2006-03-24

Quantifying Chlorophyll a Content Through Remote Sensing: A Pilot Study of Utah Lake

Tiana Davis
Brigham Young University - Provo

QUANTIFYING CHLOROPHYLL A CONTENT THROUGH REMOTE SENSING: A PILOT STUDY OF UTAH LAKE

by

Tiana Davis Secor

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Geography

Brigham Young University

March 2006

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Tiana Davis Secor

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Mark W. Jackson, Chair
Date	Renee Gluch
Date	Perry J. Hardin

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Tiana Davis Secor in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Mark W. Jackson
Chair, Graduate Committee

Accepted for the Department

Perry J. Hardin
Graduate Coordinator

Accepted for the College

David B. Magleby
Dean, College of Family, Home, and Social Sciences

ABSTRACT

QUANTIFYING CHLOROPHYLL A CONTENT THROUGH REMOTE SENSING:
A PILOT STUDY OF UTAH LAKE

Tiana Davis Secor

Department of Geography

Master of Science

Utah Lake is a really large but shallow lake located in the arid environment of the Western United States. Due to a variety of factors it is listed by the Environmental Protection Agency as an "impaired water body" and must be closely monitored. Because of its large extent and shallow depth the water quality is heterogeneous and can change rapidly. This means that traditional water quality monitoring methods, which require large investments in field personnel, equipment, and water sample analysis, cannot produce a model that is truly representative of the entire water body. This thesis examines the feasibility of using remotely sensed imagery to develop a water quality monitoring system for Utah Lake that is accurate, repeatable and cost-effective. Due to the paucity of *in situ* water quality information, this is primarily a pilot study using Landsat satellite imagery collected within a 5-day window of existing *in situ* water samples measuring chlorophyll *a*. The brightness values of the imagery were regressed against the water samples to produce a model to accurately predict chlorophyll *a* concentrations across the entire lake. The results of the pilot study conclude that Landsat imagery could be a very useful monitoring tool if sufficient *in situ* data for calibration were available.

ACKNOWLEDGEMENTS

This thesis is the product of many people, not solely the author. Primarily, I would like to thank my committee chair, Dr. Mark Jackson, for the time, expertise, and patience he offered me during this research and throughout my graduate studies. I would also like to acknowledge the contributions of my committee, Dr. Hardin and Dr. Gluch. Each of them had unique knowledge and experience to share with me and I am honored to have worked so closely with them. The Division of Water Quality deserves credit for the water samples I used as well as advice many employees offered me. Finally, I would like to thank my dad, Guy Davis, for the interest he had in my research and schooling and for his continual urging to finish it. Above all, a special thanks to my optimistic husband, Jason, who is my biggest supporter and was by my side every step of the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF CHARTS

CHAPTER 1

Introduction

Introduction

The Division of Water Quality listed Utah Lake among the list of lakes on "Utah's 2004 list of impaired waters" that do not currently meet water quality standards (DWQ, 2004). There is an increasing awareness and concern for the water quality of our lakes and rivers. As populations increase, the concern will continue to grow as the waters are further contaminated and polluted. The Clean Water Act of 1972 is a step towards protecting our waters. Under this act, all impaired water bodies, streams, and rivers must be identified and steps must be taken towards restoring the quality of the water to meet the standards stated in the act.

The U.S. Environmental Protection Agency has estimated that approximately one third of the surface waters in the United States are unsafe for drinking, swimming, and fishing (Vigil, 2003). Water quality refers to the physical, chemical, and biological properties of water (Liu *et al.,* 2003). While there are several natural factors that influence water quality such as vegetation, morphology, climate and location, human activities are one of the largest contributors to water pollution. The water is polluted from directly dumping contaminants into the water (point pollution) or from non-point sources where the harmful substances result from many different sources and are carried over land to the water supply. While numerous human activities may contribute to pollution, urban development often has a significant impact. Two of the largest sources of urban non-point source pollution include: 1) matter deposited onto impervious surfaces (from vehicles, atmospheric particulates, etc.) and subsequently washed in streams and lakes, and 2) runoff from lawns characteristic of low-density suburban development. This type of pollution poses unique problems because there are countless sources which cannot be eliminated easily and it is a form of pollution that is rooted in the North American way of life. Water is a recyclable resource and as humans continually contaminate the air, ground, and ocean, such contamination affects the water quality of our lakes, streams, and rivers as shown in Figure 1.1. Regardless of the source, the water cycle spreads these contaminates throughout the earth's entire water supply.

