4-3-2009

Assessing the Performance of Water Bodies in Hillsborough County, Florida Using Data Envelopment Analysis (DEA)

Geoffrey George Fouad
University of South Florida

Assessing the Performance of Water Bodies

in Hillsborough County, Florida

Using Data Envelopment Analysis (DEA)

by

Geoffrey George Fouad

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Geography
College of Arts and Sciences
University of South Florida

Major Professor: Kamal Alsharif, Ph.D.
Hyun Kim, Ph.D.
Philip Reeder, Ph.D.

Date of Approval:
April 3, 2009

Keywords: Benchmark, Efficiency Frontier, Geographic Information Systems (GIS), Lake, Land Use

Table of Contents

List of Tables

List of Figures

Assessing the Performance of Water Bodies in Hillsborough County, Florida Using Data Envelopment Analysis (DEA)

Geoffrey Fouad

ABSTRACT

The purpose of this thesis is to describe the relationship between surface water quality and land use. Water management recommendations will be divulged based upon the interaction of lake water quality and land use. The methodology developed for this research applied Data Envelopment Analysis (DEA), a performance measurement tool, to evaluate lake water quality in relation to surrounding land use. Lake performance ratings were generated by DEA software that assessed multiple variables describing surface water nutrient loads and surrounding land use. Results from this analysis revealed a significant trend between lake water quality and land use within the study area. Lakes located within a two mile radius of more naturally preserved land area typically attained higher performance ratings than lakes located within a two mile radius of less naturally preserved land area. The spatial quantity of naturally preserved land influenced lake nutrient concentrations. Also, lake performance ratings generally declined in two mile radius delineations that contained less naturally preserved land area indicating a direct relationship between natural land area and lake performance.

Introduction

The intention of this thesis is to explore existing scientific literature discussing previously attempted environmental and water management applications of Data Envelopment Analysis (DEA), a performance measurement tool. It is also the intention of this thesis to devise water management recommendations for Hillsborough County using a DEA methodology. This methodology was an attempt to characterize the impacts of land use on lake water quality. In doing so, the applied research provides a means to develop specific water management recommendations based on localized data from Hillsborough County, Florida. DEA was implemented as a performance measurement tool to gauge the impact of surrounding land uses. The scientific research for this thesis represents an application of DEA not previously attempted in available literature. DEA has not been previously used to assess the relationship between land use and water quality.

The applied research examined the effects of multiple variables on water quality including total chlorophyll, total nitrogen, total phosphorous, and natural land area. These variables were selected based upon internet availability and relativity. In subsequent sections, a further description and justification for the selected variables will be provided. DEA was applied as a tool to examine the previously mentioned variables in the form of a cross-sectional analysis of select

Hillsborough County lakes. In doing so, the applied research identified benchmarks in water quality that resulted in optimum DEA performance measurements. After identifying maximum performance benchmarks, the reader is provided with water management recommendations based upon recreating or sustaining the optimal conditions corresponding to the said benchmarks.

Water has become an increasingly significant natural resource. Throughout history, water has been the source of human conflict and the root of civilization meltdowns. Currently, water is a strictly managed commodity with monetary and intrinsic value. The value of water has so risen that humans are constantly exploring new and improved methods for managing it (Postel 1997; Feldman 2007; Houck 2002). This effort has been constricted by steadily shrinking budgets and man power (Postel 1997; Feldman 2007). Universally, water management has been further complicated by steadily declining water quality as a result of an assortment of human activities (Reddy and Dev 2006). It has been widely discussed and agreed that the overall quality of water resources in the United States has declined in recent years due to urbanization (Reddy and Dev 2006; Wescoat and White 2003; Gleick et al. 2006).

In light of these challenges, water managers have become more reliant on remote monitoring methods that require less cost and labor (Castelletti and Soncini-Sessa 2007). Remote monitoring is powered by advancing computer technology that allows users to process large volumes of data. The selected method of processing data can unveil statistical results that sometimes influence managerial decisions. DEA is one such method that has recently been applied to water management issues for optimization purposes (Alsharif et al. 2008; Castelletti and Soncini-Sessa 2007). DEA integrates actually observed data related to environmental quality when assessing system performance. This efficiency measurement tool is supported by numerous computer software platforms which have been typically applied to economic and industrial production assessment. Recent scientific publications have discussed the application of DEA to natural resource management and more specifically water management concerns (Alsharif et al. 2008; Shafiq and Rehman 2000; Jaenicke and Lengnick 1999; Castelletti and Soncini-Sessa 2007). This application of DEA represents a relatively new and vastly unexplored management tool that could possibly become very valuable in the future. DEA is a performance assessment tool that can be and has previously been used to optimize the beneficial aspects of a given natural resource (Alsharif et al. 2008; Shafiq and Rehman 2000; Jaenicke and Lengnick 1999; Castelletti and Soncini-Sessa 2007). In doing so, DEA focuses on actually observed data that potentially impacts the performance of an environmental system.

Literature Review

Water Management Application

In a study performed by Alsharif et al. (2008), DEA was applied to the performance of water supply systems in the Palestinian territories. The methodology of this study focused on water resources in a region experiencing population increases that have contributed to diminished water resources and increasingly negative human impacts (Alsharif et al. 2008). The methodology discussed in this paper was an attempt to improve water management strategies that must cope with a limited budget. DEA was used to evaluate the efficiency of individual water supplies. This entails the use of production ratios composed of outputs over inputs (Stolp 1990). The single output included in the DEA performed for this study was total revenue generated from water distribution activities. Input variables for this study focused on investments and losses related to water distribution systems. Water losses, energy, maintenance, and salary of workers associated with Palestinian water distribution systems were all considered by the DEA. Analyses of these ratios yield interpretable efficiency measures that can be referred to while managing water supplies.

Findings of the study determined DEA to be a highly applicable tool for the management of stressed water supplies (Alsharif et al. 2008). By referring to benchmarks known as an ‘efficient frontier’, the study successfully established the relative efficiencies of individual water supplies. Given the appropriate, time-sensitive data, DEA was characterized as a valuable method for managing water resources when confronted with limited man power and funding (Alsharif et al. 2008). The stability of these water resources were also successfully assessed (Alsharif et al. 2008).

Results of the assessment discovered that productivity for water resources within the Gaza Strip were significantly lower than that of neighboring water resources in the West Bank (Alsharif et al. 2008). Alsharif et al. (2008) identified water loss as the primary input variable affecting water supply efficiencies in the region. The input variable for municipality populations had little bearing on these results (Alsharif et al. 2008). Managerial policies recommended by the study suggested that water governing entities in the Palestinian region should concentrate on limiting water losses by making the necessary repairs to water distribution systems (Alsharif et al. 2008). In relation to the content of this thesis, the research conducted by Alsharif et al. (2008) is a direct example of how DEA can be applied to a water management issue.

Environmental Assessment Application

In a study conducted by Jaenicke and Lengnick (1999), the quality of soil was examined in relation to its agricultural productivity. The applied research necessary for completing this examination employed DEA to evaluate the performance of soils located within U.S. Department of Agriculture experimental fields in Maryland. Soil performance was measured by the crop yields of these fields. The study determined the quality of soil in an economic context. During the study, crop yields were perceived as an economic product that reflected the quality of soil in which the crop was planted.

The methodology for such an undertaking applied DEA to establish efficiency benchmarks representing the best known crop production levels. In simple applications of DEA, productivity is determined by production ratios containing single output over a single input. In this study, a simple DEA application was deemed impossible due to the complexity of the relationship between soil quality and crop yields. Applications requiring multiple inputs and outputs for each production ratio abide by a mathematical method developed by Caves, Christensen, and Diewert (1982).

Input variables included in the production ratios for this study were composed of management inputs such as fertilizer application, weather conditions such as precipitation, and soil quality properties such as soil moisture (Jaenicke and Lengnick 1999). Production ratios for this study also included output variables composed of crop production in mass yield and mass yield of crop by-products (Jaenicke and Lengnick 1999).

The study performed by Jaenicke and Lengnick (1999) relied upon an Additive, or alternative, form of DEA. After evaluating the nonparametric application of DEA, Jaenicke and Lengnick (1999) conclude that the methodology used during the study is a solution to creating a universal and practical soil quality index. This conclusion was supported by the study's acceptance of economic and quantitative terms for expressing the productive efficiency of a particular type of soil. The study demonstrates that the quality of an environmental factor can be assessed based upon quantitative figures that represent economic value. Researchers that participated in this study recommend that future soil quality indices, especially those applied to agricultural systems, should incorporate DEA as a cost-effective tool for examining crop production yields in relation to biological, chemical, and physical soil parameters (Jaenicke and Lengnick 1999). In relation to the content of this thesis, the research conducted by Jaenicke and Lengnick (1999) provides a useful example of how DEA can be applied to environmental assessments.

Agricultural Application

In a study performed by Shafiq and Rehman (2000), the sources of production inefficiencies for cotton production in the Punjab province of Pakistan were identified using DEA. Information regarding the actual farmers responsible for a particular cotton field were collected and used as inputs within the production ratios of the study. This information was collected primarily as quantitative data that expressed such factors as the age of the farmer and the amount of land attended to by the farmer. Other inputs considered during the study performed by Shafiq and Rehman (2000) included nitrogen fertilizer use, phosphorous-based fertilizer use, artificial irrigation levels, and hours of field plowing activity. Inputs were also categorized even further by using descriptive variables that framed the situation in which the cotton was being produced. Examples of these input categorizations included a classification scheme for available land to grow cotton as well as a classification scheme for precipitation levels. Similar to Jaenicke and Lengnick (1999), outputs consisted of various quantitative forms of evaluating crop yields. The various quantitative outputs measured crop production by mass and monetary profits.

Shafiq and Rehman (2000) applied an Additive DEA model during their study of inefficiencies for cotton production in the Punjab province of Pakistan. This DEA methodology is also frequently referred to as a DEA alternative model (Ramanathan 2003). Researchers determined that this application of a nonparametric DEA model is an appropriate technique for identifying production inefficiencies and the specific variables contributing to diminished crop yields (Shafiq and Rehman 2000). However, the researchers pointedly remark that the interpretation of results gathered from this form of DEA should be developed in a cautious manner (Shafiq and Rehman 2000).

Agricultural management interpretations based upon an application of DEA could be misleading if the model parameters do not reflect the actual inputs or outputs of a system. The same would be true for environmental management interpretations or any other realm of study with DEA applicability. Shafiq and Rehman (2000) acknowledge the power of which input and output variables are selected for an application of DEA. If certain variables are chosen to receive the expected results from a DEA model, the researcher could quite possibly omit a relevant variable or variables that would otherwise completely alter the outcome of the model. Subsequently, interpretations based upon that model would contribute to misguided management practices. Therefore, the scientific validity of a DEA model is heavily contingent upon the input and output variables selected for the analysis. This process is subjective and determined by the given researcher's logic. The final recommendation posed by Shafiq and Rehman (2000) demanding that researchers proceed with caution should be viewed as a universal truth during attempts to develop management strategies through the use of DEA.

Land Management Application

In a study administered by Rhodes (1986), land management issues confronted by the National Park Service (NPS) were prioritized according to the performance of individual parks. The efficiency with which parks employ their associated natural resources was examined during this DEA. The decision-making units (DMUs) examined during this study consisted of individual parks managed by the NPS. This allowed Rhodes (1986) to compare the efficiency of NPS managerial operations between parks. Previous studies that assessed the performance of NPS managerial operations were only site specific typically focusing on anywhere from one to three parks (Rhodes 1986). The study performed by Rhodes (1986) was unique from previous studies because it sought to evaluate the efficiency of NPS operations by comparing numerous parks simultaneously.

The overall objective of this study was to determine how well the NPS was fulfilling the agency's mission statement. DEA was considered a suitable production assessment tool for this purpose because it is capable of evaluating multiple inputs and outputs, performance measures produced by DEA are scalar eliminating assumptions that typically restrict other forms of performance analysis, and finally, the 'technical' and 'scale efficiency' components of DEA allow it to provide interpretation based upon the relationship between park size and efficiency (Rhodes 1986). 'Technical efficiency' refers to a system that can improve performance by increasing outputs proportionately (Cooper et al. 2000). 'Scale efficiency' refers to a system that can improve performance by increasing outputs (or inputs) without considering their proportions.

At the time of the study, the mission statement of the NPS provided a generalized notion regarding how the agency should preserve historic and natural sites for public use. Therefore, multiple DEA models were devised during this study that emphasized the various elements discussed by the NPS mission statement. Variable selection and data collection for each of the DEA models were based upon a collaborative effort between the author of the study, park policy-makers, and available park information stored by the NPS and the Department of the Interior (DOI). These variables were then grouped together by the element of the Park Service's mission that the variable describes. For instance, the number of historic buildings, engineering sites, prehistoric structures, and archaeological artifacts were grouped together as output variables that fulfilled the historical preservation element of the Park Service's mission. The alternative grouping of output variables fulfilled the educational and natural resource preservation aspects of the Park Service's mission. This grouping included variable data for number of educational activities, visitors attending educational activities, attendants using trails, recreational hours devoted to the park, and overnight campers. Input variables were also separated into two different groupings. The number of permanent full-time employees, career seasonal employees, and temporary employees were categorized together as labor input variables. For variables representing capital and land inputs, the model included data for number of buildings designated for visitor use, park operational buildings, miles of trails within the park, and miles of roads within the park.

Model results from the study revealed that parks attempting to increase visitation at the expense of natural resource management typically achieve maximum efficiency by reducing the staff members and receive short visits during daylight hours (Rhodes 1986). As expected, NPS properties that exclusively preserve historic monuments or sites typically perform more efficiently when recreational outputs are minimized along with labor inputs. Exactly five parks included in the study received optimal efficiency ratings for all DEA models ran by Rhodes (1986). Upon further investigation, the author concluded that these DMUs were actually examples of parks with labor and equipment deficiencies (Rhodes 1986). This investigation was prompted by the unlikelihood of a park obtaining optimal efficiency scores for all the various operational goals of the NPS.

Overall, Rhodes (1986) acknowledges that DEA is a valid tool for assessing the performance of land management activities. The series of models developed by Rhodes (1986) aluminates the potential for DEA to be used during land management and use studies. This study provides a viable example of how DEA can be implemented to assess managerial practices and their impacts on operational efficiency. Results from the DEA can then be referred to while adjusting managerial practices for the purpose of improving operational efficiency.

Methods Other than DEA

In previously published scientific literature, the relationship between land use and water quality has been explored using techniques other than DEA. One such study conducted by Griffith et al. (2002) examined the interrelationship shared by land cover and water quality using remotely sensed indicators known as normalized difference vegetation index (NDVI) and vegetation phenological metrics (VPM). This study focused on 290 randomly selected stream sites located throughout the U.S. Central Plain states of Nebraska, Kansas, and Missouri. These sites were sampled for water quality parameters such as conductivity, turbidity, total phosphorous, nitrate-nitrite nitrogen, a biotic integrity index, and a habitat integrity index. Water quality data collected during the study was then compared to landscape data for NDVI and VPM representative of individual sample site watersheds. The study then proceeded to perform statistical testing for significant relationships between the water quality data and the remotely sensed landscape data.

The methodology developed and performed by Griffith et al. (2002) embraced a recent transfer in scope from stream runs to the entire stream catchment basin for studies regarding water quality impacts. This shift in scope has become prevalent during recent studies concerning the degradation of water resources (Sidle and Hornbeck 1991; Johnson and Gage 1997; O'Neill et al.

1997; Wiley et al. 1997). Studies adopting this new scope operate with the understanding that water resource conditions are heavily contingent upon large-scale interrelationships that span an entire catchment basin. For the study performed by Griffith et al. (2002), a large-scale scope encompassing entire catchment basins was assumed while statistically analyzing the relationship between remotely sensed vegetation data and water quality variables.

In most cases studied by Griffith et al. (2002), a statistically significant correlation between the remotely sensed vegetation data and the water quality parameters existed. The relationship between vegetative cover within a catchment basin and water quality conditions was more strongly correlated than the relationship between overall land uses within a catchment basin and water quality conditions. Therefore, vegetative cover has more of a bearing on water quality conditions than land use according to the study. This unexpected conclusion was further explained by a significant correlation between the vegetative cover data and the land use data. Based on this correlation, the study determined that land use shared a statistically significant relationship with the vegetative cover data that was previously correlated to the stream water quality data. Therefore, the land use data provides an indirect explanation of the water quality data. The study introduced a viable methodology for interrelating water quality data to land cover data without the use of DEA.

Allan (2004) offered another example of a study that did not implement a methodology based on DEA to establish a relationship between land use and aquatic conditions. During this study, the index of biotic integrity (IBI) was identified as a water quality measure that typically shared a significant correlation with the land uses of a catchment basin. Allan (2004) agreed that freshwater resources have recently been increasingly studied from a large-scale perspective that views individual catchment basins as decision-making units. The research conducted by Allan (2004) supported this recent trend and perceives the IBI method as the most effective for evaluating the relationship between water quality conditions and land use. The process of correlating IBI to land use within a catchment basin represents a viable method for investigating water quality degradation at a variety of spatial scales. Therefore, this study has identified a method for devising resource management decisions based on the relationship between water quality and land use.

Research Design

The contents of this section will present the problem statement of the thesis along with research questions, hypotheses related to the research questions, research objectives, and justification for conducting the research. Impacts on freshwater bodies of Hillsborough County will be assessed by analyzing Geographic Information System (GIS) land use layers along with selected variables composed of environmental contaminant data collected and freely distributed via the Hillsborough County online Water Atlas. In doing so, the applied and previously unperformed research portion of this thesis will attempt to establish a relationship between land use and lake performance in terms of water quality. The applied research for this thesis will answer the following problem statement: Can a notable relationship between surrounding land uses and lake water quality be established, and if so, what impact does naturally preserved land have on lake water quality? In addressing this problem statement, benchmarks, a term that in the scope of this research describes water quality conditions that optimized a lake's performance, will be identified through the application of DEA. Along with the problem statement, various other research questions will be answered.

These research questions are listed below.

1. After analyzing the various forms of scientifically acceptable data using DEA computer software, does naturally preserved land typically contribute to a water quality benchmark optimizing lake performance?
2. How can water managers operating within the boundaries of Hillsborough County reproduce the necessary conditions to achieve an optimal water quality benchmark?
3. Short of altering the current land use surrounding a given lake through land acquisition techniques, how can localized water managers improve management techniques to achieve an ecologically optimal water quality benchmark?
4. After performing the necessary analysis, will the devised methodology be easily transposable to other study areas?

Hypotheses numerically corresponding to the above research questions are provided below.

1. Lakes surrounded by a greater proportion of naturally preserved land will attain higher DEA performance ratings than those lakes surrounded by a lesser proportion of naturally preserved land.
2. Results generated from the DEA will support water management strategies focused on preserving natural land and rehabilitating impaired natural land.

3. Water management efforts based on BMPs that reduce lake nutrient deposition and artificially simulate the pollutant filtration function of naturally preserved land will also likely be supported by the results of the DEA.
4. The methodology developed for this thesis will be readily transferable to other study areas that collect and store the required datasets.

The project will attempt to reveal the effects of land use on the overall performance and quality of water bodies within Hillsborough County. The content of this thesis will focus on determining the relationship of land use to water quality by applying DEA, a performance measurement tool. As stated previously, DEA is a performance assessment application that has been historically used to evaluate economic and industrial productivity. In recent available literature, DEA has been increasingly applied to performance evaluations concerning agricultural and environmental systems (Alsharif et al. 2008; Shafiq and Rehman 2000; Jaenicke and Lengnick 1999; Malana and Malano 2006). A literature review of the most relevant journals was performed to discover recently published scientific articles discussing the results of applying DEA methodologies to environmental and agricultural systems.

Besides providing an extensive literature review of environmental and agricultural DEA applications, the content of this thesis will research and evaluate the applicability of water management techniques that enhance the performance of freshwater bodies in Hillsborough County. This evaluation will focus on land use alteration and Best Management Practices (BMPs) intended to improve freshwater quality in lakes. Both of the aforementioned techniques have recently assumed a role on the forefront of emerging comprehensive water management strategies (Castelletti and Soncini-Sessa 2007; Gleick et al. 2006; Reddy and Dev 2006). The content of this thesis also attempted to reveal instances in which DEA has directly improved water management practices. A thorough literature review based upon this topic as well as a critical review of the applied research portion of this thesis revealed the advantages and disadvantages of applying DEA to environmental performance assessments.

As discussed earlier, the content of this thesis will consist of an applied research component in which DEA will measure the performance of Hillsborough County water bodies in relation to land use. The research objectives of this applied science element have been listed in numbered format below. The research questions posed previously during this section will be answered by completing the following research objectives.

1. The applied research portion of this thesis will first identify the land uses surrounding forty-three lakes within Hillsborough County through GIS data post-processing techniques and the use of a land use classification scheme developed by The Planning Commission of Hillsborough County.
2. Three DEA models, CCR-I, BCC-I, and Additive, will then be implemented to supply a comparative analysis in which the relationship between land use and water quality is examined.

3. Through this analysis, water quality benchmarks will be established that identify optimum environmental conditions within a freshwater, inland lake of Hillsborough County.
4. With the water quality benchmarks established, the research will then focus on identifying land use alterations and BMPs that will restore or maintain environmental conditions associated with optimum water quality performance in Hillsborough County lakes.

The content of this thesis contributed research toward a relatively unexplored application of a commonly used performance measurement tool. While DEA has been widely implemented for economic and industrial performance concerns, it has been generally ignored by those participating in environmental assessments. After a thorough review of the available scientific literature, it was determined that DEA has not been previously implemented to assess the performance of lakes in relation to land use. It is the goal of this thesis to contribute to the published scientific literature regarding environmental applications of DEA. In doing so, the results of this thesis discuss and evaluate the applicability of DEA for environmental assessments. The applied research portion of this thesis is supported by an in depth review of scientific literature discussing environmental applications of DEA.

DEA Background Information

An Introduction to DEA

Prior to an in depth discussion of the applied methodology, it will be important for the reader to gain an introductory knowledge of DEA. For that purpose, this section will summarize the basic aspects of DEA as discussed by Cooper et al. (2000), Sexton (1986), Ramanathan (2003), and Thanassoulis (2001). The information provided in this section will assist the reader's understanding of the applied methodology.

According to Cooper et al. (2000), DEA received its name from mathematical terminology that describes a scatter plot depicting an output versus a relevant input. When a line shelters all of the points of a scatter plot, the line is said to 'envelop' the points of the scatter plot. This line is termed the 'efficient frontier', which can be most easily defined as a high performance benchmark. Typically, a collection of performance ratios are analyzed with DEA computer software. After which, the highest levels of performance are identified by the 'efficient frontier'. Decision making units identified as efficient will occupy a point along the 'efficient frontier' line. The 'efficient frontier' is a concept unique to DEA that separates it from other forms of statistical analyses (Sexton 1986). 'Efficient frontier' lines are typically displayed on ordinary x- and y-axis scatter plots. Figure 1 provided below is a rudimentary example of an 'efficient frontier' line represented by an x- and y-axis scatter plot derived from the Charnes, Cooper, and Rhodes DEA model (Cooper et al. 2000).

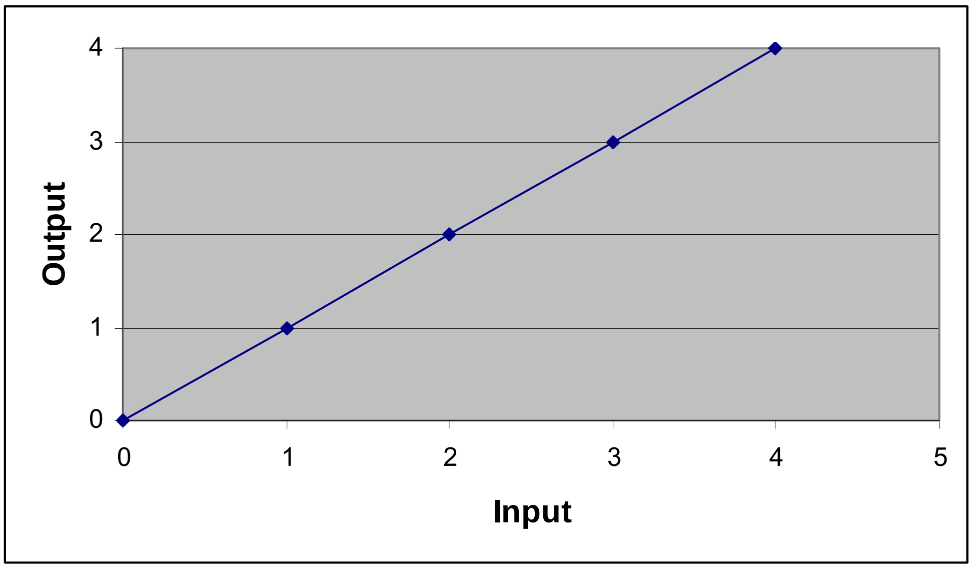


Figure 1. 'Efficient Frontier' Line Example

DEA is a mathematical platform for reviewing performance related ratios. Performance ratios are composed of a single output over a single input such as number of sales over number of employees at a store or the quantity of products generated per person employed at a factory. The concept of a performance ratio consisting of a single output over a single input stands alone as the initial idea behind DEA (Thanassoulis 2001). Performance ratios provide the foundation upon which DEA has been developed. DEA has vaulted itself to the forefront of performance measurement tools because of its capability to assimilate multiple inputs and outputs (Ramanathan 2003). As a result, this method has become widely implemented by operation managers attempting to maximize productivity (Ramanathan 2003). The modeling capability of DEA has been strongly substantiated by its ability to incorporate inputs and outputs in a multivariate fashion (Ramanathan 2003).

