) and Poll (1992)). As Ghanaian universities and university libraries in particular, are increasingly confronted with limited resources, the pressure to efficiently provide library services is rapidly increasing. Therefore, it is interesting to see whether or not the performance of Ghanaian University libraries is significant. It is also interesting whether or not there are efficiency differentials between library groups, particularly between small and large libraries and between libraries in Ghana. The analysis of performance differentials between libraries is motivated by the assumption that, due to different environmental conditions (e.g. more strongly developed “surrogate” market forces, continuous performance assessment, etc.), libraries (are forced to) perform better than others. As regards the comparison of libraries of different sizes, we assume that small libraries might be less efficient because of sub-optimal scale sizes. To check these assumptions, this study aims to investigate library performance, particularly to provide information on the ranking of individual libraries, on potential benchmarks for less efficient libraries and on efficiency differentials between different libraries.

As the use of single factor productivity measures might be unsatisfactory because these measures reflect only partial aspects of the library performance spectrum, we apply DEA to be able to assess library performance using a measure which depicts total factor productivity. DEA has become established as a tool to judge performance, especially of non-profit institutions such as university libraries where profit measures are of little value, especially due to the absence of output prices. Analysing the potential of DEA, Shim (2000) found out that the DEA technology is a suitable tool to benchmark re-search libraries in the United States. The potential of DEA goes beyond the calculation of a scalar measure of efficiency: it is part of a continuous learning process, including the thorough analysis of the service production process to identify candidate measures of inputs and services at first, the calculation of efficiency measures based on the selected input/output measures, the discussion of efficiency results and finally the derivation of strategies to improve performance via learning from the best practice performer identified through DEA. Using DEA we therefore obtain a first insight into the efficiency differentials among libraries.

1.6 STRENGTH and LIMITATIONS of DEA

As the earlier list of applications suggests, DEA can be a powerful tool when used wisely. A few of the characteristics that make it powerful are:

- DEA can handle multiple input and multiple output models.
- It doesn't require an assumption of a functional form relating inputs to outputs.
- DMUs are directly compared against a peer or combination of peers.
- Inputs and outputs can have very different units. For example, x_1 could be in units of lives saved and x_2 could be in units of Ghana Cedis without requiring an a priori tradeoff between the two.

The same characteristics that make DEA a powerful tool can also create problems. An analyst should keep these limitations in mind when choosing whether or not to use DEA.

- Since DEA is an extreme point technique, noise (even symmetrical noise with zero mean) such as measurement error can cause significant problems.
- DEA is good at estimating "relative" efficiency of a DMU but it converges very slowly to "absolute" efficiency. In other words, it can tell you how well you are doing compared to your peers but not compared to a "theoretical maximum."
- Since DEA is a nonparametric technique, statistical hypothesis tests are difficult and are the focus of ongoing research.
- Since a standard formulation of DEA creates a separate linear program for each DMU, large problems can be computationally intensive.

1.7 SCOPE OF WORK

This project work is only measuring the efficiency of selected public University Library from Kwame Nkrumah University of Science and Technology, Kumasi, University of Cape Coast, Cape Coast and University of Ghana, Legon-Accra. The analysis is based on data collected for 2010 and has nothing to do with the previous/future performance or neither efficiency nor does is have any effect on the other University libraries (public or private).

1.8 ORGANIZATION OF THE THESIS

This thesis is organized into five chapters:

- Chapter **1** introduces the problems to be addressed and the research objectives in this thesis. This chapter also includes general information of research methodology and the structure of the thesis.
- Chapter **2** provides background introduction of DEA and its application to the measurement of efficiency of libraries, also some recent studies on the topic and applications of DEA.
- Chapter **3** explains how DEA is applied to measure efficiency in this research and further study on the factors that may affect the efficiency score.
- Chapter **4** discusses the data collection and analysis of the problem
- Chapter **5** summarizes the results of this study conclude, and make recommendations. It also discusses the possible future research that needs to be studied on library efficiency topic.

CHAPTER 2

LITERATURE REVIEW

There is extensive literature on Data Envelopment Analysis (DEA), applied to a wide diversity of field in the evaluation of efficiency.

Yongjun Shen et al. (2009) evaluated Road Safety Performance using Data Envelopment Analysis.

A data set consisting of 21 indicators for 26 European countries is analyzed, and data envelopment analysis (DEA) as a performance measurement technique is applied to combine these 21 indicators into a composite index. In particular, the concept of hierarchical structure is embodied in the model thereby giving a more detailed insight into the layered hierarchy of the indicators. In addition, the presence of both quantitative and qualitative indicators is taken into account. In the end, a separate, best possible model is constructed for each country, the most optimal road safety index score is computed, and the weights assigned to each layer of the hierarchy are analyzed.

Chin Huan Huan et al. (2008) applied Data Envelopment Analysis to evaluate the Municipal Solid Waste Management Projects in Metro Manila, Taiwan.

In this paper the authurs in consultations with a Municipal Solid Waste Management (MSW) expert group, this study elucidates how governmental officials can solve the problems surrounding municipal solid waste management in Metropolitan-Manila. A crucial related issue is how the expert group can better evaluate MSW solutions and select favorable ones better evaluate and select a favourable MSW solution using a series of criteria.The study applies cost-benefit analysis (CBA) and data envelopment analysis (DEA) to determine the benefits and cost / input and output technical efficiency of alternative projects, which affords financial data information that evaluators can use for economic decision-making regarding MSW projects. Results of this study suggest that the thermal process technology is less efficient than resource recovery using DEA.

N. K.Womer et al. (2006) evaluated Benefit-cost analysis using data envelopment analysis

” The paper develops an approach for conducting benefit-cost analysis derived from data envelopment analysis (DEA).

The models and methodology proposed give decision makers a tool for evaluating alternative policies and projects where there are multiple constituencies who may have conflicting perspectives.

Chun-Yu Lin et al. (2007) used Data Envelopment Analysis for Product Line Selection

They define product line selection problem as selecting a subset of potential product variants that can simultaneously minimize product proliferation and maintain market coverage. This paper proposes a method based on Data Envelopment Analysis (DEA) for product line selection. In this study, they construct a five steps method that systematically adopts DEA to solve a product line selection problem. They then apply the proposed method to an existing line of staplers to provide quantitative evidence for managers to generate desirable decisions to maximize the company profits while also fulfilling market demands.

Mishra and Gokulananda (2009) applied Data Envelopment analysis to Suppliers Development Strategies.

This study developed an application guideline for the assessment, improvement, and control of quality in Supply Chain Management using Data Envelopment Analysis. Improvement in the quality of all supply chain processes lead to cost reductions as well as service enhancement. The data is collected from 25 suppliers of food and agro based industry for the analysis.

Joost Schalken et al. used Data Envelopment Analysis in measuring IT Infrastructure Project Size: Infrastructure Effort Points.

The objective of the research was to design a metric that can be used to measure the size of projects that install and configure commercial-of-the-shelf components COTS stand-alone software, firmware and hardware components.