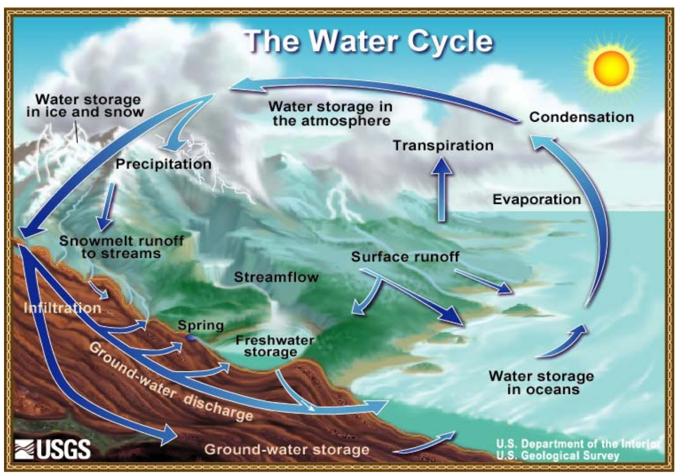


Figure 1.1 The Water Cycle

Retrieved from: http://ga.water.usgs.gov/edu/watercycle.html

Purpose

To comply with the Clean Water Act, water samples must be regularly gathered and areas of severe pollution must be identified. Traditional field sampling has its drawbacks as it is expensive, time consuming, and not representative of a large area. Impaired areas must be identified and continually monitored which drains resources and personnel. The hypothesis of this thesis is that water quality parameters of Utah Lake can be successfully measured through the use of satellite imagery and remote sensing. Specific research questions that fall under this hypothesis are:

- Can a universal model be developed to measure chlorophyll *a* through remote sensing of the water of Utah Lake across time?
- Can a model be developed to automate the measurement of this parameter for different years to be mapped and compared?
- Can a map be generated from the information derived from these models that would provide insight to the distribution of pollutants and contaminants in Utah Lake?
- Can this pilot study be used to determine sample size requirements for a full-scale chlorophyll *a* research of Utah Lake?

This thesis will demonstrate that remote sensing can be used to create a model that can predict water quality parameters across the entire extent of Utah Lake. This research will focus on the water quality parameter chlorophyll *a*.

Importance

Water is essential for life. Large cities and communities have been built around water as they rely upon it for such things as transportation, trade, drinking, crop irrigation, recreation, industry and disposal of waste. Regardless of the reason why people build near the water, their proximity causes degradation to the water.

Although 71% of the earth's surface area is covered with water, 97.22% of the water supply is contained in the ocean leaving only 2.78% fresh water. Of this 2.78%, 77.78% is surface water and groundwater and soil moisture total about 22%. Of the 77.78%

surface water, 99.3% is frozen. Surface fresh waters such as lakes, streams, and rivers make up less than 1% of Earth's fresh water (Christopherson, 2004).

Background

Utah Lake, located in Utah Valley, Utah, is one of the largest natural freshwater lakes in the western United States. It has approximately 76 miles of shoreline and has an average depth of 10 feet (The Columbia Electronic Encyclopedia, 2005). Utah Lake is part of the Jordan River/Utah Lake Watershed Management Unit shown in figure 1.2.