During its relatively short history beginning in 1978, DEA has typically been applied to issues regarding economic productivity or industrial efficiency (Cooper et al. 2000). In recently emerging scientific literature, DEA has been increasingly applied to performance-based questions related to agricultural productivity, ecosystem services, and land-use decisions (Fraser and Hone 2001; Shafiq and Rehman 2000; Malana and Malano 2006; Alsharif et al. 2008; Jaenicke and Lengnick 1999). Previous studies have referred to DEA methods when attempting to assess the efficiency of water management strategies (Alsharif et al. 2008; Tong and Chen 2002). DEA is a mathematically-based performance assessment application that incorporates production ratios (Thanassoulis 2001). These production ratios are commonly formatted with an output (or outputs) over an input (or inputs) (Thanassoulis 2001). Performance ratios evaluated by DEA measure the productivity of individual components that compile a multifaceted system. In this sense, these ratios should be considered ‘partial productivity measures’ (Cooper et al. 2000). Cooper et al. (2000) describes a collection of performance ratios as ‘partial productivity measures’ because this terminology separates DEA from other performance measurement tools that attempt to account for every output and input of a process. DEA is a ‘partial productivity measure’ because it only attempts to incorporate a select number of inputs and outputs that dictate performance. Therefore, DEA does not measure the performance of an entire system. It only measures selected inputs and outputs of a process or system. This concept has made DEA an appealing tool for environmental researchers that assess the performance of natural systems (Sexton 1986; Cooper et al. 2000; Thanassoulis 2001).

DEA does not incorporate performance ratios that assess the total productivity of a system without considering the system’s individual components such as employee efficiency or output per agricultural field (Cooper et al. 2000). In this manner, DEA is capable of identifying excesses in individual inputs as well as shortages in specific outputs. By evaluating the individual components of a system, DEA avoids assigning false or inflated values to a relatively unimportant performance factor. This is an analytically valuable aspect of DEA because a performance assessment can identify specific production areas in need of improvement. For example, a production increase might be attributed to employee labor efficiency when in actuality the individual ratios reflect that increased production was due to an increase in capital.

‘Partial productivity measures’ have frequently encountered complications or limitations because the mathematical programming for executing these evaluations has not previously been widely dispersed (Cooper et al. 2000). Advances in computer programming have made it possible to process a wide variety of variables and place quantitative values on how these variables interact (Cooper et al. 2000). The computer software currently performing DEA does not require the evaluator to assign weights and functional forms to each performance variable (Cooper et al. 2000). Computer programming improvements in DEA software have made it easier to address complicated performance-related questions. Computational progress has permitted DEA to be applied to a wider variety of managerial, social, environmental, and economic issues. Widely dispersed standardized yet flexible DEA programming frees the evaluator from the burden of creating customized software designed for a fixed evaluation and allows the evaluator to concentrate on the actual application of DEA. The body of literature related to DEA applications has also progressed and expanded in the recent past (Cooper et al. 2000), which simplifies subsequent studies that will apply DEA in a similar manner. Simplifying DEA application to a variety of fields has increased the opportunity for feedback between the analysts and those who make decisions based upon the results of the analysis (Ramanathan 2003). Increasing the feedback between analysts and those who ultimately make policy decisions has allowed more detailed and significant performance-based questions to be posed during DEA.

Performance improvements as they relate to DEA can be executed with quantitative simplicity by altering either the output (y) or the input (x). By modifying the output (y) or the input (x), the analyst can adjust underperforming points to a location within the scatter plot along the ‘efficient frontier’ line. In this quantitative manner, policy amendments need only address the quantity of outputs or inputs assigned to a specific location. By doing so, underperforming locations can improve to the best known level of performance efficiency.

From a strictly quantitative perspective, the efficiency of a point removed from the ‘efficient frontier’ can be improved by linear movement toward the ‘efficient frontier’ but not surpassing it (Ramanathan 2003). This represents the optimal placement along the ‘efficient frontier’ (Ramanathan 2003). However, efficiency improvement could be realized by altering the point’s location anywhere along the appropriate line segment of the ‘efficient frontier’ (Cooper et al. 2000). Efficiency improvement can be executed by altering either the quantity of an input or output (Cooper et al. 2000). When a decision making unit is fully efficient, it is no longer possible to improve any input or output without detracting from some other input or output (Cooper et al. 2000).

Single Output and Input Production Ratios

A simplified explanation of DEA can be accomplished by referring to an analysis composed of only a single performance ratio. This ratio places a single output over its associated input. When the ration is divided, the resulting number ranges from zero to one and expresses the productivity of a particular system component. From this point, single performance ratios from various locations are computed and can be expressed graphically with a scatter plot that places the output on the vertical line (y-axis) and the input on the horizontal line (x-axis). When the origin (0,0) and the point of each ratio are connected via a straight line, the slope of that line is compared to the slopes of the other decision making units. These slopes are compared by their rate of increase with more drastic slopes identified as more efficiently performing locations (Cooper et al. 2000). The slopes of these lines are quantitatively measured by the traditional method for calculating the slope of a line. The line connecting the origin (0,0) to the point of an individual ratio with the most drastic slope is known as the "efficient frontier" (Cooper et al. 2000). This line touches at least one of the ratio points, while the remaining ratio points are located on or below this line. Data Envelopment Analysis received its name from mathematical terminology that describes this phenomenon (Cooper et al. 2000). When a line shelters all of the points of a scatter plot underneath it, the line is said to 'envelop' the points of the scatter plot (Cooper et al. 2000).

In other performance assessment techniques, a statistical regression line can be fitted to a scatter plot. This form of statistical analysis splits the plotted data into two separate categories consisting of inferior and exemplary productivity (Cooper et al. 2000). Points above the regression line are considered exemplary, while points below the regression line are characterized as inferior. Productivity can then be assessed quantitatively by measuring the magnitude of deviation from the fitted regression line. Standard deviation is the descriptive statistic typically used to measure the distance of a sampled point from a fitted regression line (Mendenhall and Sincich 2003). By incorporating the 'efficient frontier' concept, DEA measures the deviation of points from the most productive point (Cooper et al. 2000). This represents the fundamental difference between regression analysis and DEA. The DEA for this thesis will not incorporate a regression line. It will compare lake water quality to the 'efficient frontier' line representing optimal performance. Regression analysis is focused on the central trends of a data set, while DEA avoids the use of a best-fit line and measures deviation from an actually observed line illustrating the best known performance. These two methods of statistical analysis create two very different perspectives that can greatly influence policy decisions for performance improvement. DEA identifies a line that represents the most efficient performance of a functional relationship between an output and an input. When making decisions intended to improve system performance, DEA uses an actually observed performance line as a benchmark (Cooper et al. 2000). A policy based upon the results of a DEA will attempt to improve system performance in a more dramatic fashion than a policy based upon the results of an accompanying regression analysis evaluating the same set of data.

Production Ratios with Two Inputs

Performance ratios can also reflect productivity of a system component that relies upon two inputs, which within the ratio format would be placed under one output or more practically known as the product of the two inputs. When plotting such a system component, the first input is divided by the only output to form a unitized vertical y-axis, and the second input is also divided by the only output to form a unitized horizontal x-axis (Cooper et al. 2000). From a logical perspective, systems that use fewer inputs to generate a single unit of output are considered more efficient. When two inputs are plotted in unison with a normalized output, the ‘efficient frontier’ is segmented into multiple frontiers that illustrate input tradeoffs between the two complimentary inputs (Cooper et al. 2000). The segmented line that envelops a data set including two separate inputs is located beneath the other points of the scatter plot. In this instance, the segmented frontier line represents the point at which an input cannot be increased without having a negative impact on the other input (Ramanathan 2003). By extending a vertical line down from the data point possessing the highest vertical value (y-value) and a horizontal line to the left of the data point with the highest horizontal value (x-value), a production possibility set can be established for the data set (Cooper et al. 2000). This area within the scatter plot represents all of the possible rates of production for the process being analyzed. An example of the 'efficient frontier' line for a system with two inputs and a single output is displayed in Figure 2 provided below (Cooper et al. 2000).

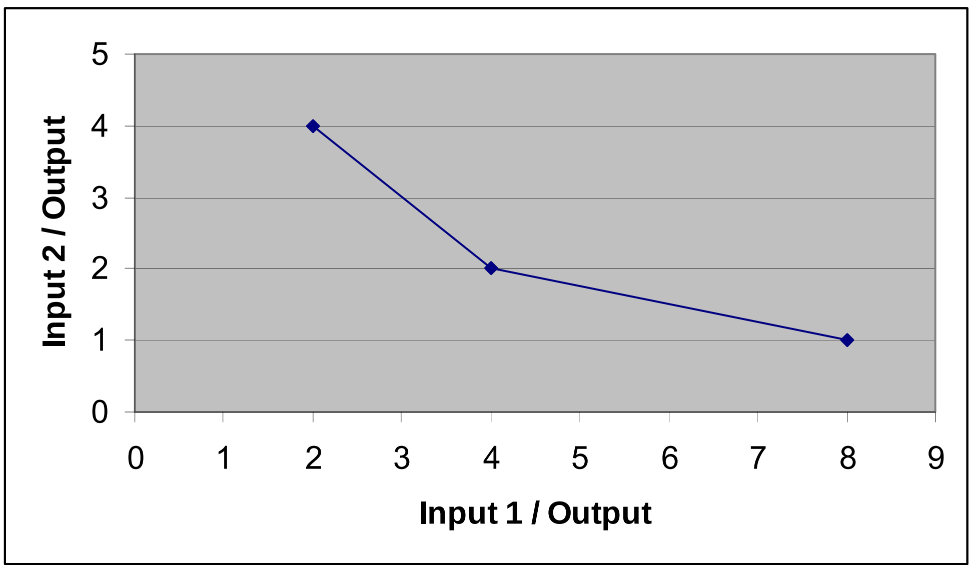


Figure 2. 'Efficient Frontier' Line for Two Inputs and One Output

The performance inefficiencies of points within the production possibility set are then measured using the 'efficient frontier' as a reference line. This task is completed by calculating two distances and dividing those distances (Cooper et al. 2000). The first distance is composed of a line from the origin (0,0) to the intersecting point along the ‘efficient frontier’ line between the origin (0,0) and the point being analyzed. The second distance is simply from the origin (0,0) to the point being analyzed. To measure the performance inefficiency of a point, the analyst must divide the first distance by the second distance. After this task has been performed, the analyst can also determine which segment of the line should be used to evaluate a point’s inefficiency. The line segment of the ‘efficient frontier’ that is intersected by the line from the origin (0,0) to the point being analyzed is the line segment of the ‘efficient frontier’ that should be used when evaluating a point’s inefficiency (Cooper et al. 2000). The two end points of the ‘efficient frontier’ line segment intersected by the line emanating from the origin (0,0) are considered the reference data set for the point being analyzed (Cooper et al. 2000). A reference data set can differ from point to point based upon the angle and distance of line segments composing the ‘efficient frontier’. Points along the ‘efficient frontier’ can also be considered more representative of the entire data set. This designation is determined by the overall distribution of points in relation to the ‘efficient frontier’ line. Points along the ‘efficient frontier’ line segments that are further removed from the majority of the points within the production possibility set likely possess unique characteristics that alter its performance from the remainder of the data set.

Production Ratios with Two Outputs

Performance ratios can also reflect productivity of a system component that relies upon two outputs, which within the ratio format would be placed over one input or practically known as the investment. Just as a ratio with two inputs, displaying a performance ratio with two outputs can also be accomplished with a scatter plot that contains a unitized x- and y-axes (Cooper et al. 2000). The first output is divided by the only input to form the x-axis, while the second output is also divided by the only input to form the y-axis. Upon plotting the data points, the 'efficient frontier' can be established by locating the outermost points and connecting them via straight line segments. The segmented 'efficient frontier' represents the outer boundary of the production possibility set, and the x- and y-axes form the innermost range of the production possibility set. Therefore, when analyzing a ratio with two outputs, one can guarantee that the line segments of the 'efficient frontier' house the other data points of the production possibility set. An example of the 'efficient frontier' line for a system with two outputs and a single input is displayed in Figure 3 provided below (Cooper et al. 2000).

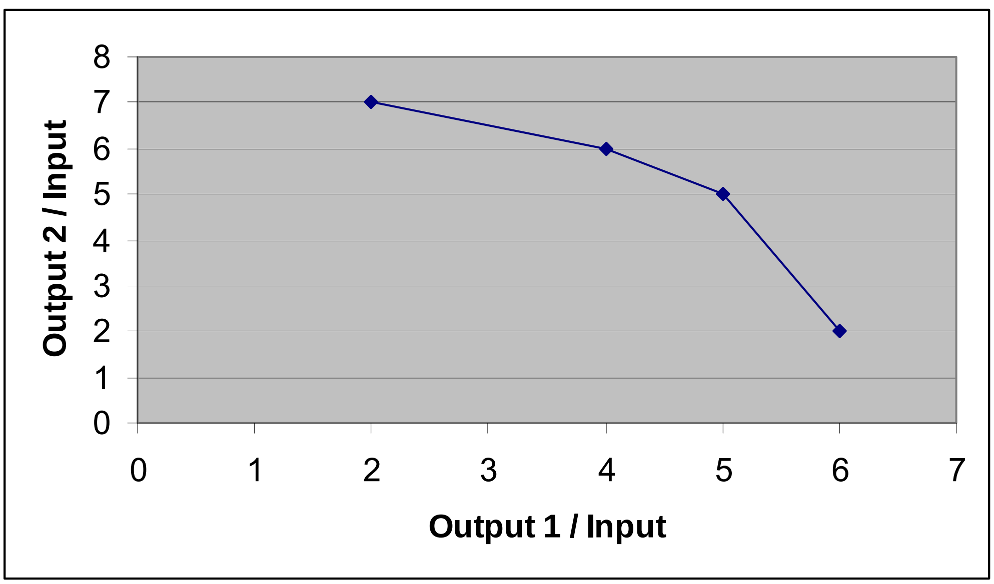


Figure 3. 'Efficient Frontier' Line for Two Outputs and One Input

Data points within the range from the x- and y-axes to the ‘efficient frontier’ are categorized as inefficient. The magnitude of such a point’s inefficiency can be calculated by referring to the ‘efficient frontier’. A line can be drawn from the plot’s origin (0,0) that intersects both the inefficient data point and one of the ‘efficient frontier’ line segments. Where this line intersects one of the ‘efficient frontier’ line segments, the evaluator can assume a point exists at this location. Once this point has been established along the ‘efficient frontier’ line segment, the distance from the origin (0,0) to the inefficient point is divided by the distance from the origin (0,0) to the point along the ‘efficient frontier’ (Cooper et al. 2000). This calculation reveals the magnitude of inadequate production efficiency for any point housed within the ‘efficient frontier’. A division calculation of this nature is commonly known as a ‘radial measure’, which in essence is a ratio composed of two distance measures (Cooper et al. 2000). Since the distance from the origin (0,0) to the inefficient point will always be shorter than the distance from the origin (0,0) to the ‘efficient frontier’ line segment, we can assume that the result of dividing these two distances will always provide a number from zero to one.

From a managerial perspective, this figure reveals information concerning two outputs of a process only when the figure’s reciprocal is interpreted (Cooper et al. 2000). Therefore, a value of three divided by four would be practically interpreted by dividing four over three. This calculation would reveal that the inefficient point would achieve optimal efficiency within the production possibility set if it were to increase outputs by a value of 1.33. When increasing outputs for a production ratio with multiple outputs, the increase should not alter the proportions of any of the ratio's outputs (Cooper et al. 2000). Inefficiencies that can be rectified by increasing outputs proportionately are referred to as 'technical inefficiencies' (Cooper et al. 2000). The term 'mix inefficiency' refers to an inefficiency that can be nullified by increasing the outputs (or inputs) without maintaining proportions (Cooper et al. 2000).

Applying Weights to Variables in DEA

Production ratios containing multiple inputs and outputs can be assessed with DEA by assigning weights or a quantitative value representing importance to the various inputs and outputs included in the ratio. Variable weights are assigned in the form of a ratio that is intended to reflect the manner in which individual outputs and inputs interact with one another (Ramanathan 2003). In DEA, ratios for a weighted variable only express how outputs and inputs interact on a separate basis. When performing an applied analysis of an actual system, values that weight specific inputs and outputs of a production process must be justified through quantitative records (Ramanathan 2003). The process of assigning weights to a variable can cast doubt on the results of a production ratio if the weight for a particular variable cannot be established through the analysis of reliable quantitative observations (Cooper et al. 2000). Another issue that clouds the analysis of production ratios including weighted variables is the level of inefficiency attributable to the assigned weights and the level of efficiency actually occurring.

DEA attempts to remove these doubt casting issues by only applying variable weights that have been directly derived from the observational data set being analyzed (Cooper et al. 2000). Weighted variables in DEA can also be chosen based upon maximizing the relative efficiencies of the entities or locations being analyzed (Cooper et al. 2000). In DEA, it is understood that an increase or improvement in an output will not negatively impact its associated input until the decision making unit has achieved optimal efficiency (Cooper et al. 2000). When a decision making unit is operating at optimal efficiency, it is not possible to augment any input or output without negatively impacting another input or output (Cooper et al. 2000). The Charnes, Cooper, and Rhodes model (CCR model) of DEA accomplishes this task by selecting variable weights that will ultimately result in the best known production levels (Sexton 1986). Using the CCR model expands the production possibility set to include all known levels of production. This model also provides an 'efficient frontier' that reflects the best known production efficiencies for a given data set. Improved performance efficiencies are accomplished through linear alterations in the ratio describing outputs over inputs.

Inefficiencies associated with entities or locations being evaluated with a multiple output and input production ratio can be labeled as a 'technical inefficiency', a 'mix inefficiency', or a 'scale inefficiency' (Cooper et al. 2000). DEA computer programs assist users by automatically identifying the form of inefficiency taking place and assigning the appropriate inefficiency values to each variable included in the ratio (Cooper et al. 2000). The benchmark reference data set that achieves the best known production efficiency is also automatically identified by DEA computer programs (Cooper et al. 2000). Another advantage of DEA computer programs is that they avoid the use of statistical assumptions based upon the trends of an entire population (Cooper et al. 2000). Avoiding these assumptions increases the accuracy of the computations performed by the program. DEA computer programs do not require the relationships between variables to be defined (Cooper et al. 2000), which can often be an arbitrary task that only weakens the results of an analysis. A final advantage of DEA is that the variables evaluated to assess performance can be expressed in different measurement units.

Summarizing DEA Production Ratios

Production ratios measure the efficiency of a process or system by dividing the output (or outputs) by the input (or inputs). In cases when multiple inputs and outputs are required to complete a production process, the variables within such a ratio are weighted according to the observed data set. These weights can be derived directly from the observational data set by employing DEA computer programs. In DEA, variable weights are not applied uniformly amongst the various outputs and inputs. Variable weights assigned by a DEA computer program reflect the best set of weights that result in the highest benchmark of efficiency. Overall, DEA is considered an advantageous method of performance analysis because it is capable of isolating sources of inefficiency and attributing a level of inefficiency to specified outputs and inputs of a production process. DEA is also a preferred measure of performance because it identifies the entities or locations with the most efficient production levels and uses these observations to form a benchmark of highest known production efficiency. This benchmark is then used as a reference to compare all other observations that fail to attain the highest known level of production efficiency. DEA requires that entities or locations being assessed include the same inputs and produced outputs. Observational data evaluated by DEA must be only composed of positive values (Sexton 1986). This limitation is also true while assigning variable weights (Sexton 1986). The selection of inputs and outputs for a designated process is determined by the evaluator performing the DEA. Inputs and outputs are commonly selected at the discretion of the performance analyst (Cooper et al. 2000). In more advanced forms of DEA, inputs and outputs are further classified as discretionary and non-discretionary (Cooper et al. 2000). Such a classification system was not used during the analytical portion of this thesis. Therefore, discretionary and non-discretionary designations will not be further defined. Categorical variables can also be applied to DEA. Variables of a categorical nature provide further differentiation between a set of production ratios and assumedly increase the level of real analytical accuracy. Application of DEA to a land cover analysis of Hillsborough County as it relates to the performance of water bodies will only include physically measurable inputs and an individual output that impact performance efficiency. At the completion of this thesis, categorical variables were not integrated into the structure of the DEA model because they were not applicable. Due to the design of the DEA model, categorical variables were not required. The subsections above provided a general description of DEA and background information useful for framing DEA in the context of the thesis research, which entails measuring the impacts of land uses on the performance of nearby water bodies.

The following paragraph will serve as a summary of the advantages and disadvantages of using DEA performance ratios during the study performed for this thesis. It will also discuss how the methodology developed during this thesis will attempt to overcome the disadvantages of applying DEA performance ratios to a cross-sectional analysis of lakes in Hillsborough County. The DEA technique is disadvantageous because it is only a partial measure of performance. This aspect of DEA poses a problem because inputs and outputs of a freshwater lake represent an intricate ecological relationship. It was not possible to include all of these inputs and outputs due to the current state of available water quality data on the Hillsborough County Water Atlas. The process of selecting relevant variables hinged upon the significance of an input or output as well as data availability. This process required a great deal of research on the Water Atlas database to view the available data for every lake within the political jurisdiction of Hillsborough County. After performing this process, it was readily apparent that the inputs of total chlorophyll, total phosphorous, and total nitrogen as well as the output of naturally preserved land surface area surrounding a particular lake represent sufficient variables for identifying a lake’s optimum performance benchmark. The other significant disadvantage of DEA is that its variable selection process can be vulnerable to scrutiny and should proceed with caution. As discussed during the examination of the study performed by Shafiq and Rehman (2000), any DEA methodology relies upon a variable selection process dependent upon the researcher's logic. If careless, the DEA user could unintentionally skew the results of the model as well as the management recommendations derived from the model's results. When using DEA, it is important to justify the variable selection process with valid arguments for each input and output chosen. In the case of this thesis, the input and output variables were selected based upon both valid and unbiased arguments. The variables were first selected based upon significance with regards to water quality and the performance of a freshwater lake. This was determined by reviewing the available literature discussing the status of lake water quality in Hillsborough County. The subsequent literature review identified specific substances most threatening to lake water quality in Hillsborough County. Following this first selection parameter, the variables were then selected based upon data availability on the Hillsborough County Water Atlas. An advantage of applying DEA performance ratios to a cross-sectional analysis of lakes is that this research is capable of identifying the levels of inputs and outputs that resulted in optimal aquatic conditions. Also, DEA enabled this research to identify the exact quantities at which a specific input or output is most beneficial to the performance of the aquatic ecosystem of a freshwater lake. Furthermore, the performance ratios used by DEA are designed to examine the most critical inputs and outputs related to an environmental system. Finally, DEA performance ratios supplied this research with the necessary evidence to suggest the most effective water management techniques for the freshwater lakes of Hillsborough County.

Methodology

The scientific research of this thesis applied DEA, a performance assessment tool, to the lakes of Hillsborough County. Because of the original nature of this research, the methodology will be provided for the first time within this paper. For the applied research of this thesis, the DEA output consisted of data measuring naturally preserved land surface area within a two mile radius of each lake selected for the study. Initially, the output variable measured natural land area within sub-basins previously determined by the Southwest Florida Water Management District (SWFWMD). This method for calculating natural land area was eliminated because it did not provide enough variability in the output data set. Lakes contained by the same sub-basin recorded the same output values. In many instances, the study lakes were located within the same sub-basin reducing the variability of the output data set. In an attempt to counteract this, it was decided to measure natural land area within a two mile radius of study lakes. This effort returned improved output data variability. A distance of two miles was selected because it reflected a size comparable to SWFWMD sub-basins. Also, a two mile radius was determined to be an appropriate size because of the sparse geographic distribution of natural land in Hillsborough County.

Several of the two mile radius delineations extend into neighboring counties surrounding Hillsborough County. Areas in which the two mile radius delineations extend beyond the Hillsborough County boundary were considered during the natural land cover selection process, however, these areas failed to yield any naturally preserved land. All naturally preserved land within two miles of a study lake was selected for the DEA regardless of county boundaries.

The thesis methodology isolates lakes according to spatially oriented polygons that represent a two mile radius surrounding each study lake. These two mile radius delineations have been automatically assigned feature identification numbers by the ‘Buffer’ tool of ArcMap. Table 1 lists each of the two mile radius delineations included in this study by feature identification numbers and provides the corresponding naturally preserved surface area. This table also displays which lake is located within each two mile radius delineation feature. The final column of the table displays the natural land use percentage within each two mile radius delineation feature.

Table 1. Two Mile Radius Delineation Feature Summary

Two Mile Radius Delineation Feature Identification Number (FID)	Surface Area (in acres)	Lake within Two Mile Radius Delineation Feature	Natural Land Use Percentage within Two Mile Radius Delineation Feature
00	8,650	Garden Lake	0.9402
01	10,277	Brant Lake	3.5855
02	10,485	Lake Hiawatha	0.9815
03	15,115	Lake Thonotosassa	0.6194
04	8,820	Flynn Lake	0.6034
05	10,446	Pretty Lake	2.2794
06	9,681	Hanna Lake	3.9308
07	9,563	Lake Josephine	2.3540
08	9,094	Echo Lake	1.2807
09	15,042	Lake Keystone	1.8600
10	9,301	Lake Armistead	1.9134
11	8,993	Lake Harvey	2.2417
12	9,213	Sunset Lake	0.8827
13	8,859	Cypress Lake	2.7560
14	8,928	Chapman Lake	0.1616
15	8,954	Lake Virginia	2.2515
16	8,602	Burrell Lake	0.6187
17	10,288	Lake Thomas	3.6078
18	9,552	Rock Lake	2.2959
19	9,818	Osceola Lake	0.8119
20	8,847	James Lake	2.7597
21	10,228	Lake Alice	1.7454
22	9,495	Lake Weeks	1.3223

Natural land area is considered an output because of the perceived notable relationship between water quality and surrounding land use. Ideally, water quality should directly relate to the amount of natural land area surrounding a particular lake. In this case, water quality is maximized by increasing amounts of natural land area. Data for the output variable was derived from the intersection of a land use shapefile layer provided by The Planning Commission of Hillsborough County, a lake polygon shapefile layer stored in the Florida Geographic Data Library and Map Server, and polygon shapefiles representing a two mile radius surrounding each of the study lakes. GIS tools supplied within the ArcMap software package were used to calculate naturally preserved land surface area positioned within a two mile radius of each study lake. The natural land surface area calculations acquired from ArcMap populated the sample data for the DEA output variable. Inputs consisted of recorded data for substances in aquatic ecosystems that typically have a negative impact on lake performance. For this particular study, total chlorophyll, total phosphorous, and total nitrogen were examined as input variables. These substances were selected as input variables because they are the three most significant indicators of impaired water quality performance in Hillsborough County lakes (Poe et al. 2005). After a certain threshold, lakes containing an excess amount of these substances experience a decline in water quality, which negatively impacts overall lake performance (Poe et al. 2005). It is expected that minimizing input concentrations and maximizing the output variable will result in higher DEA lake performance measurements.