Xiangyu Wang (2010) used Data Envelopment Analysis to measure the performance of 25 insurance companies. This Project focused on a linear programming model used in performance evaluation of 25 property and casualty insurance companies as of the year of 2007. The goal was to determine the efficiency of each company compared to the peer competitors within property and casualty insurance industry. The technique is called data envelopment analysis (DEA). The emphasis was on data selection and cleanup, mathematical approach behind the data envelopment analysis model, and the application of this model to the efficiency comparison.

Mohammad S. etal. (2009) Utilized data envelopment analysis to benchmark safety performance of construction contractors.

The purpose of this paper was to utilize data envelopment analysis (DEA) to benchmark safety performance of construction contractors.

DEA has been recognized as a robust tool that is used for evaluating the performance of business organizations. The proposed approach is deployed based on empirical data collected from 45 construction contractors. On a scale of 0–1.0, DEA analysis assesses the relative efficiency of every contractor relative to the rest of the contractors in terms of safety performance. For inefficient contractors, DEA analysis provides quantitative guidance on how to become efficient.

Brenda McCabe et al. (2004) measured Construction Prequalification using data envelopment analysis.

Data envelopment analysis (DEA) had been recognized as a useful technique to prequalify contractors by assigning relative efficiency scores. Data envelopment analysis, however, usually requires a large amount of data and has not been fully developed to achieve reliable results. An enhanced contractor prequalification model using DEA was developed together with a methodology for determining a "practical frontier" of best contractors.

The established practical frontier can be used as a regional performance standard for the owner in prequalification and as improvement guidelines for contractors.

M. Francesca Cracolici et al. (2006) made an assessment of Tourist Competitiveness by analysing destination efficiency.

Recently the notion and the measurement of destination competitiveness have received increasing attention in the economics literature on tourism. The reason for this interest emerges from both the increasing economic importance of the tourist sector and the increasing competition on the tourist market as a consequence of the transition from mass tourism to a new age of tourism that calls for a tailor-made approach to the specific attitudes and needs of tourists. The central subject of this paper concerned the efficiency of tourist site destinations. Using a dataset of 103 Italian regions for the year 2001, an economic efficiency analysis based on a production frontier approach was made in the study. The study deploys a measure of tourist site competitiveness in terms of its technical efficiency using parametric and non-parametric methods, a stochastic production function and data envelopment analysis, respectively.

Paul Lau Ngee Kiong et al. (Universiti Teknologi MARA Sarawak) measured school performance using Data Envelopment Analysis.

This paper reports the findings of a project that examined the determinants contributing to school efficiency based on the 'High Standard Quality Education' framework (PKSBSTKP) compiled by the Malaysian School Inspectorate and evaluated the relative efficiencies of all the secondary schools in the Sri Aman/Betong Division for the year 2002. The research reveals that the PKSBSTKP performances of all the participating schools were average. The frontier analysis shows that 9 schools were efficient and 7 schools were inefficient. There were obvious differences in evaluating performances of schools by DEA and the methodology of PKSBSTKP. However, there was no significant difference in the efficiencies between schools in the urban and rural area of the Sri Aman/Betong Division. Based on slack analysis, the output maximization BCC model shows that all the three principal inputs, student quality, managerial quality and school facilities were of equal importance. For the output variables, school uniqueness topped the list, followed by students' academic performance, and change of academic performance and achievement in co-curriculum.

Dr. Manuel SALAS-VELASCO (University of Granada, Spain,2006) evaluated the performance private and public schools.

Data envelopment analysis (DEA) was used to investigate the technical efficiency of the public-sector funded schools in Spain. Particular attention is paid to the role of uncontrollable factors—such as the socioeconomic status of the schools—as inputs into the production process. The dataset used came from the OECD Programme for International Student Assessment (PISA). He concluded that publicly funded private schools in Spain, operating with similar amounts of money, are more efficient than public schools. These private schools are free of bureaucratic constraints that encumber public schools, and are able to control many more decisions at the school level.

Preeti Tyagiet al. (Department of Mathematics, IIT, Roorkee, India) measured the efficiency of schools in India.

This paper assessed the technical efficiency and efficiency differences among 348 elementary schools of Uttar Pradesh state in India by a linear programming based technique, Data Envelopment Analysis (DEA). They assessed the schools with eight inputs and three outputs. Inputs include school resources (teaching, physical and ancillary facilities, teachers' qualities) and home environment of schools' students (parents' education and occupation) while output comprise school wise average marks in environment studies, mathematics, language. In preparing these inputs and outputs, Principal Component Analysis is used.

Moffat and Abbas (2009) measured the efficiency of financial institutions in Botswana.

This paper examined technical and pure technical efficiencies of ten major financial institutions in Botswana for each year during the period 2001-2006 using data envelopment analysis. In order to obtain more robust and reliable results, the sensitivity of their efficiency indices were put into test by choosing three alternative approaches in specifying the mix of inputs and outputs. The empirical results indicate that: (a) no matter which approach and year are taken into consideration, Bank of Baroda and First National Bank (which are both foreign banks) and Botswana Savings Bank (which is a publicly owned institution) are consistently among the most efficient institutions and Botswana Development Corporation, African Bank Corporation and National Development Bank are the least efficient ones; (b) the most efficient banks are either small or large institutions in terms of their asset sizes; (c) due to the small sample size, the evidence of a relationship between the age of institutions and their technical efficiencies remains inconclusive. One can conclude that financial institutions can further enhance efficiency by adopting self-service technologies such as telephone and internet banking which can substantially reduce their service delivery costs.

Jingtao Yang (2005) quantified the Technical Efficiency of Canadian Library Systems Using Data Envelopment Analysis.

This research focused on evaluating the level of efficiency of individual library systems in Canada with the specific objective of identifying the most efficient service agencies and the sources of their efficiency. By identifying the most efficient systems along with the influencing factors, it is possible that new service policies and management and operational strategies could be developed for improved resource utilization and quality of services. To achieve this objective, this research applied the analysis methodology called Data Envelopment Analysis (DEA) approach which is a mathematical programming based technique for determining the efficiency of individual systems as compared their peers involving multiple performance measures. Annual operating data from Canadian Urban Transit Association (CUTA) for Canadian library systems of year 2001, 2002 and 2003 were used in this analysis. Regression analysis was performed to identify the possible relationship between the efficiency of a library system and some measurable operating, managerial and other factors which could have an impact on the performance of library systems. The regression analysis also allows for the calculation of confidence intervals and bias for the efficiency scores in order to assess their precision.

Ahmed Salem Al-Eraqi etal. (2007) evaluated the Location Efficiency of Arabian and African Seaports Using Data Envelopment Analysis (DEA).