Table 2. Summary of Study Variables

Variable	Variable Type (Input/Output)	Measurement Units	Data Source
Total Chlorophyll (EPA method 0445.0)	Input	ug/L	Hillsborough County Water Atlas
Total Nitrogen (EPA method 0351.2)	Input	ug/l	Hillsborough County Water Atlas
Total Phosphorous (EPA method 0365.1)	Input	ug/L	Hillsborough County Water Atlas
Naturally Preserved Land Area	Output	Acreage	The Planning Commision of Hillsborough County

The measurement unit used to express raw input data was not consistent with the measurement unit used to quantify the raw output data. Input data for total chlorophyll, total nitrogen, and total phosphorous was expressed as micrograms per liter, while output data for natural land use area was quantified by acres. This inconsistency in measurement units is acceptable within the mathematical framework of DEA (Cooper et al. 2000). Measurement units used to express raw input and output data entered into a DEA are not required to be equivalent (Cooper et al. 2000; Ramanathan 2003; Thanassoulis 2001).

The ultimate goal of the applied research portion of this thesis is to link land use activity to aquatic conditions in a lake. The chosen methodology will accomplish this goal by relating the concentrations of three critical pollutants to the spatial extent of naturally preserved land surrounding a lake. By doing so, the research will unveil trends linking surface water quality to surrounding land uses. The DEA developed for this thesis consisted of input variables that should be minimized and an output variable which should be maximized to increase lake performance. Input variables assessed by this study consisted of significant pollutants that diminish lake performance as their concentrations increase. The output variable included in this DEA represents a positive influence on lake performance that should be maximized. Naturally preserved land area was selected as the output variable because this land use type maximizes water quality (Tong and Chen 2002; Castelletti and Soncini-Sessa 2007; Osborne and Wiley 1988; Lee 2002), which subsequently improves lake performance.

In the state of Florida, a naturally preserved land use category has been previously identified by the Florida Department of Transportation in a government document entitled *Florida Land Use, Cover and Forms Classification System* (1999). This classification system was referred to while spatially analyzing the land use shapefile provided by The Planning Commission of Hillsborough County. The naturally preserved land use category was selected as the output variable because it enhances lake performance as it is maximized. Multiple studies and texts have corroborated the fact that water quality typically improves with the increase of natural land uses (Xian et al. 2007; Wang 2001; Tong and Chen 2002; Gleick et al. 2006; Castelletti and Soncini-Sessa 2007; Lenat and Crawford 1994; Stauffer 1991).

Output variable data was gathered through a GIS data processing technique. Naturally preserved land use area was calculated within a polygon shapefile representing a two mile radius surrounding each of the study lakes. Initially, ArcMap was populated with three shapefiles depicting the land uses, lakes, and a two mile radius surrounding each of the study lakes. The shapefile depicting lakes was then redefined to only include those lakes selected for the study. Then, a tool by the name of ‘Clip’ was used to select only the land use polygons which fell directly within the shapefile depicting a two mile radius surrounding each of the study lakes. After redefining the land use layer, it was then feasible to select only the naturally preserved land use polygons which are spatially located within a two mile radius of a study lake and create a new layer from this selection. Finally, the surface area of naturally preserved land surrounding each lake was individually calculated. This task was accomplished using a tool known as ‘Calculate Geometry’ located in the attribute table of the most recently generated layer depicting naturally preserved land use within a two mile radius of a study lake.

Data sources that supplied the necessary information for completing this methodology are publicly accessible via the internet. Quantitative records for substances that impact the water quality of a lake populated the input variable data set. Three of the most significant inputs related to water quality in Hillsborough County populated the input variable data set. Total chlorophyll, total phosphorous, and total nitrogen represent three of the most significant aquatic pollutants currently being deposited in Hillsborough County freshwater lakes (Poe et al. 2005). The concentrations for these water pollutants have been recorded by a variety of partner agencies on a consistent basis for multiple years dependent upon the lake in question. These records were retrieved from the Water Atlas hosted and maintained by the City of Tampa and Hillsborough

County governments (http://www.hillsborough.wateratlas.usf.edu/). The surface area for naturally preserved land uses within a two mile radius of each study lake populated the output variable data set. For the purpose of this study, it was determined that naturally preserved land is the most significant land use influencing the performance of lakes in Hillsborough County. This determination was made because the natural lands surrounding lakes along with the water quality within these lakes has steadily diminished in recent decades (Poe et al. 2005). From this observation, it appears as though the spatial extent of natural lands surrounding lakes has a direct correlation with lake performance. Therefore, natural land area will be examined by the output variable of this study.

The data set for the input variables consisted of measurements from 2006 through 2008. Only data collected during the spring and summer months from March to September were included in the input variable data set. This time frame was established by a survey of the available data in the Water Atlas website. GIS land cover layers for Hillsborough County are available on publicly accessible websites hosted by a variety of governmental agencies such as Hillsborough County Planning Department, the Environmental Protection Commission of Hillsborough County, and Southwest Florida Water Management District. Certain websites designated as GIS data clearinghouses may also provide pertinent land use layers for Hillsborough County. For this thesis, one specific Hillsborough County land use layer was used. This layer has been developed, disseminated, and updated on a quarterly basis by The Planning Commission of Hillsborough County. After examining the attribute table of this land use data layer, it was determined that this shapefile contained the necessary information to calculate the areal extent of naturally preserved land uses within a two mile radius of each study lake. The polygon shapefile depicting the lakes of Hillsborough County was recovered from the Florida Geographic Data Library and Map Server, a GIS data clearinghouse for the state of Florida. And, the polygon shapefile depicting a two mile radius surrounding each of the study lakes was generated using a tool supported by the ArcMap software package known as ‘Buffer’.

The data collected from this assortment of websites was entered into two specialized computer programs that execute DEAs. These programs are named ‘DEA solver’ and ‘DEAlytics’. ‘DEA solver’ produced the results for the CCR-I and BCC-I models, while ‘DEAlytics’ produced the results for the Additive model. Three different DEA models were used to process the collected data.

The Charnes, Cooper, Rhodes (CCR) model is capable of measuring ‘technical inefficiency’ and ‘mix inefficiency’ (Cooper et al. 2000). ‘Technical inefficiency’ is eliminated without altering the proportions of system inputs and outputs, while ‘mix inefficiency’ is removed by adjusting the proportion of system inputs and outputs (Cooper et al. 2000). Multiple versions of the CCR model are used to determine both of these forms of inefficiency. For this particular analysis, the CCR-Input (CCR-I) model was applied to assess water quality inefficiencies associated with Hillsborough County lakes. The CCR-I model is designed for analyses in which the input variables are minimized and the output variables do not require any mathematical augmentations. It was determined that the input oriented version of the CCR model was the most suitable for this analysis because lake water quality optimization depends upon the minimization of the selected input variables. The CCR-I model is expressed by the following set of mathematical equations:

$$\begin{aligned}
&\max && \omega = es^{-} + es^{+} \\
&\text{subject to} && s^{-} = \theta^{*} x_o - X\lambda \\
& && s^{+} = Y\lambda - y_o \\
& && \lambda \geq 0,\ s^{-} \geq 0,\ s^{+} \geq 0 \text{ (Cooper et al. 2000),}
\end{aligned}$$

where e is equal to a vector of ones so that $es^{-} = \sum s^{-}$ and $es^{+} = \sum s^{+}$, θ^{*} is equal to the optimal objective value, s^{-} represents input excesses, s^{+} represents output shortages, x_o represents the input vector, y_o represents the output level, X and Y are the matrices of the inputs and outputs, ω represents performance maximization, and λ represents a measurement known as slack.

The BCC model is also capable of measuring both ‘technical’ and ‘mix inefficiency’ (Cooper et al. 2000). Multiple versions of the BCC model have been devised to determine both of these forms of inefficiency. For this particular analysis, the BCC-Input (BCC-I) model was applied to assess water quality inefficiencies associated with Hillsborough County lakes. Just like its other input oriented counterpart, the BCC-I model is designed for analyses in which the input variables are minimized and the output variables are maintained at actually observed levels. The input oriented version of the BCC model was selected for this analysis because lake water quality optimization depends significantly upon minimizing input variable concentrations. The BCC-I model is expressed by the following set of mathematical equations:

$$\min z_o = \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right)$$

subject to

$$\theta x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_i + s_i^- \qquad i = 1,\ldots,m$$

$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_i - s_r^+ \qquad r = 1,\ldots,s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$0 \leq \lambda_j,\ s_i^-,\ s_r^+ \ \ \forall i, j, r.$$ (Alsharif et al. 2008),

where x_{ij} and y_{rj} are inputs and outputs, respectively, s_i represents input excesses, s_r represents output shortages, z_o represents performance optimization, θ is equal to the optimal objective value, ε is equal to the sum of input and output deficiencies, and λ represents the slack measurement.

The Additive model is capable of distinguishing between efficient and inefficient DMUs, however, it differs from the previously discussed DEA models because it has no means of measuring inefficiency (Cooper et al. 2000). Instead, the Additive model specializes in directly identifying input excesses and output deficiencies through a measurement known as a ‘stability value’ (Cooper et al. 2000). These values quantitatively express the level of efficiency or inefficiency achieved by a particular decision making unit. Stability values for inefficient units are expressed as negative numbers, while stability values for efficient units are expressed as positive numbers. Increasingly negative numbers indicate higher levels of inefficiency, and increasingly positive numbers indicate higher levels of efficiency. The Additive model should be applied to performance assessments in which the input variables are minimized and the output variables are maximized. For the performance assessment conducted during this thesis, lake water quality should theoretically be optimized by minimizing inputs and maximizing outputs. Therefore, it is suitable to apply the Additive model during this performance assessment of lake water quality in Hillsborough County. The Additive model is expressed by the following set of mathematical equations:

$$\min \quad (-e^{T}s^{+} - e^{T}s^{-})$$

$$\text{subject to} \quad Y\lambda - s^{+} = Y_j$$

$$X\lambda + s^{-} = X_j$$

$$e^{T}\lambda = 1$$

$\lambda, s^{+}, s^{-} \geq 0$ (Feroz et al. 2001),

where X and Y are the matrices of the inputs and outputs, respectively, s^{-} and s^{+} are excesses in inputs and insufficiencies in outputs, respectively, e is equal to a vector of ones so that $es^{-} = \sum s^{-}$ and $es^{+} = \sum s^{+}$, and λ represents the slack measurement.

According to the framework of DEA, each lake represents a Decision-Making Unit, or DMU. In this study, the DMUs influence the 'efficient frontier' line depicting optimum lake water quality conditions. DMU selection was based on a sampling design that considered geographic location and the availability of water quality data. Initially, every lake within Hillsborough County was considered a potential DMU. This sample was then diminished by the availability of relevant water quality data on the Hillsborough County Water Atlas. Only lakes with water quality data for total chlorophyll, total nitrogen, and total phosphorous concentrations during the spring and summer months of 2006 through 2008 were selected as DMUs. Forty-three lakes within Hillsborough County satisfied this selection criterion. Finally, this sample was further diminished by the lack of any naturally preserved land use area in twenty of the two mile radius delineations. Lakes spatially contained by a two mile radius with a recorded natural land use area of zero were excluded from the sample. These lakes were excluded because they would automatically render a performance rating of zero due to the mathematical framework of DEA. After considering all of the above selection criteria, twenty-three lakes were selected as DMUs for the DEA conducted during this thesis. A list of the twenty lakes eliminated during this selection process has been provided in Table 3.

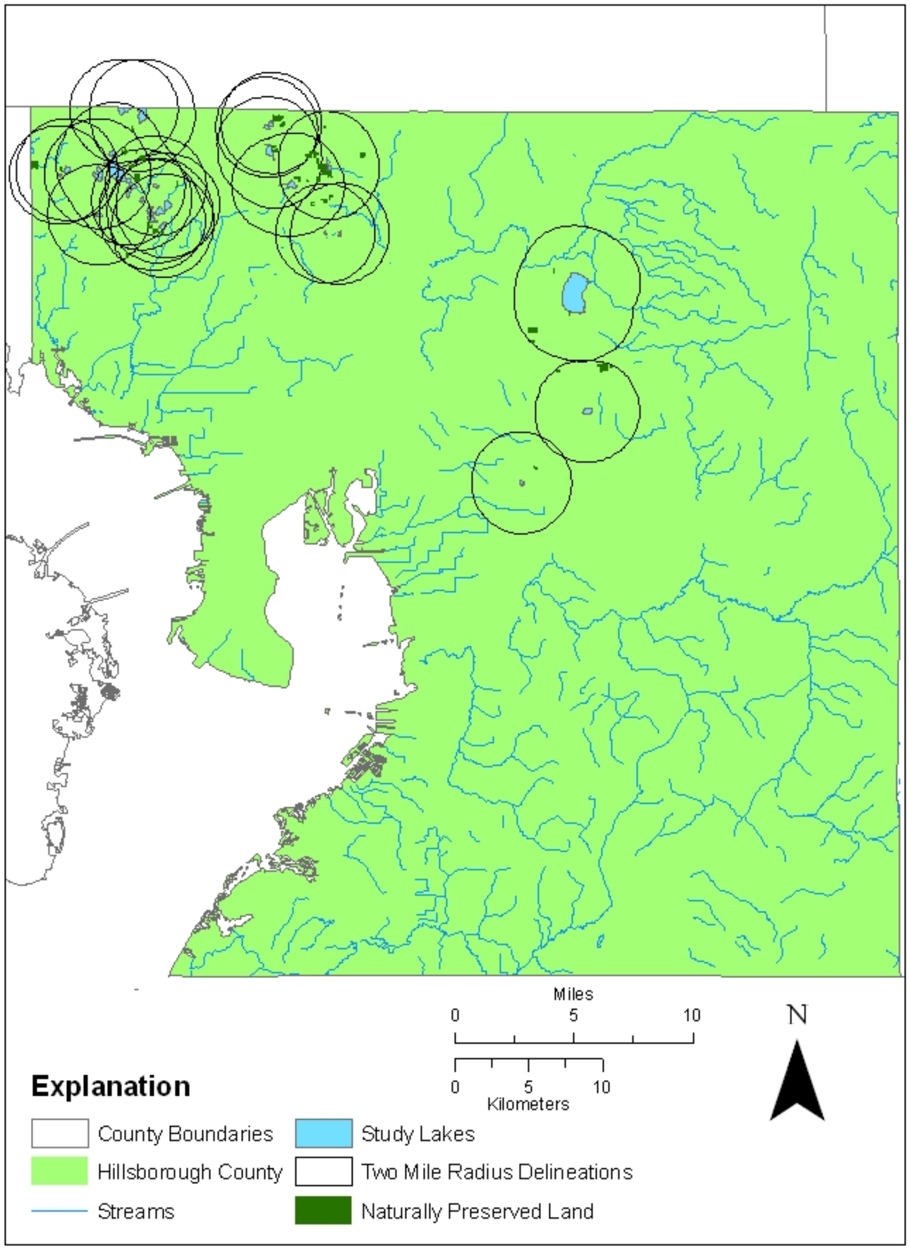
Table 3. Lakes Eliminated Due to Lack of Natural Land Use Area

Lake Name			
Boat Lake	Dorset Lake	Lipsey, Lake	Reinheimer Lake
Carroll, Lake	Eckles, Lake	Magdalene, Lake	Round Lake
Cedar Lake	Halfmoon Lake	Mango Lake	Saddleback Lake
Cooper Lake	Hobbs, Lake	Noreast Lake	Starvation Lake
Crenshaw, Lake	Leclare, Lake	North Crystal Lake	Wimauma, Lake

The final step of the applied research methodology for this thesis was to enter the data into a DEA computer software and interpret the statistical results that emerge from the analysis. Statistical results from the DEA computer software were in the form of efficiency ratings for individual lakes. Careful interpretations of these efficiency ratings allowed the researcher to recommend adjustments related to the inputs and outputs of a specific lake. The DEA software selected the lakes with the highest efficiency ratings based upon the data sets for the input and output variables. These lakes represented the 'efficient frontier'. Lakes with the most desirable water quality data were included along the 'efficient frontier'. The 'efficient frontier' served as an efficiency benchmark for all other lakes not achieving similar levels of performance. Lake inefficiency levels were determined by referring to the efficiency benchmark or 'efficient frontier'. Computer software that specializes in DEA also provided statistical measures that attribute specific amounts of inefficiency to individual input and output variables. DEA not only evaluated the total magnitude of inefficiency related to a particular lake, but it also statistically assigned numerical values describing the level of inefficiency pertaining to a particular variable. After the DEA computer software has presented this information, it was the responsibility of the researcher to properly interpret the results and formulate the appropriate water management recommendations for Hillsborough County lakes.

Study Area

As stated previously, the study area consisted of Hillsborough County, Florida. More specifically, the thesis will focus its applied research on twenty-three of the freshwater lakes located within the Hillsborough County political boundary. Figure 4 is a map of the study area that depicts the spatial distribution of lakes selected for the study along with their corresponding two mile radius delineations. The distribution of naturally preserved land located within two miles of a study lake is also displayed by the figure provided below.



Map Projection: NAD 1983 UTM Zone 17N

Figure 4. Distribution of Study Lakes, Two Mile Radii, and Natural Land

The study area for this thesis is a highly urbanized county in west-central Florida with a humid, subtropical climate characterized by a pronounced wet season from June to September. Hillsborough County is populated by approximately 1.2 million people according to the 2006 estimate provided by the United States Census Bureau. The study area occupies 1,076 square miles according to the GIS shapefile provided by The Planning Commission of Hillsborough County. This corresponds to a population density of approximately 1,115 people per square mile.

The land use shapefile examined during this thesis was also provided by The Planning Commission of Hillsborough County. According to this shapefile, Hillsborough County contains 20 specific land uses consistent with Florida's Department of Transportation land categorization scheme outlined by *Florida Land Use, Cover and Forms Classification System* (1999). The land uses included in this shapefile are listed with their corresponding surface area and percent occupying Hillsborough County in Table 4 provided below.

Table 4. Hillsborough County Land Use Summary

Land Use Category	Surface Area (in acres)	Percent of Hillsborough County (%)
Agricultural	157,079	22.80
Educational	6,011	0.8725
Heavy Commercial	2,481	0.3601
Heavy Industrial	10,466	1.519
Light Commercial	13,756	1.997
Light Industrial	8,143	1.182
Mining	26,007	3.775
Mobile Home Park	5,661	0.8217
Multi-Family	30,273	4.394
Natural	8,787	1.275
Not Classified	50,783	7.371
Public/Institutions	135,182	19.62
Public Communications/Utiliites	4,617	0.6702
Recreational/Open Space	7,675	1.114
Right of Way/Roads	1,425	0.2068
Single Family/Mobile Home	126,641	18.38
Two Family	1,000	0.1451
Unknown	227	0.03295
Vacant	58,783	8.532
Water	33,937	4.926

Historically, the predominant land uses in Hillsborough County were agriculturally related (Poe et al. 2005). However, recent trends in development over the previous two decades have transformed Hillsborough County into a metropolitan area predominantly containing urban, built-up land uses (Poe et al. 2005). It has been widely publicized that formerly agricultural lands have been converted into urbanized or industrial land uses within Hillsborough County over the past two decades (Poe et al. 2005).

The following paragraphs describe the environmental monitoring program and status of the freshwater lakes in the study area. Surface water issues have been highlighted to provide readers with a general understanding of lake conditions in Hillsborough County. Water monitoring is a necessary component of any large-scale water management effort. Currently in the Tampa Bay area, a host of agencies are responsible for monitoring the quality of surface water (Poe et al. 2005). Agencies involved in this effort include but are not limited to the Southwest Florida Water Management District (SWFWMD), Hillsborough County Environmental Protection Commission (EPC), and Florida Department of Environmental Protection (Poe et al. 2005). These agencies actively monitor the quality of surface water to assess resource performance and contribute to managerial decisions.

According to the Baywide Environmental Monitoring Report (2005), recent monitoring efforts in the Tampa Bay area, specifically in the district known as Old Tampa Bay or Hillsborough County, have revealed negative trends in water quality and clarity. Factors associated with this trend toward poorer water quality will be identified during an ongoing research project. Freshwater tributaries in the Tampa Bay Area are experiencing increasing levels of hypoxia, sediment contamination, and eutrophication. As a result, the diversity and resilience of benthic communities within negatively impacted tributaries has declined. Contaminants were detected at toxic levels in many freshwater resources. This information has prompted research to identify point and non-point sources related to the elevated levels of specific water contaminants.

The Baywide Environmental Monitoring Report (2005) identifies increased urbanization rates as the cause for the heightened levels of degradation observed in many of the freshwater bodies in the Bay area. Urbanized areas replace agricultural and forested land, which function as a natural filter for aquatic environments (Tong and Chen 2002; Castelletti and Soncini-Sessa 2007; Osborne and Wiley 1988; Lee 2002). Therefore, surface area losses in these land uses have predictably resulted in increasingly impaired water resources located within the Bay area (Poe et al. 2005). The synopsis provided above regarding the status of freshwater resources in the Tampa Bay area serves as further justification for completing the applied research portion of this thesis. The selected study area, Hillsborough County, is witnessing rapid declines in surface water quality that can only be remedied with rapid assessment techniques that contribute to adaptive management strategies.

Data for twenty-three lakes within Hillsborough County was included for the DEA conducted during this thesis. Each of these lakes was composed of freshwater. Areas surrounding these lakes were primarily composed of urbanized uses with a limited amount of naturally preserved land. Table 5 provided below contains hydrologic information corresponding to each of the lakes chosen for this study. This table also contains a column describing the overall condition of each lake in relation to its designated uses. The information for this table was retrieved from the Hillsborough County Water Atlas.

Table 5. Hydrologic Summary of Study Lakes

Lake Name	Surface Area (in acres)	Mean Depth (in feet)	Max Depth (in feet)	Approximate Volume (in gallons)	Lake Condition Category
Alice	92	9	25	248,817,000	Fully supports designated use
Armistead	34	9	28	91,865,734	Fully supports designated use
Brant	55	6	16	101,616,511	Fully supports designated use
Burrell	22	NA	NA	NA	Fully supports designated use
Chapman	42	5	11	7,356,604	Fully supports designated use
Cypress	16	12	27	59,445,932	Fully supports designated use
Echo	24	9	16	73,643,500	Fully supports designated use
Flynn	12	NA	NA	NA	Fully supports designated use
Garden	8	6	21	18,937,700	Does not support designated use
Hanna	34	5	15	52,854,890	Partially supports designated use

Table 5 (continued). Hydrologic Summary of Study Lakes

Harvey	21	10	28	73,083,950	Partially supports designated use
Hiawatha	135	11	24	494,966,000	Fully supports designated use
James	15	7	15	39,876,500	Fully supports designated use
Josephine	50	7	24	111,487,453	Fully supports designated use
Keystone	431	11	24	1,509,570,177	Fully supports designated use
Osceola	60	6	16	22,649,596	Fully supports designated use
Pretty	81	11	27	282,248,369	Fully supports designated use
Rock	53	7	21	113,864,497	Partially supports designated use
Sunset	33	8	21	93,028,200	Fully supports designated use
Thomas	60	13	27	258,565,350	Fully supports designated use
Thonotosassa	849	8	18	NA	Does not support designated use
Virginia	19	7	24	50,487,800	Partially supports designated use
Weeks	47	7	6	66,978,494	Does not support designated use

Results

The DEA for this thesis examined three input variables and a single output variable. Input variables consisted of water quality data retrieved from the Hillsborough County Water Atlas. Total chlorophyll, total nitrogen, and total phosphorous concentrations were designated as inputs during the DEA. These variables quantitatively described individual lake nutrient loading, which has been previously documented as the most severe threat to lake water quality in the Tampa Bay area (Poe et al. 2005). The single output variable consisted of acreage measurements for naturally preserved land use within a two mile radius of each study lake. The output variable was selected based upon the previously documented relationship between natural land use area and lake water quality established in the Baywide Environmental Monitoring Report of 2005. Natural land use area was selected as the output variable because surface water quality in the Tampa Bay area has historically become degraded during the same time periods in which natural land cover is more rapidly removed (Poe et al. 2005). Natural land use area and distribution have been identified as the land coverage variables that most significantly influence the status of lake water quality in the Tampa Bay area (Poe et al. 2005). Natural land area was selected as the output because it directly reflects the status of lake water quality. It is typically assumed that when natural land use area increases the quality of lake water improves.

The quantitative figures for this variable were derived from the spatial intersection of three GIS shapefiles depicting land use, lakes, and polygons representing a two mile radius surrounding each study lake. Surface analysis tools provided by the ArcInfo software package were used to calculate the surface area of naturally preserved land within a two mile radius of each study lake.

The raw data for the input and output variables is displayed in Table 6. This data was entered into two different computer programs that execute DEAs, 'DEA Solver' and 'DEAlytics'. After entering the raw data into these programs, water quality performance ratings for each study lake were generated. During this analysis, the performance rating is synonymous with the overall water quality for each study lake. Lake performance ratings are numerically expressed with values from zero to one. A performance rating of one typically indicates that a lake is operating at optimum water quality conditions, while a performance rating less than one indicates that a lake is operating at less than optimum water quality conditions. In some rare instances, a performance rating of one is not indicative of optimum performance. These instances are revealed through other performance measurements provided by 'DEA Solver' and 'DEAlytics' such as 'slack' and 'projection' ratings. In order to be considered an optimally performing DMU, both the 'slack' and 'projection' measurements must be equal to zero, and the performance rating must be equal to one.