In this paper the efficiency and performance is evaluated for 22 seaports in the region of East Africa and the Middle East. The aim of the study is was to compare seaports situated on the maritime trade road between the East and the West. These are considered as middle distance ports at which goods from Europe and Far East/Australia can be exchanged and transhipped to all countries in the Middle East and East Africa. All these seaports are regional coasters, and dhow trade was built on these locations, leading this part of the world to become an important trade centre. Data was collected for 6 years (2000-2005) and a non-parametric linear programming method, DEA (Data Envelopment Analysis) is applied. The ultimate goal of the study is: 1) to estimate the performance levels of the ports under consideration. This will help in proposing solutions for better performance and developing future plans. 2) to select optimum transhipment locations. Goncalves and Ricardo (2006) measured Management efficiency Using Data Envelopment Analysis (DEA) estimation for Banks in Brazil. This paper presented a new paradigm approach for quantifying a bank's managerial efficiency, using a data envelopment analysis (DEA) model that combines multiple inputs and outputs to compute a scalar measure of efficiency and management quality. The analysis of the largest 50 Brazilian banks over a twelve-year period from 1995 to 2006 shows significant differences in management quality scores between institutions. Hence, this new metric provides an important, but previously missing, modelling element for the early identification of troubled banks and can be used as a tool for off-site bank supervision in Brazil.

Susanne Rassouli-Currier (University of Central Oklahoma) assessed the efficiency of Oklahoma public schools.

In this paper, the efficiency of the Oklahoma school districts using two different specifications is measured by the Data Envelopment Analysis (DEA) method. To determine the possible sources of inefficiency, a second stage Tobit regression was employed. Here, the specification of the inefficiency models includes(1) environmental variables that school districts have no control over (e.g., the percentage of students in special education and the poverty rate in the district) and (2) non-traditional inputs that school districts do have control over (e.g., teachers'salaries) but were not included in the first stage DEA. The findings of the models were compared and both suggest that the key factors affecting efficiency measures among the Oklahoma school districts are primarily the students' characteristics and family environment.

For more than 20 years, libraries have been confronted with performance comparisons. Numerous publications deal with the theoretical development of performance indicators to cover the libraries' range of activities (e.g. Moore (1989), Poll and Boekhorst (1996), Brophy (1989), Ceynowa (2001), Crawford et al. (1998), Mundt and Guschker (2003) and Van House etal., (1990)). Other publications concentrate on the empirical assessment of library performance. Many authors use established performance indicators, such as the number of circulations per student or the number of requests processed per employee, with the main disadvantage being that each performance indicator only covers partial performance. Berghaus-Sprengel (2001) discussed the limits of benchmarking against the background of using multiple performance indicators in empirically assessing performance of university libraries.

Recently, several studies have attempted to derive an aggregate performance indicator based on the analysis of the overall performance of university libraries. From a methodological point of view many authors fell back on DEA: Chen (1997) compared Taiwanese university libraries, Kao and Lin (1999) particularly investigated the effect of library size on library performance, Kao and Liu (2000) addressed the problem of missing data in DEA-based performance assessment. Shim and Kantor (1998) and Shim (2000, 2003) provided an overview of the possibilities of DEA for library benchmarking. They discussed in detail the strengths and weaknesses of DEA in the context of library performance evaluation, thereby covering the fundamental problems of finding suitable input and output indicators.

Reichmann (2004) in his paper analyses the technical efficiency of 118 randomly selected university libraries from German-speaking countries (Austria, Germany, Switzerland) and English-speaking countries (the United States, Australia and Canada) using Data Envelopment Analysis (DEA). DEA efficiency scores were calculated using library staff, measured in fulltime equivalents, and book materials held as inputs, and the number of serial subscriptions, total circulations, regular opening hours per week, and book materials added as outputs. Reichmann and Sommersguter-Reichmann (2006) addressed the problems of differing environments and their effects on library performance. Vitaliano (1998) and Worthington (1999) analyzed the performance of public libraries in New York (Vitaliano) and New South Wales, Australia (Wor-thington).

Kao and Lin (1999) compared the performance of University libraries of different University sizes. In an attempt to solve these problems, they adopted the concept of the Pareto optimality (Ferguson and Gould 1980; Zeleny 1982) to calculate the expected resources and services (hereafter generalised as services) to be provided by the university libraries of different university sizes from sampled university libraries.

In their paper, Akdede and Kazancoglu (2006) investigated the relative efficiency of public (state) libraries of major cities in Turkey by applying a data envelopment analysis. Scale, technical, and overall efficiency scores are calculated. It is found that there is a negative correlation between economic and social development index of the cities and efficiency scores of state libraries of same cities.

Wonsik Shim, Kantor (1998) used DEA approach to evaluate digital libraries.

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

This chapter deals with one of the most basic DEA models, the *CCR model* which was initially proposed by Charnes, Cooper and Rhodes (CCR) in 1978. Tools and ideas commonly used in DEA are also introduced.

Here, for each Decision Making Unit (DMU), we formed the input and output by (yet unknown) weights (v_i) and (u_r)

$$Input \quad = v_1 x_{1o} + \cdots + v_m x_{mo}$$

$$Output \quad = u_1\, y_{1o} + \cdots + u_s\, y_{so}$$

Then we tried to determine the weight, using linear programming so as to maximize the ratio

$$\frac{output}{input}$$

To deal with multiple inputs and outputs, a ratio like the following may be used.

$$\frac{\sum_{r=1}^{s} u_r y_r}{\sum_{i=1}^{m} v_i x_i} = \frac{u_1 y_1 + u_2 y_2 + u_3 y_3 + \ldots + u_s y_s}{v_1 x_1 + v_2 x_2 + v_3 x_3 + \ldots + v_m x_m}$$

where

y_r = amount of output r

u_r = weight assigned to output r

x_i = amount of input i

v_i = weight assigned to input i.

The weights may be (1) fixed in advance or (2) derived from the data. The former is sometimes referred as an *a priori* determination.

The optimal weights may (and generally will) vary from one DMU to another DMU. Thus, the "weights" in DEA are derived from the data instead of being fixed in advance. Each DMU is assigned a best set of weights with values that may vary from one DMU to another.

3.2 DATA

In DEA, the organization under study is called a (Decision Making Unit). The definition of DMU is rather loose to allow flexibility in its use over a wide range of possible applications. Generically a DMU is regarded as the entity responsible for converting inputs into outputs and whose performances are to be evaluated. In managerial applications, DMUs may include banks, department stores and supermarkets, and extend to car makers, hospitals, schools, public libraries and so forth. In engineering, $DMUs$ may take such forms as airplanes or their components such as jet engines. For the purpose of securing *relative* comparisons, a group of DMUs is used to evaluate each other with each DMU having a certain degree of managerial freedom in decision making.

Suppose there are $n\ DMUs$: $DMU_1, DMU_2, \ldots, DMU_n$. Some common input and output items for each of this $j = 1,2,\ldots,n\ DMUs$ are selected as follows:

1. Numerical data are available for each input and output, with the data assumed to be positive for all$DMUs$.
2. The items (inputs, outputs and choice of $DMUs$) should reflect an analyst's or a manager's interest in the components that will enter into the relative efficiency evaluations of the $DMUs$.
3. In principle, smaller input amounts are preferable and larger output amounts are preferable so the efficiency scores should reflect these principles.
4. The measurement units of the different inputs and outputs need not be congruent. Some may involve number of persons, or areas of floor space, money expended, etc.

Suppose *m* input items and *s* output items are selected with the properties noted in 1 and 2.

Let the input and output data for DMU_j be $(x_{1j}, x_{2j}, \ldots, x_{mj})$ and $(y_{1j}, y_{2j}, \ldots, y_{sj})$ respectively.

The input data matrix X and the output data matrix Y can be arranged as follows.

$$X = \begin{pmatrix} x_{11} & x_{12} \cdots & x_{1n} \\ x_{21} & x_{22} \cdots & x_{2n} \\ \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} \cdots & x_{mn} \end{pmatrix} \tag{1}$$

and

$$Y = \begin{pmatrix} y_{11} & y_{12} \cdots & y_{1n} \\ y_{21} & y_{22} \cdots & y_{2n} \\ \vdots & \vdots & \vdots \\ y_{s1} & y_{s2} & y_{sn} \end{pmatrix} \tag{2}$$

where X is an $(m\ x\ n)$ matrix and Y and $(s\ x\ n)$ matrix.

3.3 THE CCR MODEL

To allow for applications to a wide variety of activities, we use the term Decision Making Unit $(= DMU)$ to refer to any entity that is to be evaluated in terms of its abilities to convert inputs into outputs. These evaluations can involve governmental agencies and not-for-profit organizations as well as business firms. The evaluation can also be directed to educational institutions and hospitals as well as police forces (or subdivision thereof) or army units for which comparative evaluations of their performance are to be made.

We assume that there are *n* DMUs to be evaluated. Each DMU consumes varying amounts of *m* different inputs to produces different outputs. Specifically, $\boldsymbol{DMU_j}$ consumes amount x_{ij} of input i and produces amount y_{rj} of output r. We assume that $x_{ij} \geq 0$ and $y_{rj} \geq 0$ and further assume that each DMU has at least one positive input and one positive output value.

We now turn to the "ratio-form" of DEA. In this form, as introduced by Charnes, Cooper, and Rhodes, the ratio of outputs to inputs is used to measure the relative efficiency of the DMU_j = DMU_o to be evaluated relative to the ratios of all of the $j = 1, 2, \ldots, n$. We can interpret the

CCR construction as the reduction of the multiple-output /multiple-input situation (for each DMU) to that of a single 'virtual' output and 'virtual' input. For a particular DMU the ratio of this single virtual output to single virtual input provides a measure of efficiency that is a function of the multipliers. In mathematical programming parlance, this ratio, which is to be maximized, forms the objective function for the particular DMU being evaluated, so that symbolically $max\ \theta = \sum_r u_r y_{ro} \Big/ \sum_i x_i v_{xio}$

where it should be noted that the variables are $the\ u_r$'s and the v_i's and the y_{ro}'s and x_{io} 's are the observed output and input values, respectively, of DMU_o, the DMU to be evaluated.

Of course, without further additional constraints (developed below) it is unbounded.

Given the data, we measure the efficiency of each DMU once and hence need n optimizations, one for each DMU_j to be evaluated.

Let the DMU_j to be evaluated on any trial be designated as DMU_o where *o* ranges over $1, 2, \dots, n$.

We solve the following fractional programming problem to obtain values for the input "weights" v_i $\{i = 1,\dots,m)$ and the output "weights" u_r $(r = 1,\dots,s)$ as variables.

$$(FP_o)\ \max \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \cdots + u_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \cdots + v_m x_{mo}} \quad (3)$$

subject to

$$\frac{u_1 y_{1j} + \dots + u_s y_{sj}}{v_1 x_{1j} + \dots + v_m x_{mj}} \le 1 \quad j = 1,2,\dots,n \quad (4)$$

$$v_1, v_2, \dots, v_m \ge 0 \quad (5)$$

$$u_1, u_2, \dots, u_s \ge 0 \quad (6)$$

The constraints mean that the ratio of "output" vs." input" should not exceed 1 for every DMU. The objective is to obtain weights (v_i) and (u_r) that maximize the ratio of $DMUo$, the DMU being evaluated. By virtue of the constraints, the optimal objective value $\theta^*$is at most 1. Mathematically, the nonnegativity constraint (5) is not sufficient for the fractional terms in (4) to have a positive value. We do not treat this assumption in explicit mathematical form at this time. Instead we put this in managerial terms by assuming that all outputs and inputs have some nonzero worth and this is to be reflected in the weights u_r and v_i being assigned some positive value.

3.4 FROM FRACTIONAL TO A LINEAR PROGRAM

We now replace the above fractional programs (FPo) by the following linear program (LPo).

$$(LP_o) \quad \max \theta = u_1 y_{1o} + u_2 y_{2o} + \cdots + u_s y_{so} \quad (7)$$

subject to $v_1x_{1o} + v_2x_{2o} + \cdots + v_mx_{mo} = 1$ (8)

$$u_1y_{1j} + u_2y_{2j} + \cdots + u_sy_{sj} \leq v_1x_{1j} + v_2x_{2j} + \cdots + v_mx_{mj} \quad (9)$$

$$(j = 1, \ldots, n)$$

$$(v_1, v_2, \ldots, v_m \geq 0) \quad (10)$$

$$(u_1, u_2, \ldots, u_s \geq 0) \quad (11)$$

Under the nonzero assumption of v and $X > 0$, the denominator of the constraint of (FPo) is positive for every j , and hence we obtain (3) by multiplying both sides of (4) by the denominator. Next, we note that a fractional number is invariant under multiplication of both numerator and denominator by the same nonzero number. After making this multiplication, we set the denominator of (3) equal to 1, move it to a constraint, as is done in (8), and maximize the numerator, resulting in (LPo). Let an optimal solution of (LPo) be

($v = v*, u = u*$) and the optimal objective value θ^*. The solution ($v = v^*, u = u^*$) is also optimal for (FPo), since the above transformation is reversible under the assumptions above. (FPo) and (LPo) therefore have the same optimal objective value θ^*.

Theorem (Units Invariance Theorem): The optimal values of max $\Theta = \Theta^*$ in (3) and (7) are independent of the units in which the inputs and outputs are measured provided these units are the same for every DMU (Cooper et al., Introduction to DEA and its uses.)

Thus, one person can measure outputs in miles and inputs in gallons of gasoline and quarts of oil while another measures these same outputs and inputs in kilometers and liters. They will nevertheless obtain the same efficiency value.

3.5 CCR DEFINITION

1. $DMUo$ is CCR-efficient *if* $\theta^* = 1$ *and there exists at least one optimal* (v^*, u^*), *with* $v^* > 0$ *and* $u^* > 0$.

2. *Otherwise,* $DMUo$ *is* CCR-inefficient.

Thus, CCR-inefficiency means that either (i) $\theta^* < 1$ or (ii) $\theta^* = 1$ and at least one element of (v^*, u^*) is zero for every optimal solution of (LPo).

3.6 MEANING OF OPTIMAL WEIGHTS

The (v^*, u^*) obtained as an optimal solution for (LPo) results in a set of optimal weights for the DMUo. The ratio scale is evaluated by:

$$\theta^* = \frac{\sum_{r=1}^{n} u_r^* \; y_{ro}}{\sum_{i=1}^{m} v_i^* \; x_{io}}$$

From (8), the denominator is 1 and hence

$$\theta^* = \sum_{r=1}^{s} u_r^* y_{ro}$$

As mentioned earlier, (v^*, u^*) are the set of most favorable weights for the $DMUo$ in the sense of maximizing the ratio scale, v_i^* is the optimal weight for the input item i and its magnitude expresses how highly the item is evaluated, relatively speaking. Similarly, u_r^* does the same for the output item r.