Table 6. Raw Input and Output Variable Data

Lake Name	(I) Total Chlorophyll (in ug/L)	(I) Total Phosphorous (in ug/L)	(I) Total Nitrogen (in ug/L)	(O) Natural Land Use Area (in acres)
Alice, Lake	2.6	19.0	363.0	178.5218
Armistead, Lake	4.0	17.5	656.7	177.9678
Brant, Lake	12.5	21.7	695.0	368.4807
Burrell Lake	10.0	48.0	906.7	53.2218
Chapman Lake	5.2	23.0	1,004.0	14.4276
Cypress Lake	1.5	7.0	375.0	244.1508
Echo Lake	4.5	16.3	656.7	116.4646
Flynn Lake	13.7	15.7	1,030.0	53.2218
Garden Lake	66.7	48.7	2,013.3	81.3264
Hanna Lake	15.0	44.0	1,496.0	380.5429
Harvey, Lake	38.5	27.3	1,330.0	201.5969
Hiawatha, Lake	14.0	16.5	615.0	102.9058
James, Lake	5.5	12.5	670.0	244.1508
Josephine Lake	34.7	23.3	965.0	225.1115
Keystone Lake	3.0	10.5	426.7	279.7858
Osceola, Lake	4.0	11.3	696.7	79.7158
Pretty Lake	10.0	14.0	695.0	238.1067
Rock Lake	29.3	33.3	1,176.7	219.3028
Sunset Lake	1.5	11.5	430.0	81.3264
Thomas Lake	42.0	20.3	713.3	371.1753
Thonotosassa, Lake	99.7	204.0	2,356.7	93.6222
Virginia, Lake	32.0	36.0	1,423.3	201.5969
Weeks, Lake	152.4	264.0	2,699.0	125.5515

The reader may notice that certain pairs of lakes are located within a two square mile radius of the same amount of naturally preserved land. This is a result of the spatial position of several of the lakes chosen for the study. Lakes that were surrounded by the same amount of natural land area were closely located next to each other. Four pairs of lakes were surrounded by the same amount of natural land area. These lakes were bounded by two mile radius delineations that contained the same natural land area due to comparable spatial orientations. Cypress Lake and James Lake produced two mile radius delineations that contained the same natural land area at 244.1508 acres. Lake Harvey and Lake Virginia were surrounded by the same amount of natural land area at 201.5969 acres within a two mile radius of both lakes. Garden Lake and Sunset Lake were also surrounded by the same amount of natural land area at 81.3264 acres. Finally, Burrell Lake and Flynn Lake produced two mile radius delineations that contained the same natural land area at 53.2218 acres. Each of these pairs of lakes contained the same output variable data. Therefore, these lakes will prove useful in further examination focused on isolating the influence of the input variables on lake performance. Lakes that share the same output variable data are strictly influenced by the input variable data. In instances when a lake shares the same output variable data, it is possible to solely examine the impacts of the input variables on lake performance. Within each pair of lakes sharing the same output variable data, it is expected that the lake containing lower input variable concentrations will perform at a higher level. This subject will be discussed further in subsequent model results sections.

CCR-I Model Results

Lake performance ratings derived during this analysis describe the relationship between select water quality variables and surrounding land use. The CCR-I model returned an average performance rating of 0.3822 for the entire set of lakes. This value indicates that the entire set of lakes operated at only 38.22% of optimal performance efficiency. The standard deviation for the entire set of lake performance ratings equaled 0.2968. This value indicates a relatively wide distribution of performance ratings for the entire set of lakes. The lake water quality performance ratings derived from the CCR-I model have been provided in Table 7. As expected, these ratings increase for lakes surrounded by greater amounts of naturally preserved land. Performance ratings also predictably increase for those lakes containing lower concentrations of the selected input variables. The CCR-I model determined that only two lakes were performing at optimal efficiency.

Table 7. CCR-I Lake Performance Ratings and Rank

Lake Name	Performance Rating	Lake Performance Rank
Cypress Lake	1.0000	1
Keystone Lake	1.0000	1
Brant, Lake	0.8086	3
Thomas Lake	0.7936	4
Alice, Lake	0.7500	5
James, Lake	0.5600	6
Pretty Lake	0.5253	7
Armistead, Lake	0.4142	8
Hanna Lake	0.3879	9
Josephine Lake	0.3560	10
Sunset Lake	0.3331	11
Rock Lake	0.2842	12
Echo Lake	0.2706	13
Hiawatha, Lake	0.2552	14
Harvey, Lake	0.2323	15
Virginia, Lake	0.2160	16
Osceola, Lake	0.2023	17
Flynn Lake	0.0972	18
Burrell Lake	0.0895	19
Weeks, Lake	0.0709	20
Garden Lake	0.0616	21
Thonotosassa, Lake	0.0606	22
Chapman Lake	0.0220	23

The CCR-I performance ratings generated during this study follow general trends that describe the relationship between lake water quality and land use. General trends in the data are graphically represented by Figures 5, 6, and 7 for total chlorophyll, total phosphorous, and total nitrogen, respectively. These scatter plots represent the strength of relationship between a respective input variable and lake performance. The statistical lines of best-fit are provided for each of these scatter plots. The relationship between each water quality parameter and lake performance is depicted by the statistical lines of best-fit. The best-fit lines support the trend that lakes containing lower input variable concentrations typically obtained higher performance ratings. After observing this trend, it can be stated that lake performance shared an indirect relationship with the input variables. Accordingly, lake performance optimization is achieved by minimizing inputs. This outcome supports the previously mentioned expectations that input minimization would result in optimum lake performance.

When viewing these graphs, the reader should keep in mind that the calculated R^2 values provided in each scatter plot are only a partial measure of the model's effectiveness. The R^2 values in each scatter plot quantify the strength of relationship between the plotted water quality parameter and lake performance. The R^2 values provided in Figures 5, 6, and 7 should not be considered a measure of how effective the entire model is at predicting lake performance.

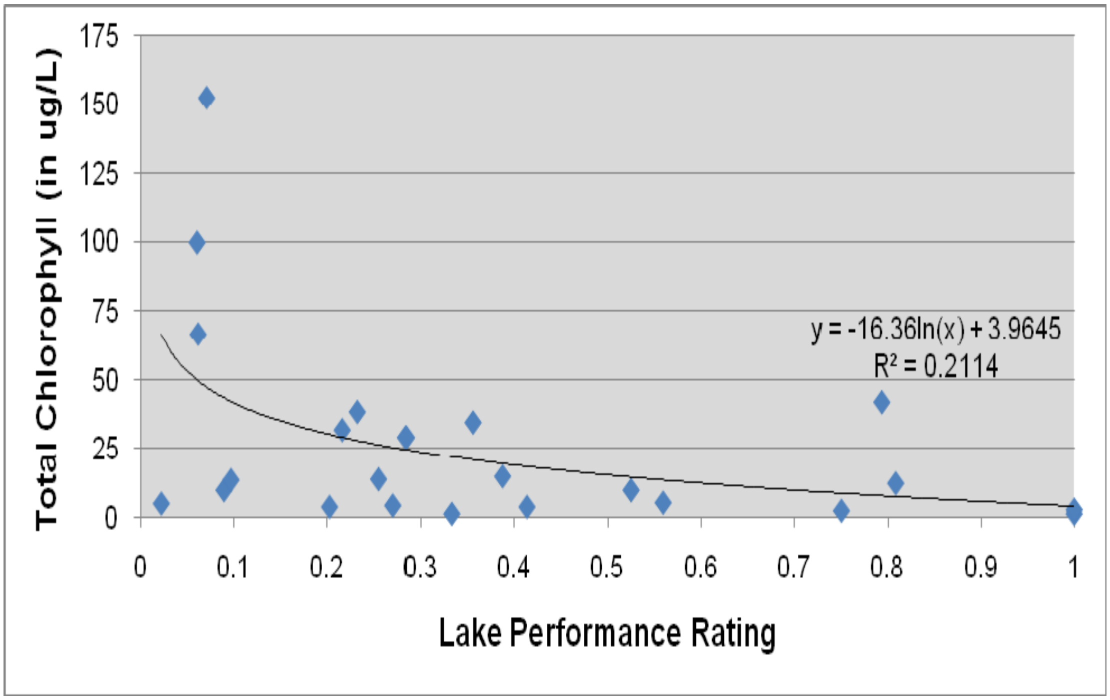


Figure 5. Total Chlorophyll Versus CCR-I Lake Performance Rating

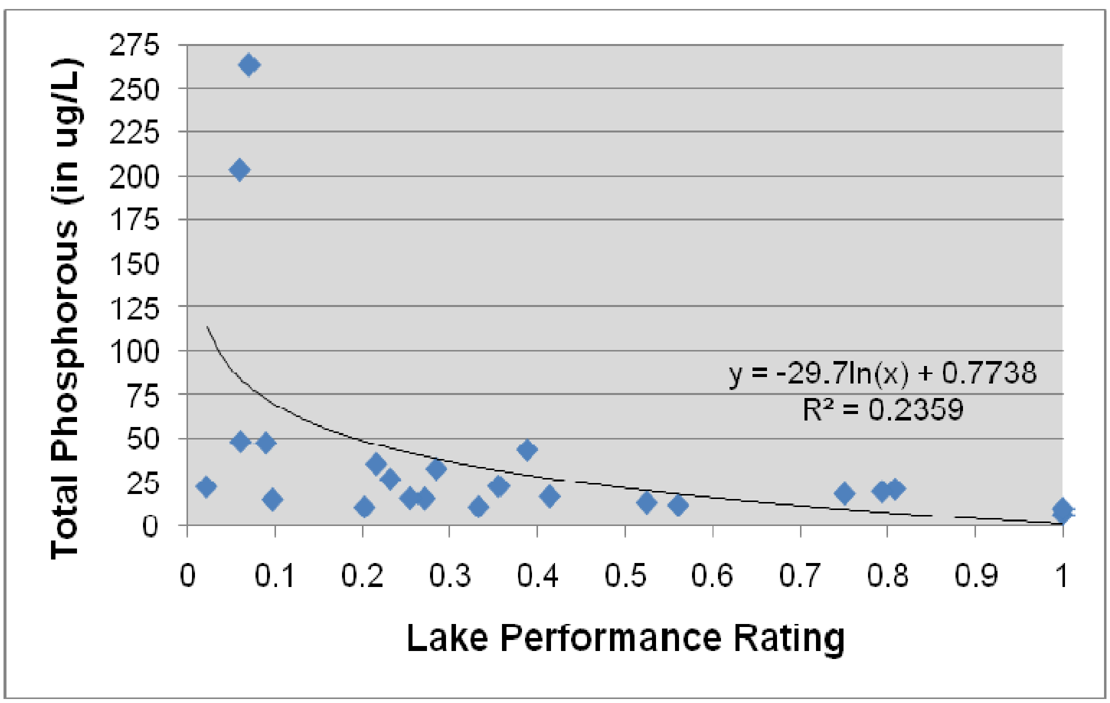


Figure 6. Total Phosphorous Versus CCR-I Lake Performance Rating

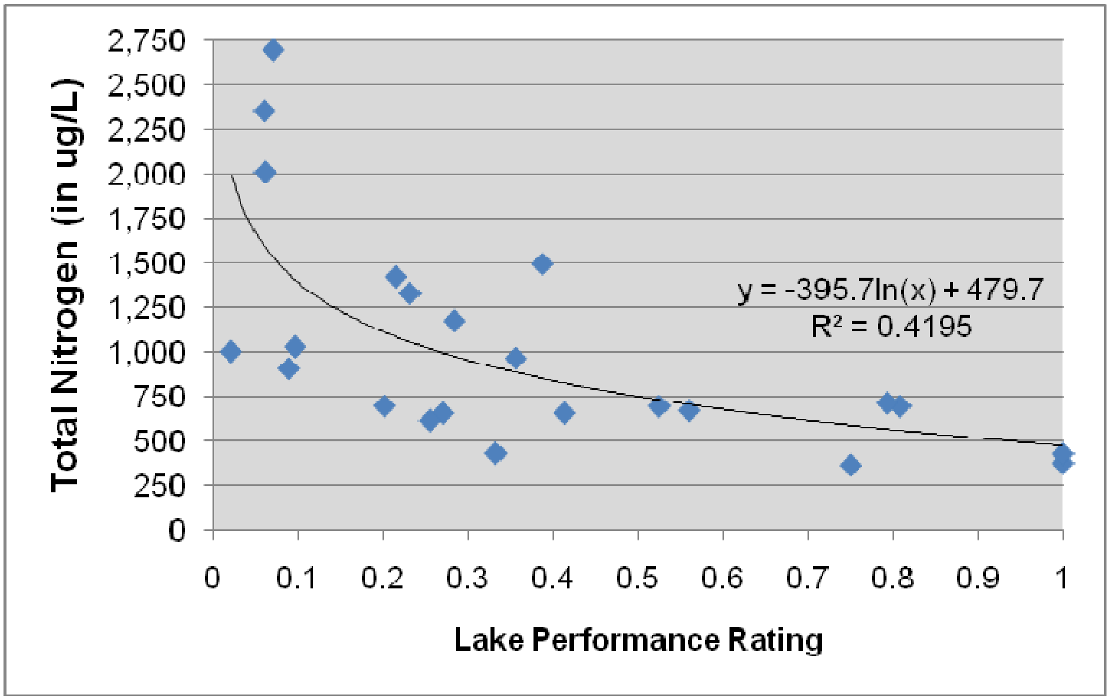


Figure 7. Total Nitrogen Versus CCR-I Lake Performance Rating

General trends in the data are graphically represented by Figures 8 and 9 for natural land area and natural land percentage, respectively. These scatter plots represent the strength of relationship between the output variable and lake performance. The statistical lines of best-fit are provided for both output oriented scatter plots. The relationship between natural land and lake performance is depicted by the statistical lines of best-fit. The best-fit lines support the trend that lakes surrounded by more natural land typically received higher performance ratings. As the single output variable increased, lake performance ratings typically improved. After observing this trend, it can be stated that lake performance shared a direct relationship with the output variable. Accordingly, lake performance optimization is achieved by maximizing outputs. This outcome supports the previously mentioned expectations that output maximization would result in optimum lake performance.

When viewing these graphs, the reader should keep in mind that the calculated R^2 values provided in each scatter plot are only a partial measure of the model's effectiveness. The R^2 values in each scatter plot quantify the strength of relationship between natural land and lake performance. The R^2 values provided in Figures 8 and 9 should not be considered a measure of how effective the entire model is at predicting lake performance.

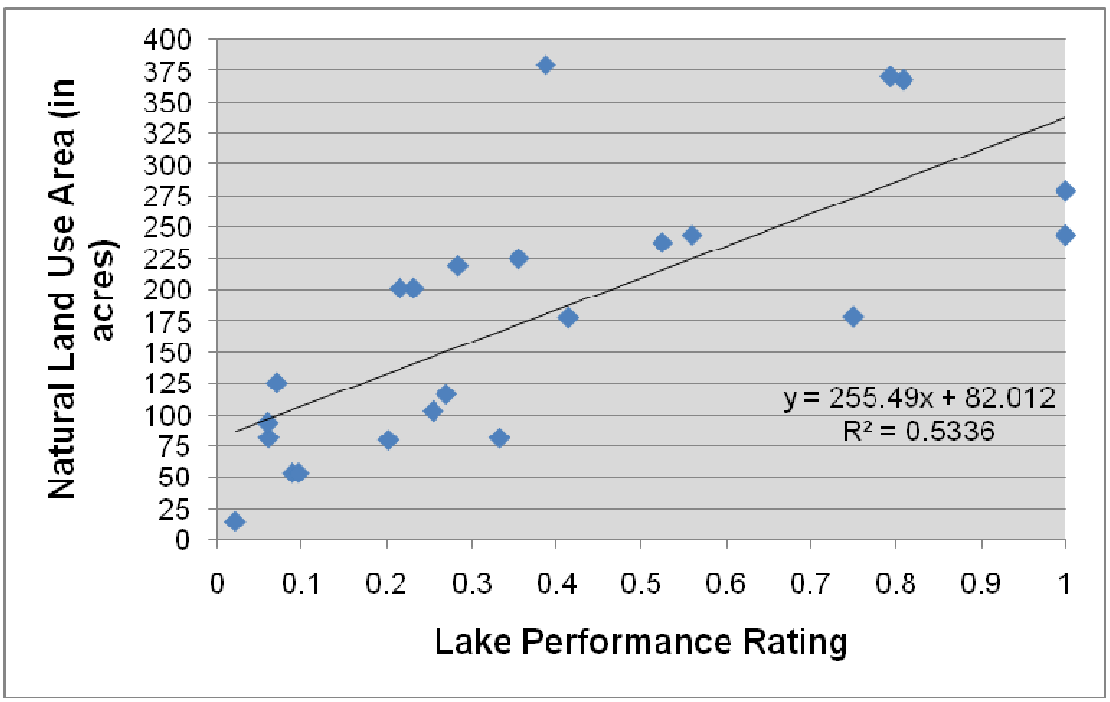


Figure 8. Natural Land Use Area Versus CCR-I Lake Performance Rating

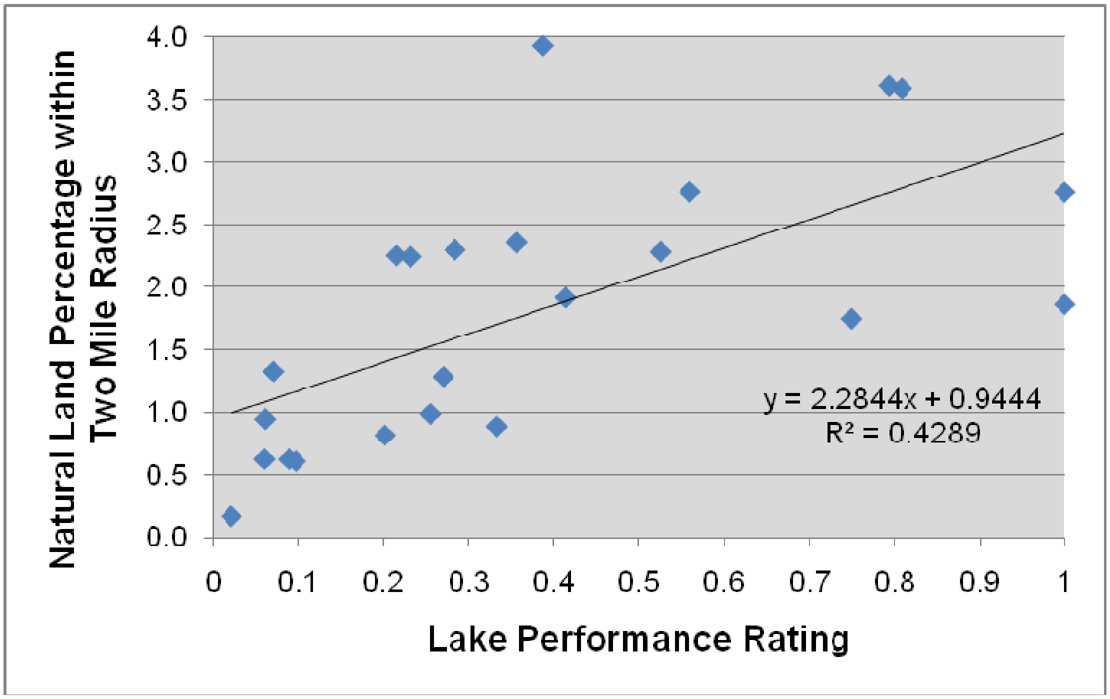


Figure 9. Natural Land Percentage Versus CCR-I Lake Performance Rating

Each of the lakes included in this study were assigned a ranking based on the DEA performance ratings. Two separate lakes claimed the number one rank as well as an optimum performance rating of one. Cypress Lake as well as Keystone Lake both achieved an optimum performance measurement. Cypress Lake was located within a two mile radius of 244.1508 acres classified as natural land, while Keystone Lake was surrounded by 279.7858 acres of natural land within a two mile radius. Naturally preserved land occupied 2.7560% of Cypress Lake's two mile radius delineation, while Keystone Lake was situated in a two mile radius delineation that contained 1.8600% natural land cover. Input variable concentrations for these two lakes were low relative to the nutrient loads of other inefficiently performing lakes included in the study.

Chapman Lake received the lowest performance rating amongst the set of lakes. This lake obtained the last ranking because it is located in a two mile radius delineation that only contained 14.4276 acres of naturally preserved land. Naturally preserved land only occupied 0.1616% of Chapman Lake's two mile radius boundary. The input variable concentrations for this lake also contributed to its poor performance rank. Each input variable concentration was high relative to the nutrient loads of other lakes included in the study. In fact, Chapman Lake contained concentrations of total phosphorous and total nitrogen higher than the sampled median. From this observation, it is evident that nutrient concentrations within Chapman Lake hindered its water quality performance.

Directly in the middle of the performance ranks, Rock Lake obtained a rating of 0.2842. This lake was representative of the average performance for the entire data set. As previously stated, the average performance rating for the entire set of lakes was equal to 0.3822. The standard deviation for lake performance ratings was equal to 0.2968. Therefore, the performance rating obtained by Rock Lake was representative of the CCR-I model results because it fell within one standard deviation of the average. Input and output variable data for Rock Lake was also representative of the averages calculated for the entire data set. Rock Lake was located within two miles of 219.3028 natural land acreage. On average, the two mile radius delineations established for this study contained 179.6641 acres of natural land. The standard deviation for natural land use area data collected during this study was equal to 103.8114 acres. Therefore, the natural land area surrounding Rock Lake was representative of the entire data set because it fell within one standard deviation of the average. Rock Lake was located in a two mile radius delineation occupied by 2.2959% natural land use. On average, the two mile radius delineations included in this study were occupied by 1.8175% natural land use. The standard deviation for natural land use percent surrounding the study lakes was equal to 1.0354%. Therefore, it can be stated that the percentage of natural land surrounding Rock Lake was representative of the entire data set because it fell within one standard deviation of the average. The total chlorophyll concentration for Rock Lake equaled 29.30 ug/L, while the average concentration of total chlorophyll for the entire data set was equal to 26.1870 ug/L. Rock Lake contained 33.30 ug/L of total phosphorous, while the average concentration of total phosphorous for the entire data set was equal to 41.1043 ug/L. Finally, Rock Lake recorded a 1,176.70 ug/L concentration of total nitrogen, which was comparable to the total nitrogen concentration average of 1,017.1217 ug/L calculated for the entire data set.

Table 8 displays a numeric value generated for the input concentrations of each DMU known as a ‘projection’. A DMU’s performance efficiency is improved when input variables are reduced or increased according to its ‘projection’. The numeric values provided for a DMU’s ‘projection’ can be either greater than or less than observed values. When a ‘projection’ is less than observed values, the performance efficiency of a DMU will be improved by radially reducing input values. Conversely, when a ‘projection’ is greater than observed values, the performance efficiency of a DMU will be improved by increasing input values. A 'projection' value is provided for each input variable. The DMU will attain optimal performance if the input variables are adjusted according to the 'projection'. 'DEA Solver' also provided the difference between the 'projection' and the actual data recorded for the input variable. The computed difference value was then converted into a percentage expressing the percent change in the input variable necessary to achieve optimal performance.