Furthermore, examining each item $v_i^* x_i$ in the input $\sum_{i=1}^{m} v_i^* x_{io} = (1)$, then we can see the relative importance of each item by reference to the value of each$v_i^* x_{io}$.

The same situation holds for $u_r^* y_{ro}$ where the u_r^* provides a measure of the relative contribution of y_{ro} to the overall value of θ^*.

These values not only show which items contribute to the evaluation of $DMUo$, but also to what extent they do so.

CHAPTER 4
DATA COLLECTION AND ANALYSIS

4.1 INTRODUCTION

This chapter describes the data used, how to apply DEA to calculate the efficiency score of University library system and the analysis of the factors which may be associated with the efficiency score.

The process starts from understanding and analysis of the data. After defining the inputs and outputs from the source data, a Linear Programming- Simplex Applet algorithm developed by Pedro Miguel Silva and Tiago Castro Guise -*Version 1.0 - Lisbon, July 1998, updated on October 1999* is employed to help calculate the technical efficiency of University Library

4.2 DESCRIPTION OF DATA

The data used in this research is provided by the University. The data used in this analysis are Library-specific and cover the period from January 1, 2009 to December 31, 2010.

Since the University Library systems in Ghana have different needs, there is no uniform information requirement. As a result, the data may vary considerably. Also, fare structure, service policies, subsidy levels and the local operating environment may vary from system to system and from province to province. Therefore, cautions must be taken in comparing the performance of 3 different public University library systems. That is also the reason why this research focuses only on technical efficiency, leaving out the issue of economic efficiency. Another issue is that economic conditions, demographic trends, development activities and differences in urban spatial structure in Ghana can cloud comparisons of economic efficiency.

4.3 DESCRIPTION OF INPUTS AND OUTPUTS

The first step in a DEA is to identify the inputs and outputs that can be potentially used to define the efficiency of a University library system. Unlike many other industries where output (e.g. consumer products) is a clearly identifiable entity, the output of a library system in general can be quantified in various ways. The basic reason for this difference is that the ‘‘output’’ of a library system is service that cannot be stored for future use. Once service is produced, it ceases to exist regardless of whether it is consumed.

For the above research to be successful, one input namely printed edition expenses (PrEdEx) and two outputs namely number of registered readers (RR), and books borrowed (BB) have been selected to be used in the analysis.

Printed edition expenses are textbooks, dictionaries, periodicals (newspapers and journals) purchased by the university plus all printed edition given as grant by a foundation or a project. All outputs are measured in numbers.

The Decision Making Units (DMU's) that have been chosen for the research work are Kwame Nkrumah University of Science and Technology, Kumasi; University of Cape Coast, Cape Coast and University of Ghana, Legon Accra.

4.4 EFFICIENCY MODELING

With the inputs and outputs identified in the previous sections, the basic DEA model for a given library system can be formulated as follows:

$$(LPo)max\ \theta = u_1 y_{1o} + u_2 y_{2o} + \dots + u_s y_{so}$$

subject to $v_1 x_{1o} + v_2 x_{2o} + \cdots + v_m x_{mo} = 1$

$$u_1 y_{1j} + u_2 y_{2j} + \cdots + u_s y_{sj} \leq v_1 x_{1j} + v_2 x_{2j} + \cdots + v_m x_{mj}$$

$$(j = 1, \dots\dots, n)$$

$$v_1, v_2, \dots, v_m \geq 0$$

$$u_1, u_2, \ldots, u_s \geq 0.$$

y_r = amount of output *r*

u_r = weight assigned to output r

x_i = amount of input *i*

v_i = weight assigned to input *i*.

The software package a '***Linear Programming- Simplex Applet algorithm***' developed by Pedro Miguel Silva and Tiago Castro Guise -*Version 1.0 - Lisbon, July 1998, updated on October 1999* was used to solve the formulated linear programming problems (refer to APPENDIX A for the program and solution.)

Data collected: Table **4.4.1** shows the data collected for 2010:

DMU	INPUT- printed edition expenses(GHS)	OUTPUT 1- registered readers	OUTPUT 2- books borrowed
KNUST	12,771.41	8612	9113
UCC	54,046.87	4020	13277
UG	334,066.392	45685	9931

Table **4.4.2** shows the data collected for 2009:

DMU	INPUT- printed edition expenses(GHS)	OUTPUT 1- registered readers	OUTPUT 2- books borrowed

KNUST	36,128.76	4,345	8308
UCC	61,391.68	4,211	14,209
UG	371,184.88	53,580	1,366

The linear programming formulated out of the data for 2009 and 2010 are as follows:

For 2010:

- Max: KNUST = $8612u_1 + 9113u_2$;

Subject to: $12771.4v_1 - 1 = 0$;

$8612u_1 + 9113u_2 - 12771.4v_1 \leq 0$;

$4020u_1 + 13277u_2 - 54046.87v_1 \leq 0$;

$45685u_1 + 9931u_2 - 334066.392v_1 \leq 0$;

$u_1 \geq 0$; $u_2 \geq 0$; $v_1 \geq 0$;

- Max UCC = $4020u_1 + 13277u_2$;

Subject to: $54046.87v_1 - 1 = 0$;

$8612u_1 + 9113u_2 - 12771.4v_1 \leq 0$;

$4020u_1 + 13277u_2 - 54046.87v_1 \leq 0$;

$45685u_1 + 9931u_2 - 334066.392v_1 \leq 0$;

$u_1 \geq 0$; $u_2 \geq 0$; $v_1 \geq 0$;

- Max: UG = $45685u_1 + 9931u_2$;

Subject to: $334066.392v_1 - 1 = 0;$

$8612u_1 + 9113u_2 - 12771.4v_1 \leq 0;$

$4020u_1 + 13277u_2 - 54046.87v_1 \leq 0;$

$45685u_1 + 9931u_2 - 334066.392v_1 \leq 0;$

$u_1 \geq 0;$ $u_2 \geq 0;$ $v_1 \geq 0;$

For 2009 we have:

- Max KNUST = $4345u_1 + 8308u_2;$

Subject to: $36128.76v_1 - 1 = 0;$

$4345u_1 + 8308u_2 - 36128.76v_1 \leq 0;$

$4211u_1 + 14209u_2 - 90391.68v_1 \leq 0;$

$53580u_1 + 1366u_2 \leq 371184.88v_1;$

$u_1 \geq 0;$ $u_2 \geq 0;$ $v_1 \geq 0$

- Max UCC=$4211u_1 + 14209u_2;$

Subject to: $61391.68v_1 -1 = 0;$

$4345u_1 + 8308u_2 - 36128.76v_1 \leq 0;$

$4211u_1 + 14209u_2 - 61391.68v_1 \leq 0;$

$53580u_1 + 1366u_2 - 371184.88v_1 \leq 0;$

$u_1 \geq 0;$ $u_2 \geq 0;$ $v_1 \geq 0;$

- Max: UG = $53580u_1 + 1366u_2;$

Subject to: $371184.88v_1 - 1 = 0;$

$4345u_1 + 8308u_2 - 36128.76v_1 \leq 0;$

$4211u_1 + 14209u_2 - 90391.68v_1 \leq 0;$

$53580u_1 + 1366u_2 - 371184.88v_1 \leq 0;$

$u_1 \geq 0;$ $u_2 \geq 0;$ $v_1 \geq 0;$

Table 4.4.3 provides the solution results indicating the relative efficiency of individual library system for each year.