Table 8. Input Concentration, 'Projection', Difference, and Percent Difference

	Concentration (in ug/L)	'Projection'	Difference	Percent Difference
Chapman Lake	2.20E-02			
Total Chlorophyll (ug/L)	5.2	0.114456	-5.08554	-97.80%
Total Phosphorous (ug/L)	23	0.463594	-22.5364	-97.98%
Total Nitrogen (ug/L)	1004	22.09875	-981.901	-97.80%
Burrell Lake	8.95E-02			
Total Chlorophyll (ug/L)	10	0.57067	-9.42933	-94.29%
Total Phosphorous (ug/L)	48	1.997345	-46.0027	-95.84%
Total Nitrogen (ug/L)	906.7	81.16832	-825.532	-91.05%
Flynn Lake	9.72E-02			
Total Chlorophyll (ug/L)	13.7	0.326981	-13.373	-97.61%
Total Phosphorous (ug/L)	15.7	1.525912	-14.1741	-90.28%
Total Nitrogen (ug/L)	1030	81.74528	-948.255	-92.06%
Osceola, Lake	0.202258			
Total Chlorophyll (ug/L)	4	0.489753	-3.51025	-87.76%
Total Phosphorous (ug/L)	11.3	2.285516	-9.01448	-79.77%
Total Nitrogen (ug/L)	696.7	122.4384	-574.262	-82.43%
Garden Lake	6.16E-02			
Total Chlorophyll (ug/L)	66.7	0.845974	-65.854	-98.73%
Total Phosphorous (ug/L)	48.7	3.001683	-45.6983	-93.84%
Total Nitrogen (ug/L)	2013.3	124.0922	-1889.21	-93.84%
Sunset Lake	0.333099			
Total Chlorophyll (ug/L)	1.5	0.499649	-1.00035	-66.69%
Total Phosphorous (ug/L)	11.5	2.331693	-9.16831	-79.72%
Total Nitrogen (ug/L)	430	124.9121	-305.088	-70.95%
Thonotosassa, Lake	0.060586			
Total Chlorophyll (ug/L)	99.7	1.003863	-98.6961	-98.99%
Total Phosphorous (ug/L)	204	3.51352	-200.486	-98.28%
Total Nitrogen (ug/L)	2356.7	142.7828	-2213.92	-93.94%
Hiawatha, Lake	0.255189			
Total Chlorophyll (ug/L)	14	1.103406	-12.8966	-92.12%
Total Phosphorous (ug/L)	16.5	3.861922	-12.6381	-76.59%
Total Nitrogen (ug/L)	615	156.9411	-458.059	-74.48%
Echo Lake	0.270585			
Total Chlorophyll (ug/L)	4.5	1.217634	-3.28237	-72.94%
Total Phosphorous (ug/L)	16.3	4.310492	-11.9895	-73.56%
Total Nitrogen (ug/L)	656.7	177.6934	-479.007	-72.94%
Weeks, Lake	7.09E-02			
Total Chlorophyll (ug/L)	152.4	1.346225	-151.054	-99.12%
Total Phosphorous (ug/L)	264	4.711786	-259.288	-98.22%
Total Nitrogen (ug/L)	2699	191.478	-2507.52	-92.91%
Armistead, Lake	0.414212			
Total Chlorophyll (ug/L)	4	1.656848	-2.34315	-58.58%
Total Phosphorous (ug/L)	17.5	6.192533	-11.3075	-64.61%

Table 8 (continued). Input Concentration, ‘Projection’, Difference, and Percent Difference

Total Nitrogen (ug/L)	656.7	272.0131	-384.687	-58.58%
Alice, Lake	0.750035			
Total Chlorophyll (ug/L)	2.6	1.914198	-0.6858	-26.38%
Total Phosphorous (ug/L)	19	6.699693	-12.3003	-64.74%
Total Nitrogen (ug/L)	363	272.2628	-90.7372	-25.00%
Harvey, Lake	0.232295			
Total Chlorophyll (ug/L)	38.5	1.528914	-36.9711	-96.03%
Total Phosphorous (ug/L)	27.3	6.341656	-20.9583	-76.77%
Total Nitrogen (ug/L)	1330	308.9525	-1021.05	-76.77%
Virginia, Lake	0.216015			
Total Chlorophyll (ug/L)	32	2.16162	-29.8384	-93.24%
Total Phosphorous (ug/L)	36	7.565671	-28.4343	-78.98%
Total Nitrogen (ug/L)	1423.3	307.4545	-1115.85	-78.40%
Rock Lake	0.284234			
Total Chlorophyll (ug/L)	29.3	2.351472	-26.9485	-91.97%
Total Phosphorous (ug/L)	33.3	8.230151	-25.0698	-75.28%
Total Nitrogen (ug/L)	1176.7	334.4577	-842.242	-71.58%
Josephine Lake	0.355964			
Total Chlorophyll (ug/L)	34.7	2.334056	-32.3659	-93.27%
Total Phosphorous (ug/L)	23.3	8.293959	-15.006	-64.40%
Total Nitrogen (ug/L)	965	343.5052	-621.495	-64.40%
Pretty Lake	0.525283			
Total Chlorophyll (ug/L)	10	1.735405	-8.26459	-82.65%
Total Phosphorous (ug/L)	14	7.353956	-6.64604	-47.47%
Total Nitrogen (ug/L)	695	365.0714	-329.929	-47.47%
Cypress Lake	1			
Total Chlorophyll (ug/L)	1.5	1.5	0	0.00%
Total Phosphorous (ug/L)	7	7	0	0.00%
Total Nitrogen (ug/L)	375	375	0	0.00%
James, Lake	0.56			
Total Chlorophyll (ug/L)	5.5	1.5	-4	-72.73%
Total Phosphorous (ug/L)	12.5	7	-5.5	-44.00%
Total Nitrogen (ug/L)	670	375	-295	-44.03%
Keystone Lake	1			
Total Chlorophyll (ug/L)	3	3	0	0.00%
Total Phosphorous (ug/L)	10.5	10.5	0	0.00%
Total Nitrogen (ug/L)	426.7	426.7	0	0.00%
Brant, Lake	0.808587			
Total Chlorophyll (ug/L)	12.5	3.95103	-8.54897	-68.39%
Total Phosphorous (ug/L)	21.7	13.82861	-7.87139	-36.27%
Total Nitrogen (ug/L)	695	561.9682	-133.032	-19.14%
Thomas Lake	0.793604			
Total Chlorophyll (ug/L)	42	3.979923	-38.0201	-90.52%
Total Phosphorous (ug/L)	20.3	13.92973	-6.37027	-31.38%
Total Nitrogen (ug/L)	713.3	566.0777	-147.222	-20.64%
Hanna Lake	0.387944			
Total Chlorophyll (ug/L)	15	4.080367	-10.9196	-72.80%
Total Phosphorous (ug/L)	44	14.28128	-29.7187	-67.54%
Total Nitrogen (ug/L)	1496	580.3642	-915.636	-61.21%

For this particular study, input variable data belonging to inefficiently performing lakes must be reduced to achieve optimal water quality conditions. The ‘projection’ values for those lakes that did not perform optimally are all less than the observed input values. Therefore, the input concentrations must be minimized to achieve optimum lake water quality performance. The ‘projection’ values for both lakes that performed at optimum efficiency are the same as the observed input values. This observation reflects that no alterations to the actual input values are necessary to achieve optimum efficiency for those lakes that already received a performance rating of one. Cypress Lake and Keystone Lake both received performance ratings of one, and their ‘projection’ values are equal to the observed input values. This indicates that both Cypress Lake and Keystone Lake require no input concentration adjustments to function at an optimal level. According to the ‘projection’ values, the two lakes performing at optimum efficiency require no further adjustments regarding input variable concentrations. Meanwhile, lakes that performed less than efficiently must reduce input variable concentrations according to ‘projection’ values that were all less than actually observed values. The ‘projection’ values computed for this model support input minimization when attempting to improve lake water quality performance.

The CCR-I model is particularly useful because it generates a measurement known as a ‘slack’. This measurement is provided for the input and output variables of each DMU, or lake in this instance. ‘Slack’ is a scalar measurement that indicates the necessary input and output augmentations to produce an optimally performing DMU (Cooper et al. 2000). Input and output variables should be adjusted according to the numeric values provided for the ‘slack’ measurement. In this particular DEA, the ‘slacks’ measure excesses in the input variables and shortages in the single output variable. For this CCR-I model, the 'slack' measurement provided for the input variables indicate the reductions necessary to obtain optimum lake water quality performance. 'Slack' measurements for the single output variable indicate the increases required to obtain optimum lake water quality performance. Table 9 provided below displays the 'slack' measurements for the input and output variables of each lake.

Table 9. CCR-I 'Slack' Measurements

Lake Name	Rating	Excess in Total Chlorophyll (in ug/L)	Excess in Total Phosphorous (in ug/L)	Excess in Total Nitrogen (in ug/L)	Shortage in Natural Land Area (in acres)
Chapman Lake	0.0220	0.0000	0.0427	0.0000	0.0000
Burrell Lake	0.0895	0.3245	2.2996	0.0000	0.0000
Flynn Lake	0.0972	1.0045	0.0000	18.3623	0.0000
Osceola, Lake	0.2023	0.3193	0.0000	18.4748	0.0000
Garden Lake	0.0616	3.2652	0.0000	0.0000	0.0000
Sunset Lake	0.3331	0.0000	1.4989	18.3204	0.0000
Thonotosassa, Lake	0.0606	5.0366	8.8460	0.0000	0.0000
Hiawatha, Lake	0.2552	2.4692	0.3487	0.0000	0.0000
Echo Lake	0.2706	0.0000	0.1000	0.0000	0.0000
Weeks, Lake	0.0709	9.4656	14.0174	0.0000	0.0000
Armistead, Lake	0.4142	0.0000	1.0562	0.0000	0.0000
Alice, Lake	0.7500	0.0359	7.5510	0.0000	0.0000
Harvey, Lake	0.2323	7.4144	0.0000	0.0000	0.0000
Virginia, Lake	0.2160	4.7509	0.2109	0.0000	0.0000
Rock Lake	0.2842	5.9766	1.2348	0.0000	0.0000
Josephine Lake	0.3560	10.0179	0.0000	0.0000	0.0000
Pretty Lake	0.5253	3.5174	0.0000	0.0000	0.0000
Cypress Lake	1.0000	0.0000	0.0000	0.0000	0.0000
James, Lake	0.5600	1.5800	0.0000	0.2000	0.0000
Keystone Lake	1.0000	0.0000	0.0000	0.0000	0.0000
Brant, Lake	0.8086	6.1563	3.7177	0.0000	0.0000
Thomas Lake	0.7936	29.3514	2.1804	0.0000	0.0000
Hanna Lake	0.3879	1.7388	2.7883	0.0000	0.0000

For this particular model, input variable data belonging to inefficiently performing lakes must be reduced to achieve optimal water quality conditions. The 'slack' measurements for input variables of those lakes that did not perform optimally represent reductions. Therefore, the input concentrations must be minimized to achieve optimum lake water quality performance. 'Slack' measurements for the output variable did not vary from zero because the input oriented CCR model was applied. The 'slacks' for the output variable were not considered by the model because it was input oriented.

Inefficiently performing lakes were capable of receiving 'slack' measurements equal to zero for individual variables, however, these lakes could not receive 'slack' measurements of zero for each variable. The 'slack' measurement for at least a single variable was not equal to zero for those DMUs that did not obtain an optimum performance rating of one. These 'slack' measurements reflected that efficiently performing lakes such as Cypress and Keystone required no adjustments to achieve optimization.

The final component explaining the results from the CCR-I model is a correlation matrix. The correlation matrix contains proportions that describe the relationship between variables included in the CCR-I model. These proportions are derived from correlation coefficients and are known as coefficients of determination. A coefficient of determination is a proportion that reveals the interrelatedness between variables. This numerical value quantifies how much a particular variable is responsible for the outcome of an accompanying variable. The correlation matrix provides numerical values for the proportion of a variable that explains another variable. For this CCR-I model, the proportions of particular interest are those that provide descriptions of the relationship between input and output variables. These proportions will quantify the level of interaction between input and output variables. Interpretation of these figures will expose how greatly natural land use area influences lake nutrient concentrations. Correlation proportions reveal the strength of relationship between natural land use area and nutrient concentrations. Proportions describing input and output interaction will quantify the influence of natural land use area in regulating lake nutrient concentrations. The correlation matrix for the CCR-I model has been provided in Table 10 below.

Table 10. CCR-I Variable Correlation Matrix

	Total Chlorophyll	**Total Phosphorous**	**Total Nitrogen**	**Natural Land Use Area**
Total Chlorophyll	1	0.9224	0.9019	0.1188
Total Phosphorous	0.9224	1	0.8507	0.2007
Total Nitrogen	0.9019	0.8507	1	0.1982
Natural Land Use Area	0.1188	0.2007	0.1982	1

BCC-I Model Results

Performance ratings computed by the BCC-I model describe the relationship between lake nutrient loads and surrounding land use. The BCC-I model returned an average performance rating of 0.5824 for the entire set of lakes. This value indicates that the entire set of lakes operated at 58.24% of optimal performance efficiency, which was a significant improvement over the results from the CCR-I model. The standard deviation for the entire set of lake performance ratings equaled 0.3069. This value indicates a relatively wide distribution of performance ratings for the entire set of lakes. Performance ratings from the previously conducted CCR-I model had a standard deviation equal to 0.2968, which was similar to that of the BCC-I model. Both input oriented models obtained performance rating standard deviations that reflected a wide range of results. The lake water quality performance ratings derived from the BCC-I model have been provided in Table 11. As expected, these ratings increase for lakes surrounded by greater amounts of naturally preserved land. Performance ratings also predictably increase for those lakes containing lower concentrations of the selected input variables.

Table 11. BCC-I Lake Performance Ratings and Rank

Lake Name	Performance Rating	Lake Performance Rank
Alice, Lake	1.0000	1
Brant, Lake	1.0000	1
Cypress Lake	1.0000	1
Hanna Lake	1.0000	1
Keystone Lake	1.0000	1
Thomas Lake	1.0000	1
Sunset Lake	1.0000	7
Osceola, Lake	0.6195	8
Hiawatha, Lake	0.6049	9
Echo Lake	0.5676	10
Armistead, Lake	0.5666	11
James, Lake	0.5600	12
Pretty Lake	0.5388	13
Flynn Lake	0.4459	14
Burrell Lake	0.4003	15
Josephine Lake	0.3865	16
Chapman Lake	0.3720	17
Rock Lake	0.3157	18
Harvey, Lake	0.2814	19
Virginia, Lake	0.2618	20
Garden Lake	0.1853	21
Thonotosassa, Lake	0.1540	22
Weeks, Lake	0.1345	23

The BCC-I performance ratings generated during this study follow general trends that describe the relationship between lake water quality and land use. General trends in the data are graphically represented by Figures 10, 11, and 12 for total chlorophyll, total phosphorous, and total nitrogen, respectively. These scatter plots represent the strength of relationship between a respective input variable and lake performance. The statistical lines of best-fit are provided for each of these scatter plots. The relationship between each water quality parameter and lake performance is depicted by the statistical lines of best-fit.

The best-fit lines support the trend that lakes containing lower input variable concentrations typically obtained higher performance ratings. After observing this trend, it can be stated that lake performance shared an indirect relationship with the input variables. Accordingly, lake performance optimization is achieved by minimizing inputs. This outcome supports the previously mentioned expectations that input minimization would result in optimum lake performance.

When viewing these graphs, the reader should keep in mind that the calculated R^2 values provided in each scatter plot are only a partial measure of the model's effectiveness. The R^2 values in each scatter plot quantify the strength of relationship between the plotted water quality parameter and lake performance. The R^2 values provided in Figures 10, 11, and 12 should not be considered a measure of how effective the entire model is at predicting lake performance.

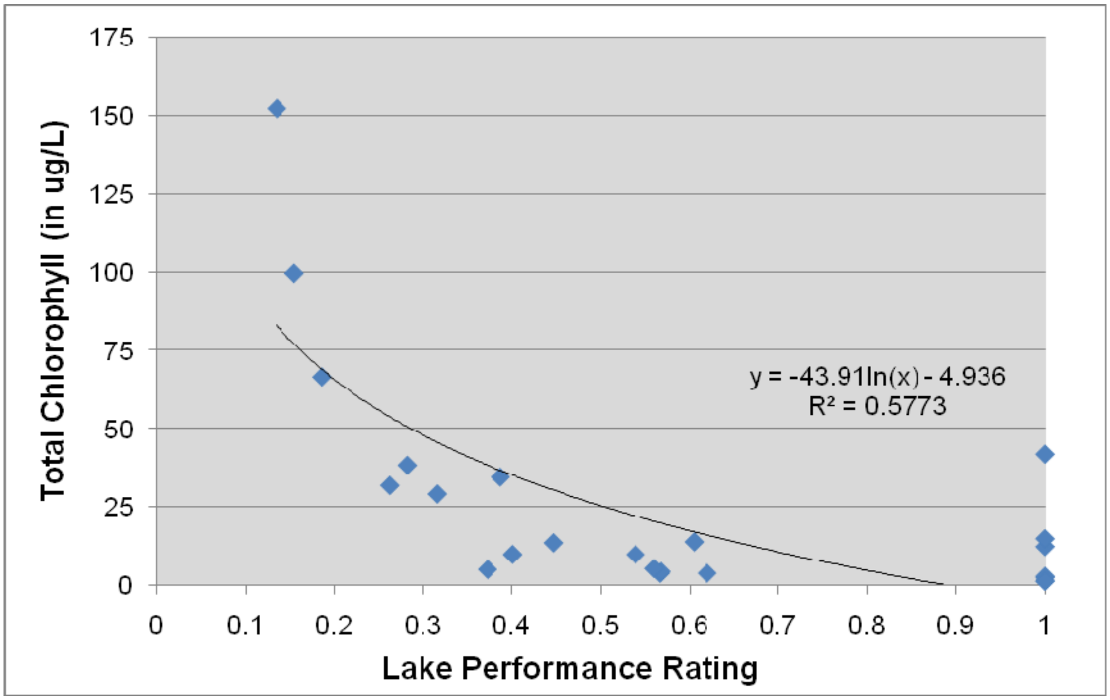


Figure 10. Total Chlorophyll Versus BCC-I Lake Performance Rating

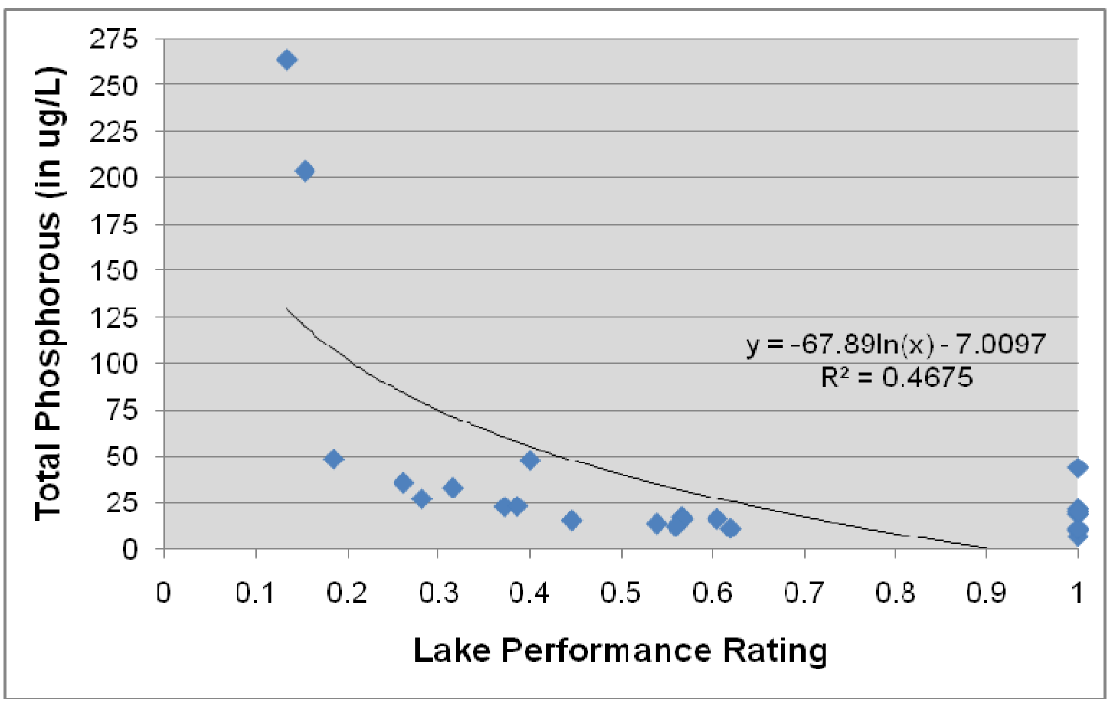


Figure 11. Total Phosphorous Versus BCC-I Lake Performance Rating

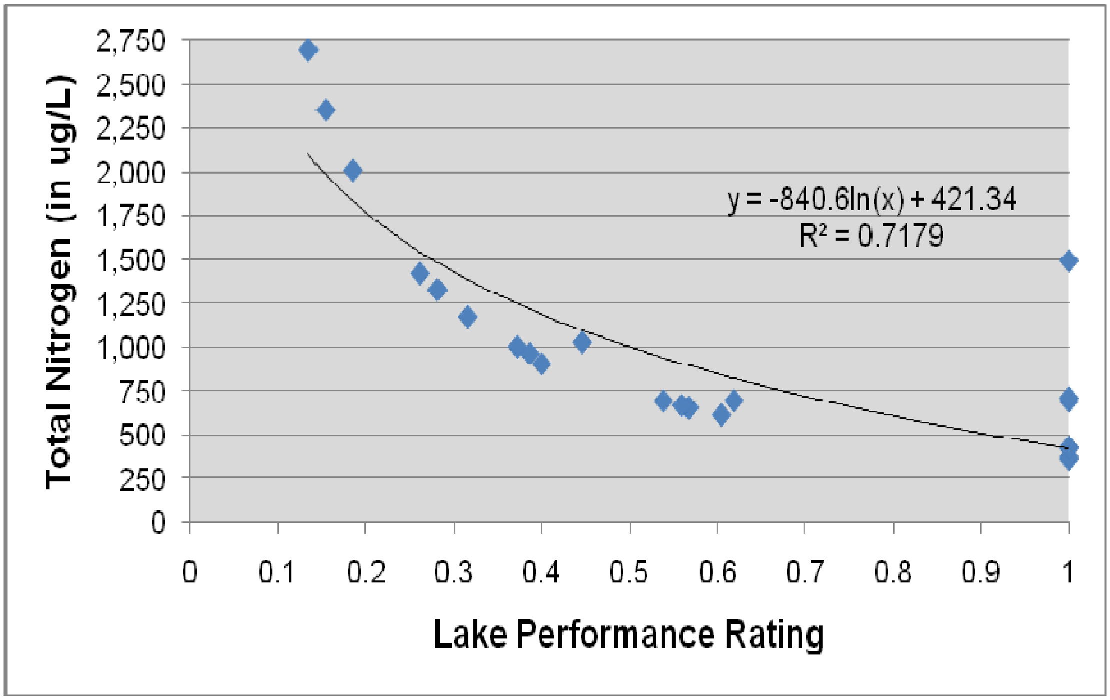


Figure 12. Total Nitrogen Versus BCC-I Lake Performance Rating

General trends in the data are graphically represented by Figures 13 and 14 for natural land area and natural land percentage, respectively. These scatter plots represent the strength of relationship between the output variable and lake performance. The statistical lines of best-fit are provided for both output oriented scatter plots. The relationship between natural land and lake performance is depicted by the statistical lines of best-fit. The best-fit lines support the trend that lakes surrounded by more natural land typically received higher performance ratings. As the single output variable increased, lake performance ratings typically improved. After observing this trend, it can be stated that lake performance shared a direct relationship with the output variable. Accordingly, lake performance optimization is achieved by maximizing outputs. This outcome supports the previously mentioned expectations that output maximization would result in optimum lake performance.

When viewing these graphs, the reader should keep in mind that the calculated R^2 values provided in each scatter plot are only a partial measure of the model's effectiveness. The R^2 values in each scatter plot quantify the strength of relationship between natural land and lake performance. The R^2 values provided in Figures 13 and 14 should not be considered a measure of how effective the entire model is at predicting lake performance.

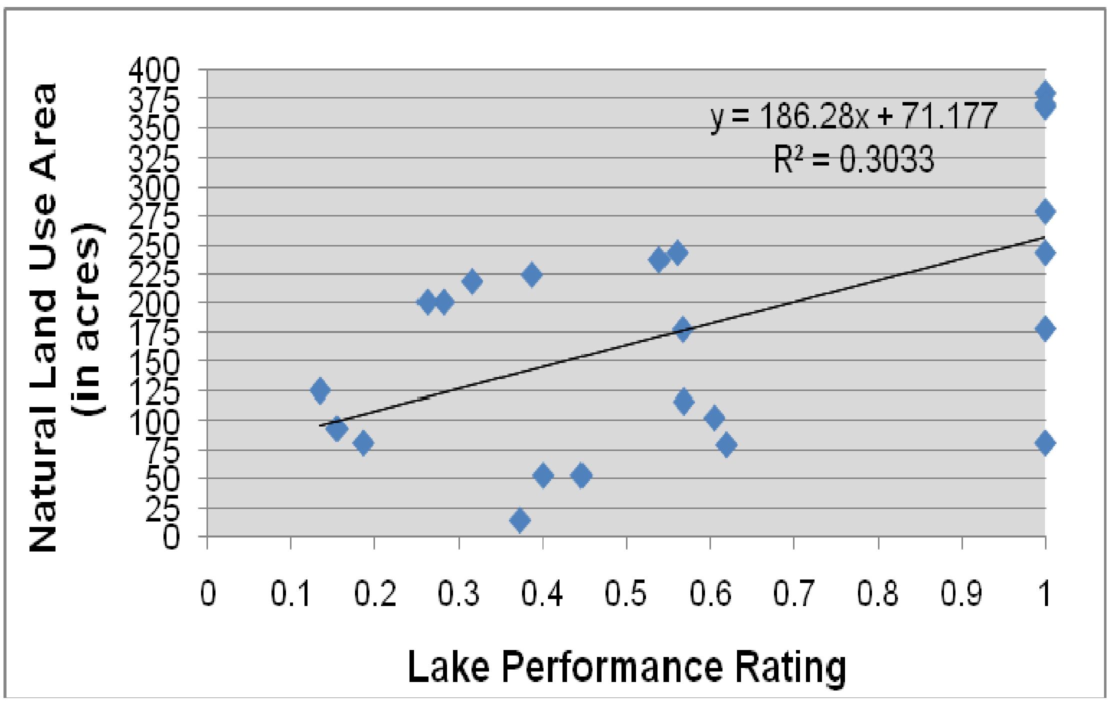


Figure 13. Natural Land Use Area Versus BCC-I Lake Performance Rating

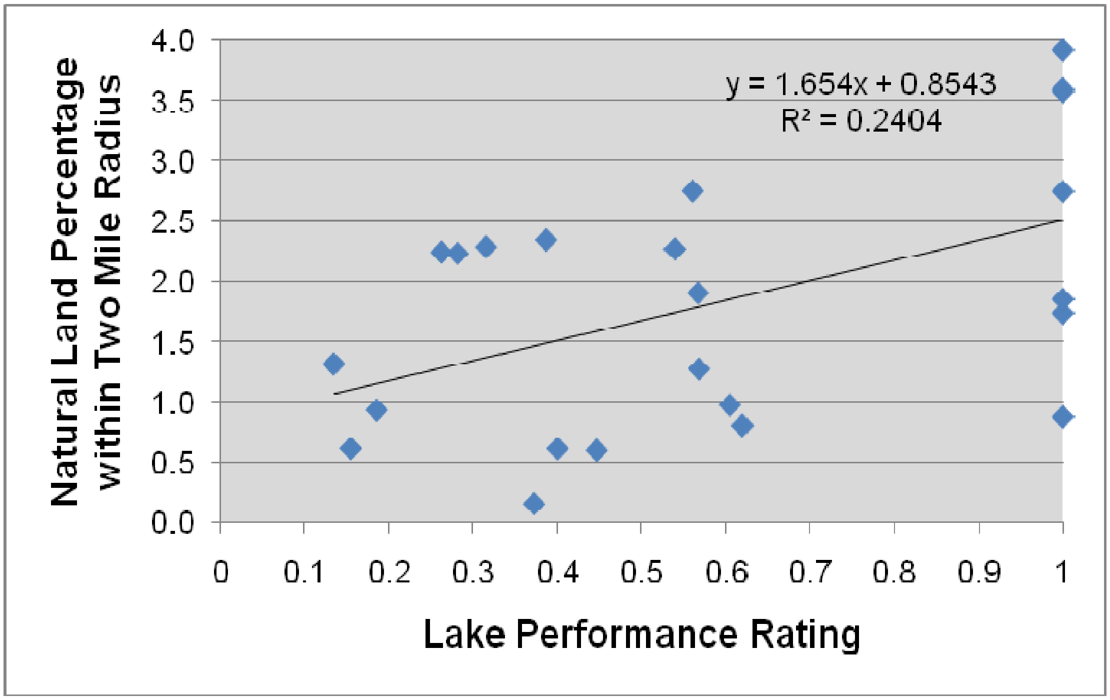


Figure 14. Natural Land Percentage Versus BCC-I Lake Performance Rating

Each of the lakes included in this study were assigned a ranking based on the performance ratings. Seven separate lakes claimed a performance rating of one, however, only six were classified as optimally performing. Sunset Lake achieved a performance rating of one, but it failed to obtain 'slack' measurements equal to zero for each of the input variables. Therefore, the number one rank as well as an optimum performance rating of one was only achieved by the following six lakes: Lake Alice, Lake Brant, Cypress Lake, Hanna Lake, Keystone Lake, and Thomas Lake. 178.5218 acres of natural land were located within a two mile radius from Lake Alice translating into 1.7454% of Lake Alice's two mile radius delineation. Lake Brant was situated in a two mile radius delineation that contained 368.4807 acres of naturally preserved land. Naturally preserved land occupied 3.5855% of Lake Brant's two mile radius delineation. Cypress Lake was located within a two mile radius of 244.1508 acres classified as natural land. Naturally preserved land occupied 2.7560% of Cypress Lake's two mile radius delineation. 380.5429 acres of natural land were located within a two mile radius from Hanna Lake. Hanna Lake was situated in a two mile radius delineation that contained 3.9308% natural land cover. Keystone Lake was surrounded by 279.7858 acres of natural land within a two mile radius. Keystone Lake was situated in a two mile radius delineation that contained 1.8600% natural land cover. Finally, Thomas Lake was situated in a two mile radius delineation that contained 371.1753 acres of naturally preserved land. Natural land occupied 3.6078% of the two mile radius surrounding Thomas Lake. Input variable concentrations for these lakes were low relative to the nutrient loads of other inefficiently performing lakes included in the study.

Lake Weeks received the lowest performance rating amongst the set of lakes at 0.1345. This lake obtained the last ranking because it is located in a two mile radius delineation with only 125.5515 acres of naturally preserved land. This acreage of naturally preserved land was significantly below the average calculated for the entire data set which was 179.6641. Naturally preserved land only occupied 1.3223% of Lake Week's two mile radius delineation. This percentage of naturally preserved land was significantly below the average calculated for the entire data set which was 1.8175%. The input variable concentrations for this lake also contributed to its poor performance rank. Each input variable concentration was high relative to the nutrient loads of other lakes included in the study. In fact, Lake Weeks consistently contained the highest concentrations of total chlorophyll, total phosphorous and total nitrogen. From this observation, it is blatantly evident that nutrient concentrations within Lake Weeks hindered its water quality performance.

Directly in the middle of the performance ranks, Lake James obtained a rating of 0.5600. This lake was representative of the average performance for the entire data set. As previously stated, the average performance rating for the entire set of lakes was equal to 0.5824. Input and output variable data for Lake James was also representative of the averages calculated for the entire data set. Lake James was located within two miles of 244.1508 acres of natural land. On average, the two mile radius delineations established for this study contained 179.6641 acres of natural land. The standard deviation for natural land use area data collected during this study was equal to 103.8114 acres. Therefore, the natural land area surrounding Lake James was representative of the entire data set because it fell within one standard deviation of the average. Lake James was located in a two mile radius delineation occupied by 2.7597% natural land use. On average, the two mile radius delineations established in this study were occupied by 1.8175% natural land use. The standard deviation for natural land use percent surrounding the study lakes was equal to 1.0354%. Therefore, it can be stated that the percentage of natural land surrounding Lake James was representative of the entire data set because it fell within one standard deviation of the average. The total chlorophyll concentration for Lake James equaled 5.50 ug/L, while the average concentration of total chlorophyll for the entire data set was equal to 26.1870 ug/L. Lake James contained 12.50 ug/L of total phosphorous, while the average concentration of total phosphorous for the entire data set was equal to 41.1043 ug/L. Finally, Lake James recorded a 670.00 ug/L concentration of total nitrogen, which was somewhat comparable to the total nitrogen concentration average of 1,017.1217 ug/L calculated for the entire data set.

Like the CCR-I model, the BCC-I model is particularly useful because it generates a measurement known as a ‘slack’. This measurement is provided for the input and output variables of each DMU, or lake in this instance. A lake’s performance efficiency is optimized when input and output variables are reduced or increased according to its ‘slack’ (Cooper et al. 2000). Input and output variables should be adjusted according to the numeric values provided for the ‘slack’ measurement. In this particular DEA, the ‘slacks’ measure excesses in the input variables and shortages in the single output variable. Therefore, input variable ‘slacks’ represent reductions required to accomplish optimum DMU performance, and output variable ‘slacks’ represent increases required to accomplish optimum DMU performance. Table 12 provided below displays the ‘slack’ measurements for the input and output variables of each lake.

Table 12. BCC-I ‘Slack’ Measurements

Lake Name	Rating	Excess in Total Chlorophyll (in ug/L)	Excess in Total Phosphorous (in ug/L)	Excess in Total Nitrogen (in ug/L)	Shortage in Natural Land Area (in acres)
Chapman Lake	0.3720	0.2916	0.0000	0.0000	221.2163
Burrell Lake	0.4003	1.4035	0.2169	0.0000	125.2982
Flynn Lake	0.4459	4.6082	0.0000	84.2348	190.9266
Osceola, Lake	0.6195	0.9779	0.0000	56.5835	164.4326
Garden Lake	0.1853	10.6712	0.0000	0.0000	151.7635
Sunset Lake	1.0000	0.0000	4.5000	54.9995	162.8220
Thonotosassa, Lake	0.1540	12.7566	12.4218	0.0000	84.8978
Hiawatha, Lake	0.6049	6.6954	0.0000	0.0000	124.9394
Echo Lake	0.5676	0.8478	0.0000	0.0000	115.3675
Weeks, Lake	0.1345	17.8967	16.5063	0.0000	52.9685
Armistead, Lake	0.5666	0.4991	0.0000	0.0000	50.2358
Alice, Lake	1.0000	0.0000	0.0003	0.0000	0.0000
Harvey, Lake	0.2814	9.2727	0.0000	0.0000	38.8142
Virginia, Lake	0.2618	6.6544	0.0000	0.0000	29.2962
Rock Lake	0.3157	7.4280	0.0000	0.0000	5.6334
Josephine Lake	0.3865	11.7283	0.0000	0.0000	8.0662
Pretty Lake	0.5388	3.8381	0.0000	0.0000	3.0719
Cypress Lake	1.0000	0.0000	0.0000	0.0000	0.0000
James, Lake	0.5600	1.5800	0.0000	0.2000	0.0000
Keystone Lake	1.0000	0.0000	0.0000	0.0000	0.0000
Brant, Lake	1.0000	0.0000	0.0000	0.0000	0.0000
Thomas Lake	1.0000	0.0000	0.0000	0.0000	0.0000
Hanna Lake	1.0000	0.0000	0.0000	0.0000	0.0000

For this particular model, input variable data belonging to inefficiently performing lakes must be reduced to achieve optimal water quality conditions. Output variable data belonging to inefficiently performing lakes must be increased to achieve optimal water quality conditions. Inefficiently performing lakes were capable of receiving ‘slack’ measurements equal to zero for individual variables, however, these lakes could not receive ‘slack’ measurements of zero for each variable. With the exception of Sunset Lake, ‘slack’ measurements for each variable of those lakes that achieved optimum performance ratings equaled zero. These ‘slack’ measurements reflected that efficiently performing lakes required no adjustments to achieve optimization. Sunset Lake received a performance rating of one, however, several ‘slack’ measurements belonging to both these DMUs were not equal to zero and indicated that changes were necessary to achieve optimization. This represents a rare case in which a DMU obtains a performance rating of one, but the ‘slack’ measurements indicate that variable adjustments are still necessary to acquire optimized performance. It should be mentioned that Lake Alice did not obtain a ‘slack’ measurement of zero for each variable. Lake Alice’s ‘slack’ measurement for total phosphorous was equal to 0.0003. It was determined that such a small measurement was negligible when considering the overall performance of Lake Alice. For the purpose of this analysis, Lake Alice was considered an optimally performing DMU, and its negligible non-zero ‘slack’ measurement for total phosphorous was ignored. Overall, the ‘slack’ measurements support input minimization and output maximization when attempting to achieve optimum lake performance.

The final component explaining the results from the BCC-I model is a correlation matrix. The correlation matrix contains proportions that describe the relationship between variables included in the BCC-I model. The correlation matrix provides numerical values for the proportion of a variable that explains another variable. For this BCC-I model, the proportions of particular interest are those that provide descriptions of the relationship between input and output variables. These proportions will quantify the level of interaction between input and output variables. Interpretation of these figures will expose how greatly natural land use area influences lake nutrient concentrations. Correlation proportions reveal the strength of relationship between natural land and nutrient concentrations. Proportions describing input and output interaction will quantify the influence of natural land use area in regulating lake nutrient concentrations. The correlation matrix for the BCC-I model of this thesis has been provided in Table 13 below. If the reader refers to the correlation matrix provided for the CCR-I model, it will be evident that the proportions within this table are equal between the two input oriented models.

Table 13. BCC-I Variable Correlation Matrix

	Total Chlorophyll	**Total Phosphorous**	**Total Nitrogen**	**Natural Land Use Area**
Total Chlorophyll	1	0.9224	0.9019	0.1188
Total Phosphorous	0.9224	1	0.8507	0.2007
Total Nitrogen	0.9019	0.8507	1	0.1982
Natural Land Use Area	0.1188	0.2007	0.1982	1

Additive Model Results

The Additive model produces performance ratings influenced equally by both input and output variables. These ratings describe the relationship between lake performance and nutrient loads as well as lake performance and natural land use area. Primarily, the performance ratings derived from the Additive model are intended to describe the relationship between lake nutrient loads and surrounding land use. The Additive model returned an average performance rating of 0.6333 for the entire set of lakes, which was similar to the average performance rating obtained during the BCC-I model. This value indicates that the entire set of lakes operated at 63.33% of optimal performance efficiency, which was a significant improvement over the results from the CCR-I model but similar to the results obtained during the BCC-I model. The standard deviation for the entire set of lake performance ratings equaled 0.2671. This value indicates a relatively wide distribution of performance ratings for the entire set of lakes. Performance ratings from the previously conducted CCR-I and BCC-I models had standard deviations equal to 0.2968 and 0.3069, respectively. Performance rating standard deviations calculated for each of the models are relatively similar and reflect a wide range of results. The lake water quality performance ratings derived from the Additive model have been provided in Table 14 along with a column for ‘stability’ values. The meaning and interpretation of ‘stability’ values will be discussed further in latter portions of this section. As expected, these ratings increase for lakes surrounded by greater amounts of naturally preserved land. Performance ratings also predictably increase for those lakes containing lower concentrations of the selected input variables.

Table 14. Additive Lake Performance Summary

Lake Name	Performance Rating	'Stability' Value	Lake Performance Rank According to 'Stability'
Brant, Lake	1.0000	0.1713	1
Cypress Lake	1.0000	0.0851	2
Hanna Lake	1.0000	0.0650	3
Keystone Lake	1.0000	0.0393	4
Thomas Lake	1.0000	0.0273	5
Alice, Lake	1.0000	0.0118	6
Sunset Lake	0.6463	0.0000	7
James, Lake	0.8014	-0.0936	8
Armistead, Lake	0.7009	-0.0955	9
Osceola, Lake	0.6203	-0.0955	10
Echo Lake	0.6377	-0.1146	11
Pretty Lake	0.7345	-0.1293	12
Chapman Lake	0.4910	-0.1413	13
Flynn Lake	0.4955	-0.2117	14
Hiawatha, Lake	0.5748	-0.2311	15
Josephine Lake	0.4991	-0.3091	16
Burrell Lake	0.4462	-0.3246	17
Harvey, Lake	0.4318	-0.4167	18
Rock Lake	0.4713	-0.4851	19
Virginia, Lake	0.4275	-0.5613	20
Garden Lake	0.2803	-0.9820	21
Weeks, Lake	0.1356	-1.4043	22
Thonotosassa, Lake	0.1715	-1.5493	23

Each of the lakes included in this study were assigned a ranking based on the 'stability' value. The 'stability' value is a measurement unique to the Additive model. This measurement quantifies DMU efficiency at a finer scale than the traditional DEA performance rating. DMUs that receive a maximum performance rating of one are assigned positive 'stability' values, while inefficiently performing DMUs are assigned negative 'stability' values. 'Stability' values directly indicate DMU efficiency levels. DMUs with higher levels of efficiency receive greater 'stability' values. The major advantage of 'stability' values is that it allows the analysis to further rank the efficiency of those DMUs with a maximum performance rating of one.

The Additive 'stability' values generated during this study follow general trends that describe the relationship between lake water quality and land use. General trends in the data are graphically represented by Figures 15, 16, and 17 for total chlorophyll, total phosphorous, and total nitrogen, respectively. These scatter plots represent the strength of relationship between a respective input variable and lake 'stability' values. The statistical lines of best-fit are provided for each of these scatter plots. The relationship between each water quality parameter and lake 'stability' values is depicted by the statistical lines of best-fit. The best-fit lines support the trend that lakes containing lower input variable concentrations typically obtained higher 'stability' values. After observing this trend, it can be stated that lake performance shared an indirect relationship with the input variables. Accordingly, lake performance optimization is achieved by minimizing inputs. This outcome supports the previously mentioned expectations that input minimization would result in optimum lake performance.

When viewing these graphs, the reader should keep in mind that the calculated R^2 values provided in each scatter plot are only a partial measure of the model's effectiveness. The R^2 values in each scatter plot quantify the strength of relationship between the plotted water quality parameter and lake ‘stability’ values. The R^2 values provided in Figures 15, 16, and 17 should not be considered a measure of how effective the entire model is at predicting lake performance.

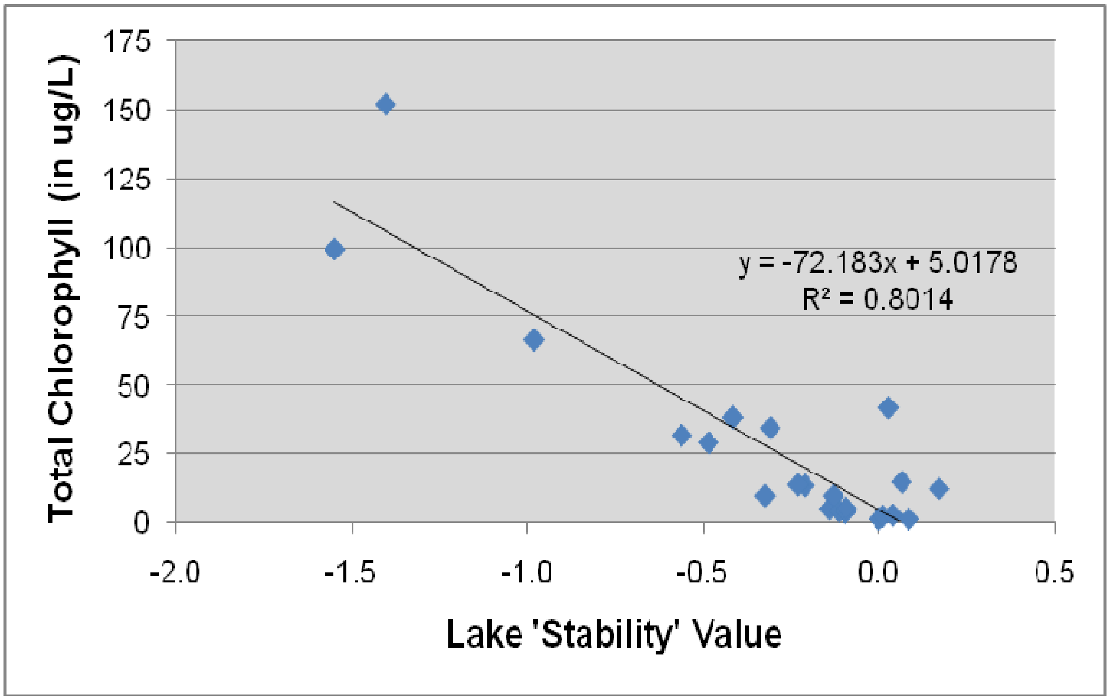


Figure 15. Total Chlorophyll Versus Additive Lake ‘Stability’ Value

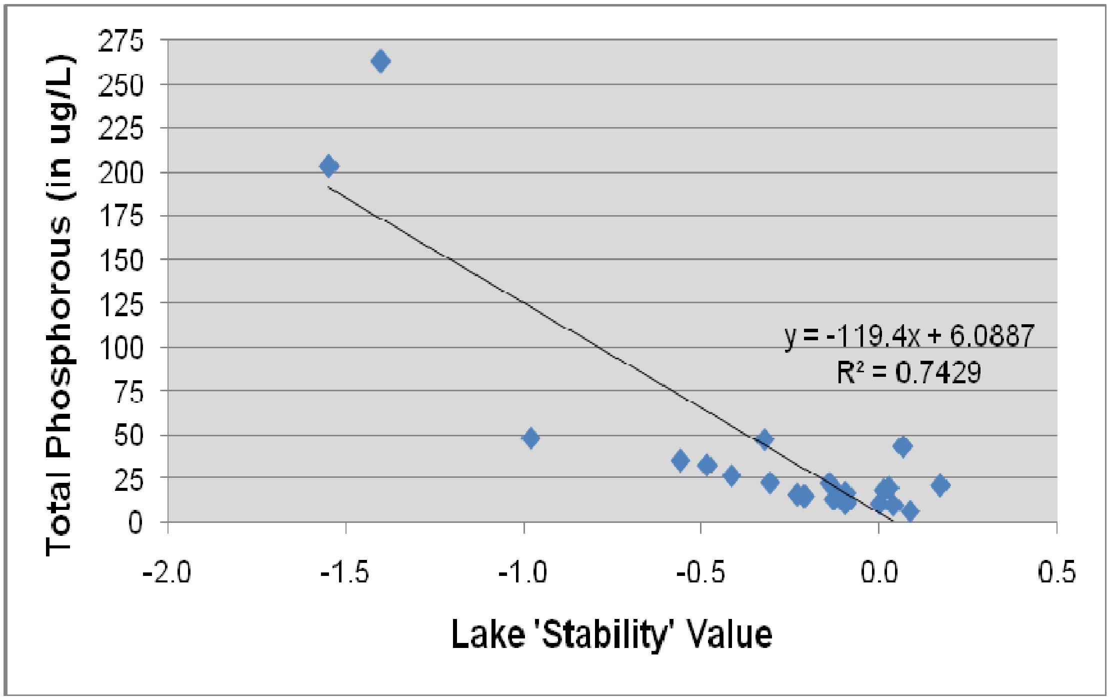


Figure 16. Total Phosphorous Versus Additive Lake ‘Stability’ Value

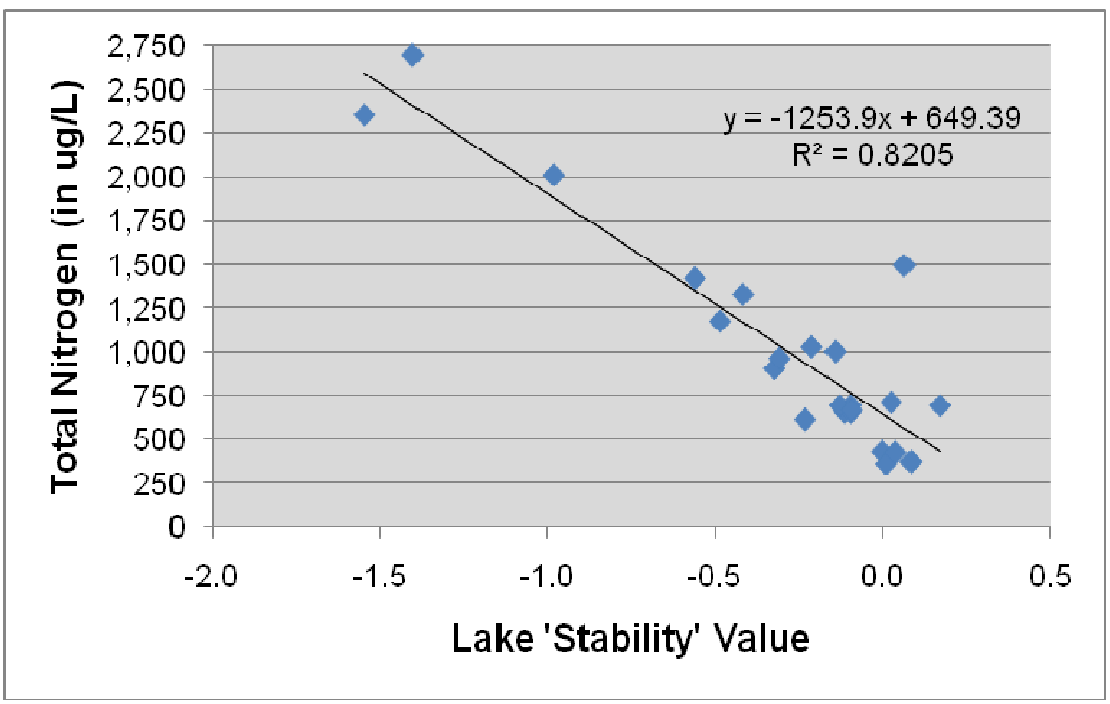


Figure 17. Total Nitrogen Versus Additive Lake ‘Stability’ Value

General trends in the data are graphically represented by Figures 18 and 19 for natural land area and natural land percentage, respectively. These scatter plots represent the strength of relationship between the output variable and lake ‘stability’ values. The statistical lines of best-fit are provided for both output oriented scatter plots. The relationship between natural land and lake ‘stability’ values is depicted by the statistical lines of best-fit. The best-fit lines support the trend that lakes surrounded by more natural land typically received higher ‘stability’ values. As the single output variable increased, lake ‘stability’ values typically improved. After observing this trend, it can be stated that lake performance shared a direct relationship with the output variable. Accordingly, lake performance optimization is achieved by maximizing outputs. This outcome supports the previously mentioned expectations that output maximization would result in optimum lake performance.

When viewing these graphs, the reader should keep in mind that the calculated R^2 values provided in each scatter plot are only a partial measure of the model’s effectiveness. The R^2 values in each scatter plot quantify the strength of relationship between natural land and lake ‘stability’ values. The R^2 values provided in Figures 18 and 19 should not be considered a measure of how effective the entire model is at predicting lake performance.

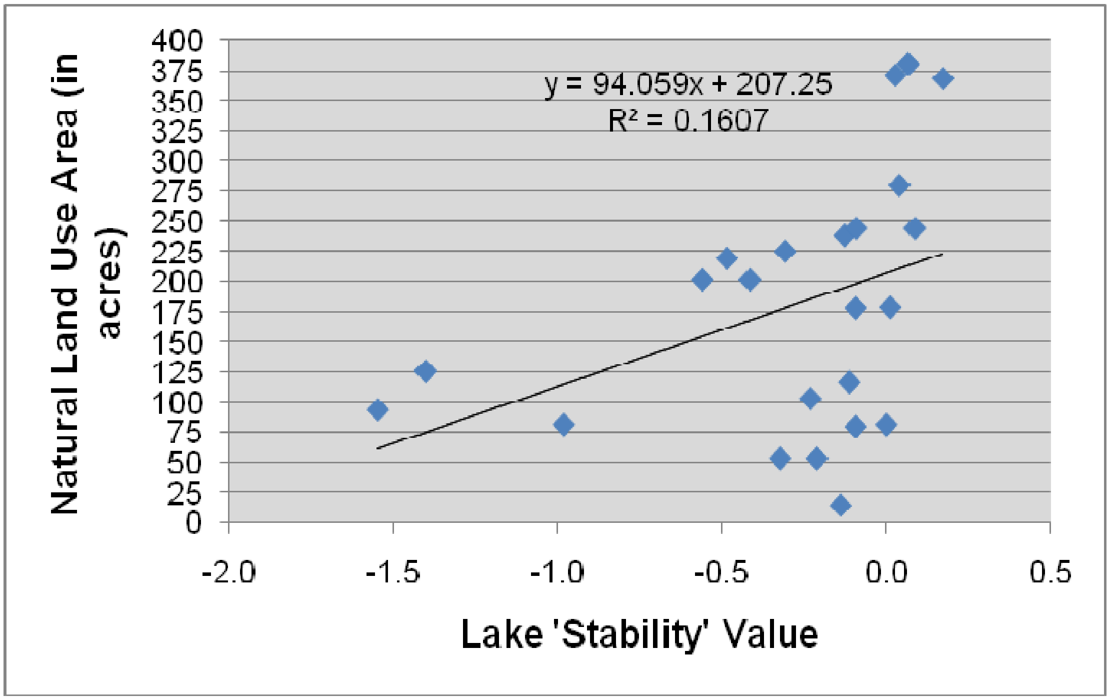


Figure 18. Natural Land Use Area Versus Additive Lake ‘Stability’ Value

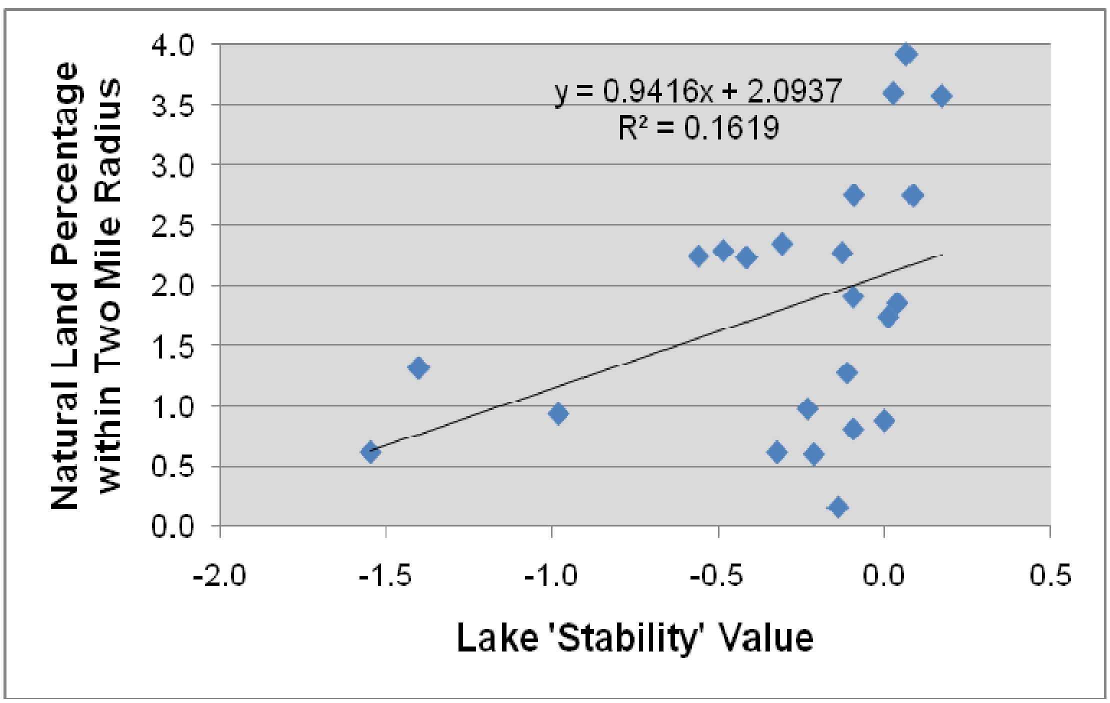


Figure 19. Natural Land Percentage Versus Additive Lake ‘Stability’ Value

In the Additive model performed for this thesis, six separate lakes claimed an optimum performance rating of one, which is similar to the BCC-I model. Lake Alice, Lake Brant, Cypress Lake, Hanna Lake, Keystone Lake, and Thomas Lake all achieved an optimum performance rating of one. The ‘stability’ values corresponding to each of these lakes allowed them to be further ranked beyond the performance rating of one. According to the ‘stability’ values, Lake Brant achieved the highest performance efficiency followed in order by Cypress Lake, Hanna Lake, Keystone Lake, Thomas Lake, and Lake Alice. Lake Brant received the greatest ‘stability’ value at 0.1713. Therefore, it can be stated that Lake Brant performed at the highest level of efficiency in the Additive model. Lake Brant was located within a two mile radius that contained 368.4807 acres of naturally preserved land translating into 3.5855% of Lake Brant’s two mile radius delineation. As anticipated, input variable concentrations for Lake Brant were low relative to the nutrient loads of other lakes included in the study.