Table 4.4.3. Results of Linear Programming.

DMU	**2009**–Relative. Effic.	**2010**- Relative. Effic.	% change in Relative. Effic.(2010-2009)
KNUST	0.999	1.000	0.100%
UCC	1.000	0.344	-65.600%
UG	0.999	0.609	-39.039%

From these results, the following observations can be made;

- The technical efficiency of library systems varies significantly across systems with values ranging from 0.344 to 1.000. The average efficiency of all systems is 0.999 and 0.651 for year 2009 and 2010, respectively.

The variation over the two years is quite significant with a standard deviation of 0.246.

- Among all of the systems, the library systems operated by UCC for 2009 and that of KNUST for 2010 outperformed other systems (100% efficient) for the period. These systems are the best performers that other library systems may consider as a benchmark for improving their efficiency for the year under consideration. This is because the efficiency score is a measure of "relative" efficiency on how well or badly a library system is operated as compared to the most efficient ones.
- In terms of change in efficiency score over the two years, there were 2 system, UCC library and UG library, that had experienced noticeable decrease in efficiency. There was also one system, KNUST library whole efficiency scores had increased significantly (approximately 0.100%) per year
- The library system with the lowest efficiency score is offered by the UCC library with an efficiency score of between 0.344 and 1.000.

4.5 IMPROVEMENT

A fundamental assumption behind the computation of relative efficiency is that if a given firm, A, is capable of producing $Y\ (A)$ units of output using $X\ (A)$ of inputs, then other firms should also be able to do the same if they were to operate efficiently. We can set Performance Targets for inefficient libraries to enable them to reach 100 percent relative efficiency in comparison with KNUST, the most efficient. KNUST has operated in an environment similar to the others and hence using its performance as a benchmark is realistic. Input Target for inefficient Libraries will be the amount of input that will enable the library to have the same ratio of efficiency as KNUST.

Using the input from 2010 and the corresponding efficiencies:

$$Input\ Target\ =\ Actual\ Input \times Relative\ Efficiency$$

For UCC:

$$Input\ target\ =\ 54{,}046.87 \times 0.344 = 18{,}592.12$$

This means that if UCC operates using GHS 18,592.12 as input with the same output, then it will be considered as efficient as KNUST.

The difference between actual input and input target is Input Slack.

ie.

$$Imput\ slack =\ actual\ input\ -\ target\ input = 54{,}046.87 - 18{,}592.12 = GHS\ 35454.75$$

Input Slack can also be expressed as a percentage.

$$Input\ Slack\ Percentage = \frac{Input\ slack}{Actual\ Input} \times 100$$

ie. $\frac{35454.75}{54{,}046.87} \times 100 = 65.60$

Thus, if UCC has to be as efficient as KNUST, it should produce the same output using 65.60 percent less input.

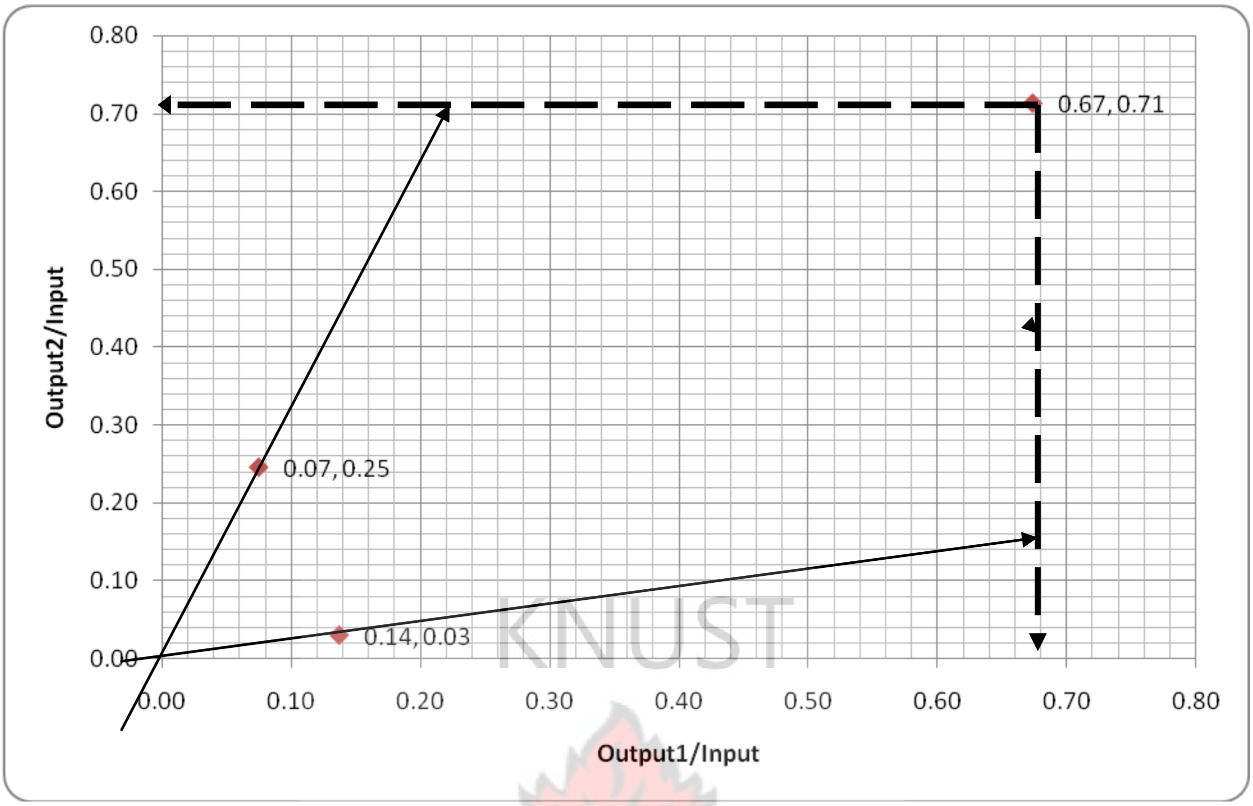
Similarly for UG its input target is GHS 203,446.43 and input slack of GHS 130,599.96, with an input slack percentage of 39.10. That is UG it should produce the same output using 39.10 percent less input.

Looking at the situation from the output point of view using the same data for 2010, to obtain a unitized frontier in this case we divide the value of printed edition expenses (PrEdEx) which is considered to be the only input of interest. The efficient frontier consist of the coordinate (0.67, 0.71) as shown in figure **4.5.1**

Table 4.5.1 Table of Output2/Input and Output1/Input

DMU	PrEdEx	RR	BB	RR/PrEdEx	BB/PrEdEx
KNUST	12771.41	8612	9113	0.67	0.71
UCC	54046.87	4020	13277	0.07	0.25
UG	334,066.392	45685	9931	0.14	0.03

Fig. 4.5.1: Graph of Output2/Input against Output1/Input



The frontier region is bounded by the axes and the frontier coordinate. The library systems in University of Cape Coast and that of University of Ghana are inefficient and their efficiency can be evaluated by the efficiency frontier point.

From figure 4.4 the efficiency of A is evaluated by

$$\frac{d(O, A)}{d(O, Q)}$$

where d(O,A) and d(O,Q) means "distance from zero to A" and "distance from zero to Q" respectively. The above ratio is refered to as a "radial measure" and can be interpreted as the ratio of two distance measures.

The Euclidean measures given by $d(O, A) = \sqrt{(0.07)^2 + (0.250)^2} = 0.26$

$$d(O, Q) = \sqrt{(0.2)^2 + (0.71)^2} = 0.73$$

$$\frac{d(O, A)}{d(O, Q)} = \frac{0.26}{0.73} = 0.36$$

which is showing the proportion of inefficiency in both output of *A*.

Because the ratio is formed relative to the Euclidean distance from the origin over the production possibility set, we will always obtain a measure between zero and unity.

We can also interpret the results for managerial (or other) uses in a relatively straightforward manner. The value of the ratio in will always have a value between zero and unity. Because we are concerned with output, however, it is easier to interpret in terms of its reciprocal

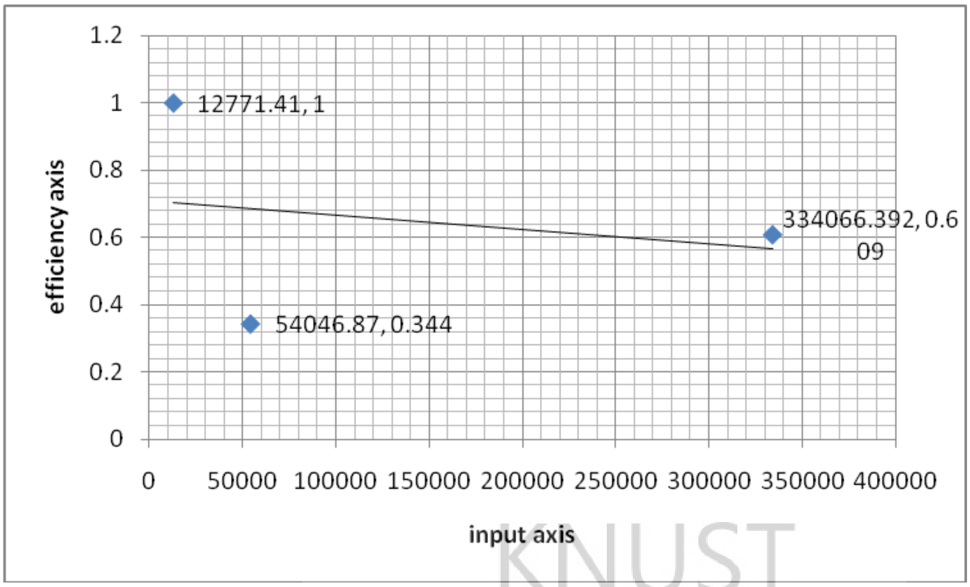
$$\frac{d(O,A)}{d(O,Q)} = \frac{0.73}{0.26} = 2.8$$

This result means that, to be efficient, UCC Library system would have had to increase both of its outputs by 2.8. To confirm that this is so we simply apply this ratio to the coordinates of A and obtain the Q(0.2, 0.7) which would bring coincidence with the coordinates of Q , the point on the efficient frontier .

Similarly the proportion of inefficiency in UG Library system is 0.21 and for it to be efficient; it would have had to increase both of its output by 4.86.

In attempt to find the correlation between the efficiency and the inputs, Pearson's product moment correlation$r$, was calculated (the steps in calculating r is shown in appendix B). It was found that$r = -0.226787817$, a weak correlation in the sample data. The minus sign tells us that it is a negative correlation, i.e. Values of efficiency tend to decrease as values of input increases and vice versa. This can also be seen by the fact that the line appears to slope downwards to the right in the figure below.

Fig 4.5.2. Graph of Regression of Efficiency against Input



The equation of the line if best fit (y on x) is $y = 0.708 - 4.28E - 07x$ (calculation shown in appendix B) which can be used to predict efficiency where x is a given input.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

Data envelopment analysis seems to be a useful tool for small data sets estimation.

The method identifies best practices for the purpose of benchmarking. The analysis provides the precise corrective figure for every output and input in order to improve the efficiency of an inefficient University Library. The Library administration might choose a new strategy based on the result of DEA, in order to operate in a more efficient mode. However, this does not mean that the result is transformed into attainable recommendation. In our case we apply Data Envelopment Analysis, using three (3) variables, which are not related to internal service quality. This analysis estimates the relative operating efficiency of University Library irrespective of quality comparisons.

This chapter summarizes the major findings and conclusions of the study. It also provides a discussion on some future research options on the subject of library system efficiency.

Two years of operating data from three (3) Universities were used in this analysis. The following is lists of the major findings and conclusions that can be drawn with respect to the efficiency of library systems and the improvement of library efficiency:

1. DEA was found to be effective and relatively easy to use for quantifying the technical efficiency of library systems.
2. To identify the improvement scale factor in each of the library systems a "radial measure" was used.
3. A correlation and regression analysis was also used to see how best a line fit into the sample data.

5.1 Future Research

This research is limited in a number of aspects due to limited availability of operating data. Future research is needed and should focus on the following areas:

- It would be valuable to conduct a survey of library systems to obtain more accurate and detailed information on inputs, outputs, environmental factors, service management factors and distinctive operating practices, etc.

- It is necessary to perform more extensive analysis of the sources of efficiency / inefficiency and influencing factors;
- Future efforts should also devoted to the development of guidelines that library systems can use to improve their service performance;
- It is important to investigate the effects of other independent variables on library efficiency, such as whether or not the employees are unionized, etc.
- The relative efficiency calculated in this study is based on the standard DEA method without placing any restrictions on the input and output weights.

Incorporation of reasonable weights restrictions into this DEA model may generate more reliable results.

- Economic efficiency of a library system is also of critical important to the future development of library services. However the technical efficiency score calculated using DEA is a "relative" score. It only represents how well or badly a firm is operated comparing with its peers, and it is not necessarily related to economic efficiency of a system. The relative efficiency analysis can be used as a basis for evaluating economical efficiency.

RFERENCES:

Akdede and Kazancoglu (2006)- Measuring Relative Efficiency of Public (state) Libraries of major cities in Turkey.

Al-Eraqi A.S. et al. (2007)- Evaluation the Location Efficiency of Arabian and African Seaports Using Data Envelopment Analysis (DEA).

Chin Huan Huan et al. (2008) -Evaluation of Municipal Solid Waste Management Projects in Metro Manila, Taiwan using Data Envelopment Analysis.