Lake Thonotosassa received the lowest ‘stability’ value amongst the set of lakes at -1.5493. This lake obtained the last ranking because it is located within a two mile radius of only 93.6222 acres of naturally preserved land. Naturally preserved land only occupied 0.6194% of Lake Thonotosassa’s two mile radius delineation. Lake Thonotosassa’s ‘stability’ value was the lowest amongst the entire set of lakes partially due to its input variable concentrations. Each input variable concentration was high relative to the nutrient loads of other lakes included in the study. In fact, Lake Thonotosassa consistently contained the second highest concentration of each type of input variable. Input variable concentrations for Lake Thonotosassa negatively impacted its water quality performance and subsequently reduced its ‘stability’ value.

Directly in the middle of the performance ranks, Pretty Lake obtained a rating of 0.7345 and a ‘stability’ value of -0.1293. This lake was representative of the average performance for the entire data set. As previously stated, the average performance rating for the entire set of lakes was equal to 0.6333. Also, the average ‘stability’ value for the entire set of lakes was equal to -0.2933. Input and output variable data for Pretty Lake was also representative of the averages calculated for the entire data set. Pretty Lake was located within a two mile radius containing 238.1067 acres of natural land. On average, the two mile radius delineations established during this study contained 179.6641 acres of natural land. The standard deviation for natural land area was equal to 103.8114 acres. Therefore, the natural land area surrounding Pretty Lake was representative of the entire data set because it fell within one standard deviation of the average. Pretty Lake was located in a two mile radius delineation occupied by 2.2794% natural land use. On average, the two mile radius delineations established in this study were occupied by 1.8175% natural land use. The standard deviation for natural land use percent surrounding the study lakes was equal to 1.0354%. Therefore, it can be stated that the percentage of natural land surrounding Pretty Lake was representative of the entire data set because it fell within one standard deviation of the average. The total chlorophyll concentration for Pretty Lake equaled 10.00 ug/L, while the average concentration of total chlorophyll for the entire data set was equal to 26.1870 ug/L. Pretty Lake contained 14.00 ug/L of total phosphorous, while the average concentration of total phosphorous for the entire data set was equal to 41.1043 ug/L. Finally, Pretty Lake recorded a 695.00 ug/L concentration of total nitrogen, which was somewhat comparable to the total nitrogen concentration average of 1,017.1217 ug/L calculated for the entire data set.

The Additive model is particularly useful because it generates a measurement known as a 'stability' value. This measurement pertains to the individual efficiency of a DMU. The 'stability' value is a more accurate measurement of efficiency than the traditional DEA performance rating. DMU efficiency is directly measured by the 'stability' value. The 'stability' value is a useful measurement because it can be used to further classify the efficiency of individual DMUs beyond their DEA performance ratings. For this set of lake data, it was observed that in some instances the 'stability' value rankings did not directly correspond to the rankings that would have been determined by the DEA performance ratings. For example, simply referring to the traditional performance ratings would have led one to rank Sunset Lake at tenth in overall efficiency, however, the 'stability' value unique to the Additive model ranked Sunset Lake at seventh in overall efficiency. This discrepancy between the traditional DEA performance rating and 'stability' value is due to the mathematical platform of the Additive model. Values for 'stability' are derived from a variation of the equations used to produce the traditional performance ratings. Therefore, discrepancies between the two forms of measurement occur in rare instances.

Comparing DEA Model Results to Trophic State Index

According to the Hillsborough County Water Atlas website, the Trophic State Index (TSI) assigns quantitative ratings to individual lakes based on measurements of biological productivity. TSI ratings are derived from data for total chlorophyll, total phosphorous, and total nitrogen. Essentially, TSI rates individual lakes according to nutrient loads that contribute to eutrophic conditions. TSI lake ratings are specifically derived from measurements for total phosphorous and chlorophyll-A concentrations along with Secchi depth. The TSI functions as a classification system that evaluates a lake based upon its nutrient loads. TSI measurements specifically focus on nutrient loads while rating the quality of water within individual lakes. The TSI uses a numeric scale from one to one hundred to express the quality of lake water. Lower values along the scale from one to one hundred are equivalent to lower nutrient loads and environmentally beneficial lake water quality. Higher values along the scale from one to one hundred are equivalent to higher nutrient loads that contribute to environmentally harmful lake water quality. Aquatic conditions in lakes that receive lower TSI measurements are more environmentally beneficial. Higher TSI measurements reflect aquatic conditions that are not beneficial to naturally functioning ecologic systems. Table 15 displays the TSI measurements for each of the twenty-three lakes studied during this thesis. The TSI measurements for each of the study lakes were retrieved from the Hillsborough County Water Atlas website. The table below also displays study lake performance ratings derived from each of the DEA models applied during the study.

Table 15. Lake Performance Ratings and TSI Measurements

Lake Name	CCR-I Performance Rating	BCC-I Performance Rating	Additive Performance Rating	Trophic State Index (TSI)
Alice, Lake*	0.7500	1.0000	1.0000	33
Armistead, Lake	0.4142	0.5666	0.7009	41
Brant, Lake*	0.8086	1.0000	1.0000	53
Burrell Lake	0.0895	0.4003	0.4462	53
Chapman Lake	0.0220	0.3720	0.4910	45
Cypress Lake*	1.0000	1.0000	1.0000	27
Echo Lake	0.2706	0.5676	0.6377	42
Flynn Lake	0.0972	0.4459	0.4955	51
Garden Lake	0.0616	0.1853	0.2803	73
Hanna Lake*	0.3879	1.0000	1.0000	61
Harvey, Lake	0.2323	0.2814	0.4318	63
Hiawatha, Lake	0.2552	0.6049	0.5748	53
James, Lake	0.5600	0.5600	0.8014	40
Josephine Lake	0.3560	0.3865	0.4991	58
Keystone Lake*	1.0000	1.0000	1.0000	35
Osceola, Lake	0.2023	0.6195	0.6203	38
Pretty Lake	0.5253	0.5388	0.7345	58
Rock Lake	0.2842	0.3157	0.4713	60
Sunset Lake	0.3331	1.0000	0.6463	27
Thomas Lake*	0.7936	1.0000	1.0000	51
Thonotosassa, Lake	0.0606	0.1540	0.1715	82
Virginia, Lake	0.2160	0.2618	0.4275	64
Weeks, Lake	0.0709	0.1345	0.1356	85

* - indicates efficient DMUs according to both BCC-I and Additive models

DEA performance ratings range from zero to one. Theoretically, higher ratings along this scale are indicative of environmentally beneficial lake water quality. Performance ratings closer to zero theoretically signify that a lake contains higher nutrient loads that contribute to eutrophic conditions. Lakes that obtained performance ratings closer to one should contain lower nutrient loads that fail to establish eutrophic conditions. Therefore, higher DEA performance ratings should correspond to lower TSI measurements. DEA performance ratings in this study should share an indirect relationship with TSI measurements. Lakes that obtained a higher performance rating should have received a lower TSI measurement. The performance ratings derived from each of the DEA models used during this thesis are compared to TSI in Figures 20, 21, and 22 displayed below. These scatter plots represent the strength of relationship between DEA performance ratings and TSI. The statistical lines of best-fit along with their associated R^2 value are provided for each of these scatter plots. The relationship between DEA performance ratings and TSI is depicted by the statistical lines of best-fit and R^2. The best-fit lines support the trend that DEA performance ratings correspond appropriately to TSI measurements. After observing this trend, it can be stated that DEA lake performance ratings shared an indirect relationship with TSI. DEA performance ratings derived during this thesis correspond appropriately to study lake TSI. Therefore, it can be stated that a majority of the study lakes received performance ratings reflecting actual lake water quality conditions as described by TSI measurements. This statement is appropriate for performance ratings derived from each of the models used during this thesis. Overall, performance ratings derived from each of the DEA models provide an accurate representation of lake water quality. By comparing DEA performance ratings to TSI measurements, it was verified that results from each of the models accurately describe lake water quality conditions.

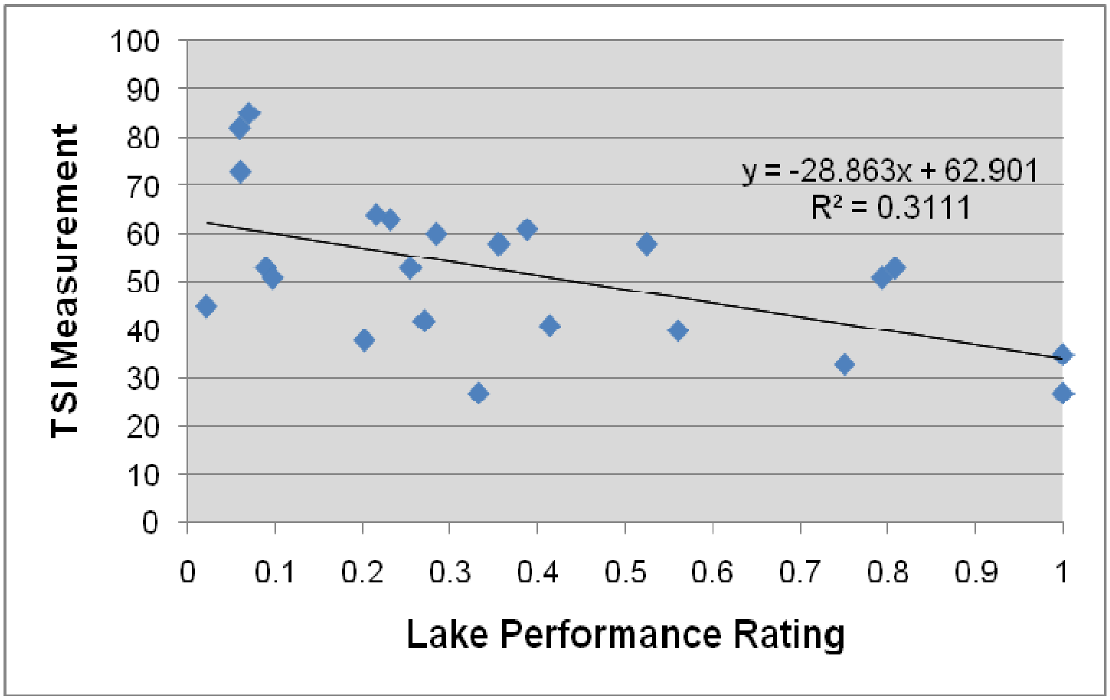


Figure 20. CCR-I Performance Rating Versus TSI

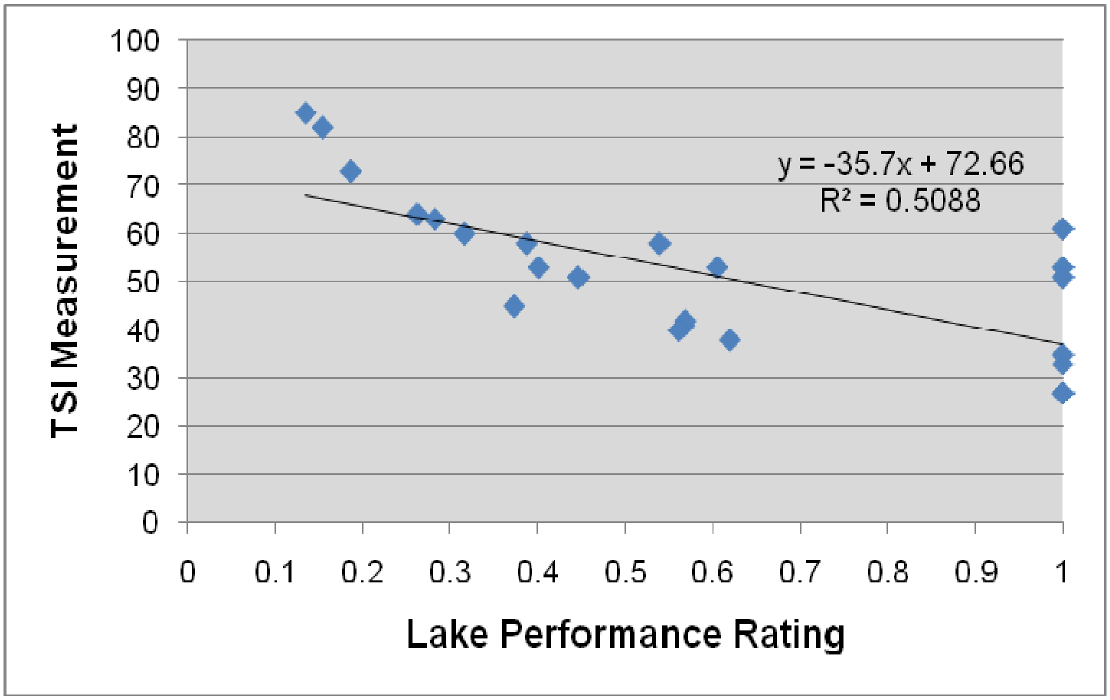


Figure 21. BCC-I Performance Rating Versus TSI

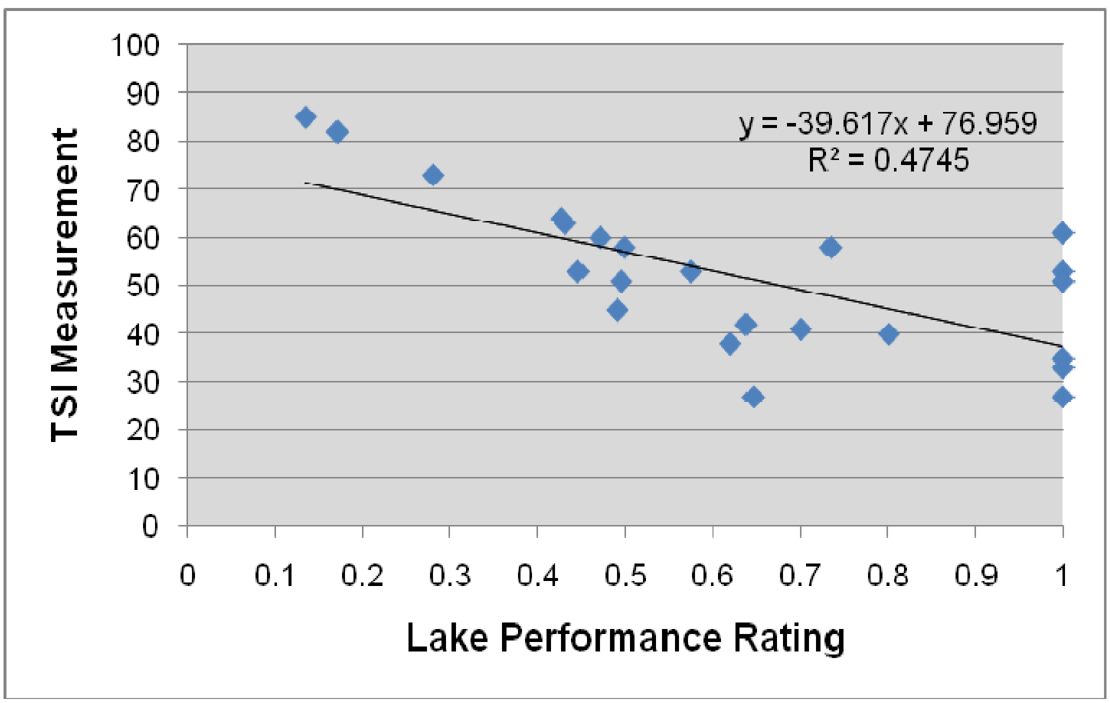


Figure 22. Additive Performance Rating Versus TSI

Discussion

Input and output variables influenced the performance ratings derived from the DEA. Differences in performance ratings for each lake reflect changes in input and output variable data. DEA performance ratings generated during this study fluctuate according to trends in the input and output variable data. Minimizing input data and maximizing the single output variable resulted in higher performance ratings. Therefore, lake water quality was optimized by minimizing total chlorophyll, total phosphorous, and total nitrogen, while maximizing natural land area within a two mile radius. This general trend is exposed by Figures 5 through 19 comparing the individual variables to each lake's performance rating.

DEA performance ratings reflect lake water quality conditions. Performance ratings provided a numerical scale for quantifying the level of lake water quality optimization. In this case, higher performance ratings represented higher levels of lake water quality optimization. The interaction between nutrient loads and natural land uses determined the level of lake water quality optimization.

Results from the DEA revealed notable trends that describe the relationship between land use and surface water quality of lakes in Hillsborough County. Lakes located within a two mile radius containing higher amounts of natural land area typically received a higher performance rating. Figures 8, 9, 13, 14, 18, and 19 graphically support this trend in the relationship between lake water quality performance and natural land use area. An optimum performance rating of one was achieved by only two separate DMUs, Cypress Lake and Keystone Lake, in the CCR-I model. Keystone Lake and Cypress Lake were located in two mile radius delineations containing the fourth and fifth highest amount of natural land area, respectively. This statistic undoubtedly contributed to their optimum performance ratings. An optimum performance rating of one was achieved by seven DMUs, Lake Alice, Lake Brant, Cypress Lake, Hanna Lake, Keystone Lake, Sunset Lake, and Thomas Lake, in the BCC-I model. Upon further scrutiny of the BCC-I model results, the 'slack' measurements indicated that Sunset Lake was not in fact operating at optimum efficiency. Therefore, it should be stated that only six lakes were optimized by the BCC-I model. With the exception of Lake Alice, optimally performing lakes were located in advantageous two mile radius delineations that consistently contained at least 64.4867 acres above the average for the entire data set. In fact, five of the six remaining optimally performing lakes were surrounded by the five highest measurements for natural land area. Hanna Lake, Thomas Lake, Lake Brant, Keystone Lake, and Cypress Lake represent the five optimally performing lakes that were situated within two mile radius delineations of the first, second, third, fourth, and fifth highest amounts of natural land area, respectively. When focusing attention on these five particular lakes, it became especially evident that natural land uses positively influenced water quality performance ratings. An optimum performance rating of one was obtained by six DMUs, Lake Alice, Lake

Brant, Cypress Lake, Hanna Lake, Keystone Lake, and Thomas Lake, in the Additive model. With the exception of Lake Alice, optimally performing lakes were located in advantageous two mile radius delineations that consistently contained at least 64.4867 acres above the average for the entire data set. The reader may notice that the same six lakes isolated by the Additive model were also specifically identified during the BCC-I model discussion above. When focusing attention on these six particular lakes, it became especially evident that natural land uses positively influenced water quality performance ratings. The Additive model further corroborates that optimum performance ratings were significantly dependent upon the amount of natural land surrounding a lake. Just as the other two input oriented models have indicated, the Additive model confirms that natural land uses possess a positive influence on lake water quality performance.

The trend discussed in the previous paragraph indicates that a designated land cover composed of a greater percentage of natural use area will generally contain lakes with lower nutrient load concentrations. Therefore, nutrient loading within lakes shares an indirect relationship with natural land use area. Each of the DEA models applied in this thesis supported the assumption that natural lands protect lakes from environmentally harmful nutrient loads. Natural lands assimilate soluble pollutants that would otherwise be directly deposited into aquatic ecosystems. In this manner, natural lands function as a pollution filtration mechanism for freshwater resources. This concept represents an area of potential future scientific research. Future studies could focus on determining how spatial distribution and habitat type affect the potential for natural land to protect against lake nutrient loading.

Lakes containing lower concentrations of total chlorophyll, total nitrogen, and total phosphorous typically received higher performance ratings. This significant trend revealed an indirect relationship between nutrient loads and lake water quality optimization. Lake water quality was optimized when nutrient concentrations were minimized. Lakes with elevated nutrient concentrations typically occurred in two mile radius delineations containing lower percentages of natural land area. This significant trend is explained by the nutrient filtration properties of natural land types (Reddy and Dev 2006; Osborne and Wiley 1988; Lenat and Crawford 1994). Natural land types assimilate nutrients that would otherwise be deposited in hydrologically connected surface waters such as lakes (Reddy and Dev 2006; Osborne and Wiley 1988; Lenat and Crawford 1994). Results from the individual DEA models established a significant relationship between natural land use and aquatic nutrient contaminants. Typically, nutrient concentrations diminished in those lakes surrounded by a higher percentage of natural land use area.

The interaction between nutrient concentrations and natural land use area was the focus of each DEA model applied during this thesis. Results from the DEA models revealed that natural land use area improves the performance of lakes and diminishes the presence of soluble nutrients. 'Projection' figures generated by the CCR-I model indicated that the input variables for each lake except Cypress Lake and Keystone Lake required reductions to obtain optimal water quality performance. Non-coincidentally, Cypress Lake and Keystone Lake were the only lakes to achieve an optimal water quality performance rating of one. Therefore, the ‘projection’ figures for both these lakes reflected that no change in the input variables were necessary to accomplish optimal efficiency.

With the exception of both Cypress Lake and Keystone Lake, each lake included in the study possessed inflated input variable data according to the ‘slack’ measurements returned by both the CCR-I and BCC-I models. Cypress Lake and Keystone Lake were the only DMUs to return all ‘slack’ measurements equal to zero for both input oriented models. The only two lakes to receive all ‘slack’ measurement equal to zero for both the CCR-I and BCC-I models were Cypress and Keystone. Also, each DMU except for both Cypress Lake and Keystone Lake was classified as inefficient by at least one of the models analyzed during this thesis. Therefore, it can be stated that each lake except for Cypress Lake and Keystone Lake contained excesses in nutrient concentrations that prohibited optimal performance ratings across all three of the DEA models. To achieve optimal performance for these failing lakes, nutrient concentrations would have to be reduced according to the value indicated by the ‘slack’ measurement produced in the CCR-I model. Nutrient load augmentations should obey the ‘slack’ measurements returned by the CCR-I model because it produced a more stringent ‘efficient frontier’ line than the BCC-I model. Therefore, adjusting nutrient concentrations according to the ‘slack’ measurements returned by the CCR-I model would improve lake water quality most drastically. For the most drastic improvement in lake water quality, the CCR-I ‘slack’ measurements should be consulted. For a detailed account of ‘slack’ measurements computed by both these models, the reader should refer to the ‘Results’ section of this thesis. The ‘slack’ measurements corresponding to each input as well as output variable can be viewed in Tables 9 and 12. Table 9 provides the ‘slack’ measurements related to the CCR-I model, while Table 12 reveals the ‘slack’ measurements produced by the BCC-I model.

Reducing the input variables according to the amount indicated by the ‘slack’ measurement would result in each underperforming lake obtaining an optimal performance rating. This augmentation in lake nutrient concentration could be accomplished by preserving natural buffer areas surrounding lakes as well as increasing the overall acreage of natural land within an entire two mile radius surrounding a lake (Tong and Chen 2002; Castelletti and Soncini-Sessa 2007; Osborne and Wiley 1988; Lee 2002). In this aspect, lake nutrient concentrations are functionally dependent upon the surrounding land uses (Reddy and Dev 2006; Osborne and Wiley 1988; Allan 2004). As this study and others have revealed, nutrient concentrations share an indirect relationship with the spatial quantity of natural land use surrounding a lake (Reddy and Dev 2006; Griffith et al. 2002; Allan 2004). Nutrient concentrations typically decline in lakes surrounded by greater proportions of naturally preserved land (Tong and Chen

2002; Lenat and Crawford 1994; Griffith et al. 2002; Allan 2004). In the DEA model developed for this study, every lake with the exception of two, Cypress Lake and Keystone Lake, contained elevated nutrient concentrations that required reductions to achieve optimal water quality performance.

The correlation matrices provided as Tables 10 and 13 contain proportions that quantify the significance of the relationship between natural land use and lake nutrient concentrations. Table 10 contains the correlation proportions produced during the CCR-I model, while Table 13 contains the correlation proportions produced during the BCC-I model. The values within both tables are identical. The 'DEAlytics' software used to compute the Additive model did not produce correlation proportions, however, it can be assumed that the correlation proportions for the Additive model remained the same as those in the CCR-I and BCC-I models. Therefore, the following analysis of correlation proportions applies to all three DEA models referred to during this thesis. The correlation proportion representing the relationship between natural land use and total phosphorous was equal to 0.2007, or 20.07%. This figure can be interpreted by stating that approximately 20.07% of the variable data for total phosphorous can be explained by natural land use area. The correlation proportion representing the relationship between natural land use and total chlorophyll was equal to 0.1188, or 11.88%. This figure can be interpreted by stating that approximately 11.88% of the variable data for total chlorophyll can be explained by natural land use area. Finally, the correlation proportion representing the relationship between natural land use and total nitrogen was equal to 0.1982, or 19.82%.

This value can be interpreted by stating that approximately 19.82% of the variable data for total nitrogen can be explained by natural land use area.

The correlation percentages provided in the above paragraph establish a notable relationship between natural land use and lake nutrient concentrations. While these correlation percentages do not appear impressive initially, further evaluation of the model and its variable data reveals that correlation percentages for the relationship between natural land use and nutrient concentrations have been limited by various factors. Correlation between natural land use and nutrient concentrations is limited due to the restricted spatial distribution of naturally preserved land in Hillsborough County. If the county contained a more balanced distribution of natural and built-up land uses, the correlation percentages would be capable of more accurately representing the relationship between natural land use and lake nutrient concentrations. Also, the deposition of soluble nutrients into lakes is dictated by an intricate system of hydrologic exchanges (Lee 2002; Sacks et al. 1998; Sacks 2002). This system is composed of hydrologic sinks and sources that control the movement of soluble nutrients (Lee 2002; Sacks et al. 1998; Sacks 2002). When considering the intricacy of this system and amount of potential sinks and sources involved, the correlation percentages provided by the model appear a great deal more significant.

Natural land uses as sinks explain anywhere between 11.88% and 20.07% of nutrient deposition in Hillsborough County lakes. Given the multitude of possible sinks and sources in the hydrologic exchange system of soluble nutrients, the range of correlation percentages depicting the significance of the relationship between natural land use and lake nutrient loads becomes more impressive than previously perceived at first glance. According to the DEA model, the functioning of natural lands as hydrologic sinks influences lake nutrient loads between 11.88% and 20.07% in Hillsborough County. This is a significant figure when considering the intricacy of the system that dictates soluble nutrient deposition into surface water lakes.

Completion of this study exposed both disadvantages and advantages associated with applying DEA to accomplish the research objectives of this thesis. A disadvantage with the DEA developed during this study was that the model failed to include a variable representing a potential source of nutrient contamination. Within the model devised for this study, the single output variable only represents a potential net sink for nutrient loads in the form of natural land use area. Unfortunately, this disadvantage could not be addressed because all remaining land uses within Hillsborough County have been traditionally classified as net nutrient contaminant sources (Tong and Chen 2002; Lenat and Crawford 1994; Osborne and Wiley 1988). Therefore, it would not have been prudent to include the remaining area of each lake's two mile radius delineation as a potential nutrient load source. Also, a variable such as this would not have been compatible with the overall scheme of the model, which sought to generate optimal water quality performance ratings by minimizing the inputs and maximizing the outputs. A variable representing a net nutrient source could not have been included as an input because this DEA's inputs were restricted to nutrient contaminants. Simply put, the variable for land uses that act as net nutrient sources could not be classified as an input variable because it is not a type of nutrient contaminant. Conversely, this variable could not be included as an output variable because outputs were maximized when generating optimal performance ratings. In theory, the variable for land uses that act as net nutrient sources would have to be minimized to improve water quality performance. Therefore, this variable could not be classified as an output.

Another disadvantage encountered by the models used during this study was due to the study area in which the model was applied. Natural land area has been greatly diminished in the predominantly urbanized county examined by the model. The limited spatial distribution of natural lands in Hillsborough County restricted the model's applicability. Numerous lakes within Hillsborough County failed to contain any natural land area within a two mile radius. Due to the mathematical framework of DEA based on production ratios, lakes contained by a two mile radius without any natural lands could not be examined by the model because they failed to fit on the 'efficient frontier' line. Performance ratings for lakes within these two mile radius delineations would have automatically been zero, which would not have accurately represented the lake's water quality performance. Therefore, lakes within two mile radius delineations that did not contain any natural land could not be included in the model. Essentially, variable data collected for the model had to be positive, non-zero numbers for the purpose of generating relevant performance ratings capable of being interpreted. Due to the lack of natural land use distribution in Hillsborough County, certain lakes located within highly urbanized portions of the County could not be included in the model. The model devised for this study should only be applied to study areas that are known to contain some form of natural land use. This application limitation represents a disadvantage associated with the model.

An additional disadvantage related to this application of DEA in assessing the relationship between lake performance and land use was that the models failed to examine how much of the nutrient loads were received from external sources. Lake nutrient loading can occur through internal processes independent of external nutrient sources. Therefore, it would be useful to devise a DEA model that distinguishes between internal and external nutrient loads. This could be accomplished using a DEA model that incorporates categorical input variables representing internal and external loading. The categorical variables would function to distinguish nutrient loads originating from either an external or internal source. By designing the model such as this, the DEA would be able to examine the impact on lake water quality from either external or internal nutrient deposits. For the study area of this thesis, it was not possible to make this distinction because there was no available data concerning internal and external nutrient loads for each individual lake examined by the DEA. Given the appropriate internal and external nutrient load data, it would be possible to devise a DEA that distinguishes between the two types of nutrient deposits. This could be accomplished by creating separate input variables for internally and externally deposited nutrient concentrations.

The final disadvantage of this DEA application was that the analysis failed to include a spatial component. The spatial distribution of naturally preserved land was not considered by the DEA constructed for this thesis. The DEA applied in this thesis simply evaluated the overall amount of natural land area located within a two mile radius from each study lake. It ignored the spatial distribution of net nutrient sinks within a two mile radius of study lakes. This can be considered a disadvantage because the spatial distribution of a specific land use has been linked to the quality of freshwater resources in previous studies (Lee 2002; Griffith et al. 2002). The spatial distribution of particular land uses likely influence the overall surface water quality of a lake (Lee 2002; Griffith et al. 2002). Therefore, it is a disadvantage that this DEA application neglected to consider the spatial distribution of natural land when attempting to describe the relationship between lake water quality and land use.

Certain advantages associated with the model were identified after completing the applied research component of this thesis. The model examined the impact of natural land use area on lake nutrient loading. The design of the model successfully isolated the variable for natural land use area to evaluate its correlation with nutrient contaminant concentrations. Correlation figures produced by the model described the significance of the relationship between natural land use area and lake nutrient loading. Also, the model could be easily transformed to evaluate the significance of other land coverage types that typically function as net sinks for soluble nutrients. Any land type that should be maximized for the purpose of establishing optimal water quality conditions could be included as an output variable in the model. For instance, the output variable could consist of land use area data for wetland habitats within a two mile radius of each study lake.

Model inputs could also be replaced in similar fashion. Any type of aquatic contaminant that should be minimized for the purpose of establishing optimal water quality conditions could be integrated as an input variable in the model. In theory, this means that the model could be used to establish the significance of the relationship between any water contaminant and any land type classified as a net sink of soluble materials. For instance, the input variables could consist of aquatic contaminant data for arsenic concentrations in a lake.

Another advantage of the model was its simplistic design. Data generated by the model was readily interpreted for the purpose of describing the relationship between natural land use and lake nutrient loading. Variables included in the model were classified as inputs or outputs according to their typically observed impact on lake water quality. Aquatic contaminant variables that required minimization to achieve optimal water quality performance were designated as inputs, while the single natural land use variable that required maximization was categorized as an output. This division of variables enabled the model to produce data that specifically examined the relationship between water quality and natural land use. The model generated data that directly measured the correlation between natural land use and nutrient loading in Hillsborough County.

The final advantage of the model is that it can be readily applied to other study areas. This advantage is contingent upon the land uses examined by the model as well as the spatial distribution of those land uses within the study area. As previously discussed, variable data for the model must be positive or non-zero to produce production ratings capable of being interpreted. Therefore, the model can only be transposed to a study area that contains the particular land use of interest. This contingency is the only limitation related to applying the DEA model developed for this thesis to other study areas. The advantage of being able to transpose the model to other study areas answers one of the overall research questions addressed by this thesis (refer to the 'Research Design' section of this document).

The DEA model completed during this thesis generated data that was interpreted to describe the significance of the relationship between natural land use and lake nutrient concentrations in Hillsborough County. In doing so, the discussion produced by this thesis contributed previously uncovered knowledge regarding the relationship between land use and lake water quality. Overall, the research discovered that a moderately significant positive correlation exists between natural land use area and lake water quality performance. Also, it was confirmed that indicators of soluble nutrient pollution such as total chlorophyll, total phosphorous, and total nitrogen detract from the performance of lake water quality. Finally, the research conducted for this thesis reaffirmed that natural land uses enhance lake water quality (Tong and Chen 2002; Lenat and Crawford 1994; Osborne and Wiley 1988).

Recommendations and Conclusion

The research questions listed below were posed in the 'Research Design' section of this thesis.

1. After analyzing the various forms of scientifically acceptable data using DEA computer software, does naturally preserved land typically contribute to a water quality benchmark optimizing lake performance?
2. How can water managers operating within the boundaries of Hillsborough County reproduce the necessary conditions to achieve an optimal water quality benchmark?
3. Short of altering the current land use surrounding a given lake through land acquisition techniques, how can localized water managers improve management techniques to achieve an ecologically optimal water quality benchmark?
4. After performing the necessary analysis, will the devised methodology be easily transposable to other study areas?

These questions were all directly answered during the research conducted for this thesis. After analyzing the selected input and output variables with DEA, naturally preserved land was identified as a positive contributor to water quality. The methodology performed during this thesis indicated that natural land enhances lake water quality and ecologic performance. These findings should persuade water managers to preserve or rehabilitate natural land whenever feasible. When natural land preservation is not feasible, water managers should implement water quality BMPs as well as efforts to artificially simulate the nutrient filtration properties of natural land. Finally, it was discovered that the methodology conducted during this thesis would be readily transferable to other study areas given the required data sets.

As with any applied research effort, certain hypotheses are identified prior to conducting the analysis. In this case, it was expected that lakes surrounded by a higher proportion of natural land would receive higher DEA performance ratings. This result was expected because it has been previously documented that the surface area of natural land surrounding a lake shares a direct relationship with the performance of that lake (Reddy and Dev 2006; Wescoat and White 2003; Gleick et al. 2006). Meaning, the performance of a lake declines as the natural lands surrounding that lake are removed. The DEA performed for this thesis supported the relationship between lake water quality performance and natural land area mentioned in the previous sentence. A significant trend was established by the DEA in which DMUs surrounded by a greater amount of natural land area typically received higher performance ratings. It was also hypothesized that results generated from the DEA would support water management strategies focused on preserving natural lands and rehabilitating impaired natural lands. This prediction was supported by the DEA model because lakes surrounded by a greater proportion of natural land typically received higher performance ratings. Therefore, the preservation and rehabilitation of natural land would likely enhance lake water quality performance.

Another hypothesis predicted that water management efforts based on BMPs designed to artificially simulate the pollutant filtration function of naturally preserved land would be supported by the results of the DEA. This hypothesis was indeed supported by the DEA because natural lands improved lake water quality performance by functioning as a net sink for nutrient loads. In this manner, management efforts proven to simulate the assimilative qualities of natural lands would also function to enhance lake water quality. Assumedly, these water management techniques would provide net sinks for nutrient loading thereby improving the quality of water in nearby lakes. Finally, it was hypothesized that the methodology developed for this thesis would be readily transferable to other study areas that collect and store the required datasets. After performing the methodology developed for this thesis, it is evident that the model could be readily transposed to any study area containing the appropriate datasets. Cities or counties concerned with lake nutrient loading due to land use could potentially refer to the model developed during this thesis.

Interpretations of the results from the three DEA models revealed notable trends between natural land use and lake water quality performance. These interpretations contribute information that supports specific water management actions. In general, water management recommendations based on these interpretations would encourage the preservation of natural lands whenever possible and especially surrounding those lakes that received lower DEA performance ratings.

Lakes that achieved lower performance ratings should be the initial focus of any water management recommendations derived from the DEA conducted for this study. Lakes performing at a lower level represent situations in which the most improvement to water quality can be accomplished. Conditions should be maintained for those lakes determined to be functioning at an optimum level. However, lakes that did not achieve optimum performance should be subjected to water management actions that reduce nutrient loading and increase the positive impacts from naturally preserved land. In a highly urbanized setting such as Hillsborough County, the economic motivation to develop residential, commercial, and industrial facilities may often create situations in which it is not feasible to preserve natural lands. Natural land preservation in the study area of this thesis frequently fails due to the economic pressures of development. Therefore, it may be more prudent to rely upon BMPs that protect water resources from harmful pollutants.

Results from the DEA suggest that rehabilitation to restore the assimilative properties of natural land would be an effective water management tactic when attempting to improve lake water quality. In frequent instances when this is not possible, improvements to lake water quality in the study area could also be realized if BMPs were implemented that simulated the filtration function of natural lands. After interpreting the results provided by the DEA, it was apparent that water managers should attempt to improve the permeability of land surfaces as a preventive measure to reduce soluble nutrient deposition in Hillsborough County lakes. Land surface permeability is improved by removing impenetrable surfaces such as concrete. Exact measures to improve land surface permeability can be accomplished by complying with environmentally conscious construction practices. Land surface permeability is improved by any technique that reduces both water run-off and the use of impermeable concrete materials. Increased natural land area was discovered to improve lake water quality performance. The permeability of natural land is typically higher than that of built-up land uses (Tong and Chen 2002). This physical property of natural land contributes to its ability to assimilate nutrient loads thereby protecting freshwater bodies from eutrophication (Tong and Chen 2002). Therefore, completing the necessary measures to improve land surface permeability within a designated area surrounding a lake would contribute to enhanced water quality and a reduction in soluble nutrient concentrations.

Lake water quality improvement is typically witnessed on seasonal scales (Castelletti and Soncini-Sessa 2007; Reddy and Dev 2006). When actions are conducted to enhance lake water quality, improvements are typically observable after the passage of a wet season (Castelletti and Soncini-Sessa 2007; Reddy and Dev 2006). After an attempt at improving lake water quality, the time necessary to witness any positive impacts is highly dependent upon the targeted water resource and the type of management action used to improve the quality of the lake's water.

In the DEA model, lakes that contained lower concentrations of soluble nutrients obtained higher performance ratings. This would indicate that regulations limiting the dispersal of soluble nutrients would improve the performance of lakes within the study area. Water management efforts to lower nutrient loading would improve the water quality performance of area lakes. Efforts to do so might include regulations that mandate a reduction in residential, agricultural, and commercial fertilizer use. Water managers could also require specific fertilizer techniques or products that typically generate lesser volumes of soluble nutrient run-off. Any fertilizer regulations supported by the results of this DEA would enforce BMPs that protect freshwater resources. Results from the model support previously enacted BMPs that have been proven to reduce soluble nutrient deposition in freshwater resources. Such BMPs may include natural buffer areas surrounding surface waters, fertilizer techniques that encourage application only during the growing season, and restrictions that prohibit fertilizer application within a designated proximity of an aquatic ecosystem.

The literature review and DEA application conducted for this thesis revealed additional research opportunities that would extend scientific understanding of how land use and lake water quality interact. These research opportunities would focus on determining the significance of the relationship between land use and lake water quality through the application of DEA. Complimentary studies would attempt to determine the significance of the relationship between lake water quality and various built-up land uses through DEA. Also, future research should examine the influence of spatial distribution on the relationship between land use and lake water quality. The spatial distribution of particular land uses surrounding a lake likely influence its water quality (Lee 2002; Griffith et al. 2002). Therefore, future studies should focus on determining the significance of naturally preserved buffer areas on lake water quality. Additionally, studies of this nature would also have to examine how built-up areas directly surrounding lakes influence water quality. Future studies should apply DEA to itemize the impact from each land use type contained in a study area on lake water quality. This all inclusive model could be developed given the appropriate data. Such a study would assist water managers in identifying specific land uses that negatively or positively impact lake water quality. Potential results from this study would allow water managers to devise strategies that either negate or enhance the influence of particular land uses on lake water quality. Future DEA modeling could incorporate water quality data collected during different time periods. The water quality performance of individual lakes could then be compared between the different time periods examined by the DEA. This would allow water managers to monitor the water quality performance of an individual lake during different time periods. Finally, future applications of DEA in describing the relationship between lake water quality and land use could examine a variety of other types of input variables such as Total Suspended Solids (TSS), Total Suspended Volatile Solids (TSVS), and other forms of aquatic contaminants. Incorporating these different types of input variables will broaden the scope of future research concerning the application of DEA in studying the relationship between surface water quality and land use.

Of the three different DEA models applied during this study, the CCR-I model would be the most effective for supporting drastic management efforts to protect or improve lake water quality. According to the CCR-I model, only two lakes achieved optimum performance, whereas, six lakes achieved optimum performance using the other two models. This observation alone indicates that the CCR-I model is less lenient than the other two models. Therefore, the CCR-I model should be used to support stronger measures aimed at improving lake water quality. In this sense, the CCR-I model represents the most effective option for instituting change in water management emphasizing the improvement of lake conditions. The more stringent rating system of the CCR-I model would encourage preventive water management actions reducing the likelihood of lake impairment. From a water management perspective, the CCR-I model represents the most stringent of the three models that could be used to justify the most protective water quality policies.

An original application of a performance assessment methodology was explored during the research conducted for this thesis. The applied research and accompanying literature review for this thesis provided previously unexplored and overdue dialogue regarding the application of DEA to water management concerns. Assessing the status of Hillsborough County lakes proved to be a valuable task in the face of growing water demands and intensifying human impacts (Poe et al. 2005). The ultimate objective of this thesis was to provide water management recommendations and interpretations based on a DEA assessment, GIS land use layers, and preexisting scientific literature. While accomplishing this primary research objective, it was discovered that DEA has been increasingly applied to natural resource performance related questions in a variety of environmental disciplines (Alsharif et al. 2008; Fraser and Hone 2001; Jaenicke and Lengnick 1999; Malana and Malano 2006; Rhodes 1986; Shafiq and Rehman 2000). From a review of the available literature, it became apparent that the application of DEA to environmental concerns is a burgeoning endeavor with vast stores of potential research yet to be conducted.

The applied research component of this thesis revealed that DEA can be effectively applied to water management issues that specifically address the relationship between land use and lake water quality. The DEA model produced during this thesis serves as a viable example of how DEA can be applied to assess the performance of lake water quality in relation to surrounding land uses. This research revealed the potential to generate additional studies based on DEA modeling techniques that assess the relationship between land use and lake water quality. Through the applied research portion of this thesis, it was determined that the performance measurement capabilities of DEA provide an effective platform for assessing land use and lake water quality interactions.

Through the use of DEA, notable trends were identified that described how water quality parameters are impacted by land use. The relationship between land use and lake water quality was effectively examined by the DEA methodology developed during this thesis.

References Cited

Allan, J.D. 2004. Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annual Review of Ecology, Evolution and Systematics* 35: 257-284.

Alsharif, K. 2008. *Assessment of the Lake Ecosystems in Hillsborough County by Using Frontier Estimation Technique*. Application for Pollution Recovery Fund Assistance.

Alsharif, K., E.H. Feroz, A. Klemer, and R. Raab. 2008. Governance of water supply systems in the Palestinian Territories: A data envelopment analysis approach to the management of water resources. *Journal of Environmental Management* 87: 80-94.

Castelletti, A. and R. Soncini-Sessa. 2007. *Topics on System Analysis and Integrated Water Resources Management*. Oxford, United Kingdom: Elsevier Ltd.

Caves, D.W., L.R. Christensen, and W.E. Diewert. 1982. The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometrica* 50 (6): 1393-1414.

Cooper, W.W., L.M. Seiford, and K. Tone. 2000. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Boston, Massachusetts: Kluwer Academic Publishers.

Feldman, D.L. 2007. *Water Policy for Sustainable Development*. Baltimore, Maryland: The John Hopkins University Press.

Feroz, E.H., R. Raab, and S. Haag. 2001. An income efficiency model approach to the economic consequences of the OSHA cotton dust regulation. *Australian Journal of Management* 26: 69-89.

Florida Department of Transportation Geographic Mapping Section. 1999. *Florida Land Use, Cover, and Forms Classification System*. Tallahassee, Florida: Florida Department of Transportation.

Fraser, I. and P. Hone. 2001. Farm-level efficiency and productivity measurement using panel data: wool production in south-west Victoria. *The Australian Journal of Agricultural and Resource Economics* 45 (2): 215-232.

Gleick, P.H., H. Cooley, D. Katz, and E. Lee. 2006. *The World's Water 2006-2007: The Biennial Report on Freshwater Resources*. Washington, D.C.: Island Press.

Griffith, J.A., E.A. Martinko, J.L. Whistler, and K.P. Price. 2002. Interrelationships among Landscapes, NDVI, and Stream Water Quality in the U.S. Central Plains. *Ecological Applications* 12 (6): 1702-1718.

Houck, O.A. 2002. *The Clean Water Act TMDL Program: Law, Policy, and Implementation*. Washington, D.C.: Environmental Law Institute.

Jaenicke, E.C. and L.L. Lengnick. 1999. A Soil-Quality Index and Its Relationship to Efficiency and Productivity Growth Measures: Two Decompositions. *American Journal of Agricultural Economics* 81 (4): 881-893.

Johnson, L. and S. Gage. 1997. Landscape approaches to the analysis of aquatic ecosystems. *Freshwater Biology* 37: 113-132.

Lee, T.M. 2002. *Factors Affecting Ground-Water Exchange and Catchment Size for Florida Lakes in Mantled Karst Terrain*. Tallahassee, Florida: United States Geological Survey.

Lenat, D.R. and J.K. Crawford. 1994. Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. *Hydrobiologia* 294 (3): 185-199.

Malana, N.M. and H.M. Malano. 2006. Benchmarking Productive Efficiency of Selected Wheat Areas in Pakistan and India Using Data Envelopment Analysis. *Irrigation and Drainage* 55: 383-394.

Mendenhall, W. and T. Sincich. 2003. *A Second Course in Statistics: Regression Analysis*. Upper Saddle River, New Jersey: Pearson Education, Inc.

O'Neill, R., C. Hunsaker, K. Jones, K. Riitters, J. Wickham, P. Schwartz, B. Jackson, and W. Baillargeon. 1997. Monitoring environmental quality at the landscape scale. *BioScience* 47: 513-519.

Osborne, L.L. and M.J. Wiley. 1988. Empirical relationships between land use/ cover and stream water quality in an agricultural watershed. *Journal of Environmental Management* 26 (1): 9-27.

Poe, A., A.J. Janicki, H. Greening. 2005. *Baywide Environmental Monitoring Report, 2002-2005*. Tampa, Florida: Tampa Bay Estuary Program.

Postel, S. 1997. *Last Oasis: Facing Water Scarcity*. New York City: W.W. Norton and Company.

Ramanathan, R. 2003. *An Introduction to Data Envelopment Analysis: A Tool for Performance Measurement*. Thousand Oaks, California: Sage Publications.

Reddy, V.R. and S.M. Dev. 2006. *Managing Water Resources: Policies, Institutions, and Technologies*. New Delhi, India: Oxford University Press.

Rhodes, E.L. 1986. An Exploratory Analysis of Variations in Performance Among U.S. National Parks. In *Measuring Efficiency: An Assessment of Data Envelopment Analysis*. ed. Silkman, R.H., 47-71. San Francisco: Jossey-Bass.

Sacks, L.A. 2002. *Estimating Ground-Water Inflow to Lakes in Central Florida Using the Isotope Mass-Balance Approach*. Tallahassee, Florida: United States Geological Survey.

Sacks, L.A., A. Swancar, and T.M. Lee. 1998. *Estimating Ground-Water Exchange with Lakes Using Water-Budget and Chemical Mass-Balance Approaches for Ten Lakes in Ridge Areas of Polk and Highlands Counties, Florida*. Tallahassee, Florida: United States Geological Survey.

Sexton, T.R. 1986. The Methodology of Data Envelopment Analysis. In *Measuring Efficiency: An Assessment of Data Envelopment Analysis*. ed. Silkman, R.H., 7-29. San Francisco: Jossey-Bass.

Shafiq, M. and T. Rehman. 2000. The extent of resource use inefficiencies in cotton production in Pakistan's Punjab: an application of Data Envelopment Analysis. *Agricultural Economics* 22: 321-330.

Sidle, R. and J. Hornbeck. 1991. Cumulative effects: a broader approach to water quality research. *Journal of Soil and Water Conservation* 46: 268-271.

Stauffer, R.E. 1991. Effects of Citrus Agriculture on Ridge Lakes in Central Florida. *Water, Air, and Soil Pollution* 59: 125-144.

Stolp, C. 1990. Strength and weaknesses of data envelopment analysis: an urban and regional perspective. *Computers, Environment, and Urban Systems* 14: 103-116.

Thanassoulis, E. 2001. *Introduction to the Theory and Application of Data Envelopment Analysis*. Boston, Massachusetts: Kluwer Academic Publishers.

The Planning Commission of Hillsborough County. 2008. Available from internet, http://www.theplanningcommission.org/.

Tong, S.T.Y. and W. Chen. 2002. Modeling the relationship between land use and surface water quality. *Journal of Environmental Management* 66: 377-393.

Wang, X. 2001. Integrating water quality management and land use planning in a watershed context. *Journal of Environmental Management* 61: 25-36.

Wescoat, J.L. and G.F. White. 2003. *Water for life: water management and environmental policy*. Cambridge, United Kingdom: Cambridge University Press.

Wiley, M., S. Kohler, and P. Seelbach. 1997. Reconciling landscape and local views of aquatic communities: lessons from Michigan trout streams. *Freshwater Biology* 37: 133-148.

Xian, G., M. Crane, and J. Su. 2007. An analysis of urban development and its environmental impact on the Tampa Bay watershed. *Journal of Environmental Management* 85: 965-976.