Chun-Yu Lin et al. (2007) - Product Line Selection using Data Envelopment Analysis.

Cooper, W.W., Seiford, L.M. and Tone, K. (2006) –Introduction to DEA and its uses.

Cracolici, M. F., Nijkamp, P. and Rietveld P. (2006) - Assessment of Tourist competitiveness by Analysing destination efficiency.

Goncalves and Ricardo (2006) - Measuring Management Efficiency Using Data Envelopment Analysis (DEA) Estimation for Banks in Brazil.

Mishra and Gokulananda (2009) -Suppliers Development Strategies usingData Envelopment Analysis.

Moffat and Abbas (2009)- Measuring the Efficiency of Financial Institutions in Botswana.

Mohammad S. etal. (2009) - Safety Performance of Construction.

N. K. Womer et al. (2006) -Benefit-Cost Analysis using DEA.

Reichmann (2004) -Measuring Technical efficiency of 118 randomly selected university libraries from German-speaking countries (Austria, Germany, Switzerland) and English-speaking countries (the United States, Australia and Canada) using Data Envelopment Analysis.

SALAS-VELASCO M. (University of Granada, Spain, 2006) Evaluation of the Performance of Private and Public Schools.

Townend J., (2002) - Practical Statistics for Environmental and Biological Scientists

Vitaliano (1998)- Analyzing the Performance of Public Libraries in New York (Vitaliano) and New South Wales, Australia (Wor-thington).

Xiangyu Wang (2010) -Measuring the Performance Insurance Companies using Data Envelopment Analysis.

Yang J. (2005) Quantifying the Technical Efficiency of Canadian Library Systems Using Data Envelopment Analysis.

Yongjun Shen et al. (2009) -Evaluation of Road Safety Performance using Data Envelopment Analysis.

APPENDIX A:

SIMPLEX PROGRAM FOR 2010

Max: KNUST =8612u_1 + 9113u_2;

c: 12771.4v_1 = 1;

8612u_1 + 9113u_2 $\leq$ 12771.4v_1;

$4020u_1 + 13277u_2 \leq 54046.87v_1$;

$45685u_1 + 9931u_2 \leq 334066.392v_1$;

$u_1 \geq 0$; $u_2 \geq 0$; $v_1 \geq 0$;

Constraints: 7 Variables: 3 Slack Surplus Variables: 6 Artificial: 4

Value of objective function: 1.0

u_1 = 0.0

u_1 = 1.0973334796444 64E-4

v_1 = 7.829994905915424E-5

Using Simplex

Max: UCC = $4020u_1 + 13277u_2$;

c: $54046.87v_1 = 1$;

$8612u_1 + 9113u_2 \leq 12771.4v_1$;

$4020u_1 + 13277u_2 \leq 54046.87v_1$;

$45685u_1 + 9931u_2 \leq 334066.392v_1$;

$u_1 \geq 0$; $u_2 \geq 0$; $v_1 \geq 0$;

Constraints: 7 Variables: 3 Slack Surplus Variables: 6 Artificial: 4

Value of objective function: 0.3442758417663344

u_1 = 0.0

u_2 = 2.5930243410886073E-5

v_1 = 1.8502458694887158E-5

Using Simplex

Max: UG=$45685u_1 + 99314u_2$;

c: $334066.392v_1 = 1$;

$8612u_1 + 9113u_2 \leq 12771.4v_1$;

$4020u_1 + 13277u_2 \leq 54046.87v_1$;

$45685u_1 + 9931u_2 \leq 334066.392v_1$;

$u_1 \geq 0$; $u_2 \geq 0$; $v_1 \geq 0$;

Constraints: 7 Variables: 3 Slack Surplus Variables: 6 Artificial: 4

Value of objective function: 0.6092116405144383

u_1 = 1.3335047400994602E-5

u_2 = 0.0

v_1 = 0.0

Using Simplex

SIMPLEX PROGRAM FOR 2009:

Max KNUST = $4345u_1 + 8308u_2$;

c: $36128.76v_1 = 1$;

$4345u_1 + 8308u_2 \leq 36128.76v_1$;

$4211u_1 + 14209u_2 \leq 90391.68v_1$;

$53580u_1 + 1366u_2 \leq 371184.88v_1$;

$u_1 \geq 0$; $u_2 \geq 0$; $v_1 \geq 0$;

Constraints: 7 Variables: 3 SlackSurplusVariables:6 Artificial:4

Value of objective function: 0.9999999999999999

u_1 = 0.0

u_2 = 1.2036591237361578E-4

v_1 = 2.7678778691189486E-5

Using Simplex

Max UCC= $4211u_1 + 14209u_2$;

c: $61391.68v_1$=1;

$4345u_1 + 8308u_2 \leq 36128.76v_1$;

$4211u_1 + 14209u_2 \leq 61391.68v_1$;

$53580u_1 + 1366u_2 \leq 371184.88v_1;$

$u_1 \geq 0;$ $u_2 \geq 0;$ $v_1 \geq 0;$

Constraints:7 Variables:3 SlackSurplusVariables:6 Artificial:4

Value of objective function: 1.0

u_1 = 0.0

u_2 = 7.037792948131466E-5

v_1 = 1.628885225311095E-5

Using Simplex

Max: $UG=53580u_1 + 1366u_2;$

c: $371184.88v_1 = 1;$

$4345u_1 + 8308u_2 \leq 36128.76v_1;$

$4211u_1 + 14209u_2 \leq 90391.68v_1;$

$53580u_1 + 1366u_2 \leq 371184.88v_1;$

$u_1 \geq 0;$ $u_2 \geq 0;$ $v_1 \geq 0;$

Constraints: 7 Variables: 3 SlackSurplusVariables:6 Artificial:4

Value of objective function: 0.9999999999999999

u_1 = 1.866368047779022E-5

u_2 = 0.0

v_1 = 0.0

Using Simplex

APPENDIX B (correlation and regression analysis**)**

x	12771.41	54046.87	334066.39
y	1	0.344	0.609

xy	12771.41	18592.12328	203446.43
y^2	1	0.118336	0.370881
x^2	163108913.4	2921064157	1.116E+11

$n = 3$

$\bar{y} = 0.651$ $\bar{x} = 133628.2$ $\sum y = 1.953$ $\sum x = 400884.67$

$\sum xy = 234809.966$ $(\sum y)^2 = 3.814209$ $(\sum x)^2 = 1.60709\text{E} + 11$

$\sum y^2 = 1.489217$ $\sum x^2 = 1.14685\text{E} + 11$

Pearson's product moment correlation r_{n-2} is given by:

$$\frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{(\sum x^2 - \frac{(\sum x)^2}{n})\ (\sum y^2 - \frac{(\sum y)^2}{n})}} = -0.226787817$$

Regression of y on x.

$y = \alpha + \beta x$

$$\beta = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 \sum (x)^2} = -4.28E - 07$$

$\alpha = \bar{y} - \beta\bar{x} = 7.08E - 01$

$\Rightarrow y = 7.08E - 01 - 4.28E - 07$