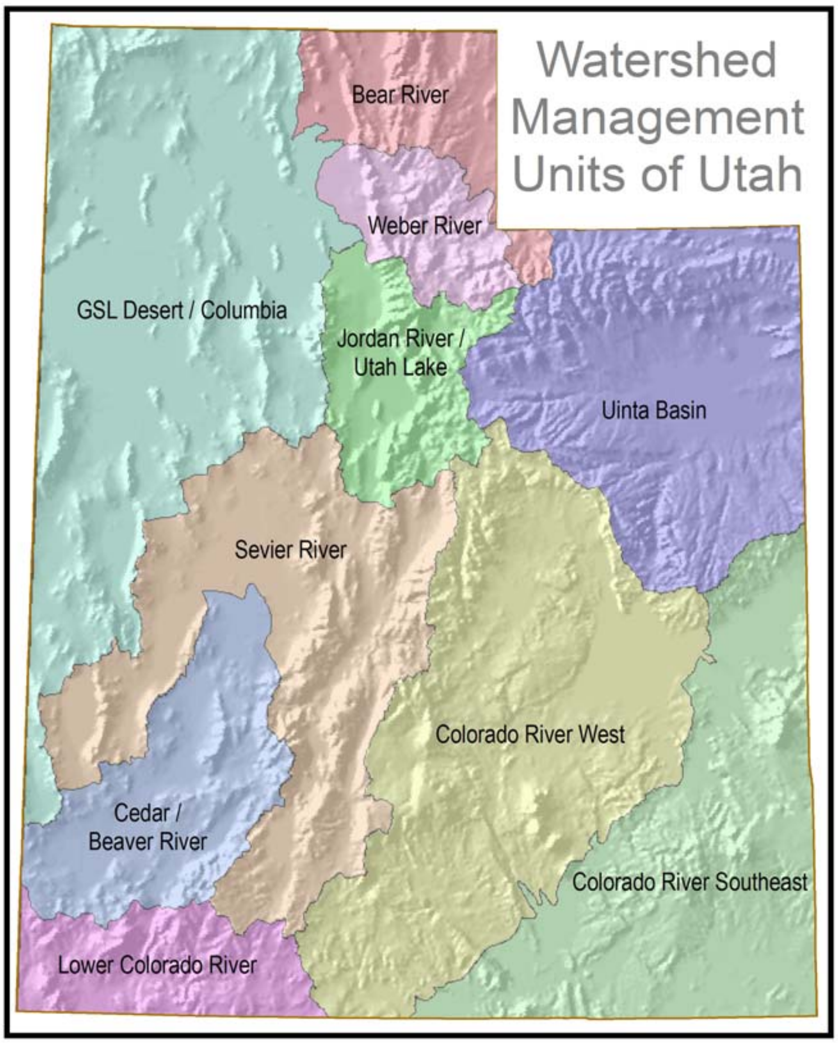


Figure 1.2 Utah Watershed Management Units

The Division of Water Quality listed Utah Lake as a high priority on the 2004 list of impaired waters. This is because many of the industrial permits required parameter limits that could be toxic to aquatic life and to humans (DWQ, 2004). The water quality of Utah Lake only partially supports its designated beneficial uses of protecting warm water species of aquatic life and providing irrigation for agriculture and stock. The main pollutants contributing to this impairment are total phosphorus and total dissolved solids (DWQ, 2004).

Utah Lake is perceived as highly polluted and unimportant (Warnock, 2003). This perception is, in part, a result of the turbidity of the water, making the water look murky most of the time. The shallow water also allows the lake to warm rapidly in the summer. These higher temperatures in-turn encourage the growth of algae blooms. While these characteristics of the lake may encourage the perception of polluted water, pollutants from land uses in the drainage basin can exacerbate the situation. Most of the urban pollutants (other than excess fertilizer from grass) don't really pollute the water in a visible way. The majority of visible pollution is from larger sources such as runoff from agricultural land, steel mill effluent, and sewage treatment facilities.

The Provo-Orem area that surrounds Utah Lake is growing rapidly with an increasing population. The US Census Bureau lists the Provo-Orem area surrounding Utah Lake as one of the ten fastest growing metropolitan areas from 1990 to 2000 with an increase in population of nearly 40 percent during that decade (Perry and Mackun, 2001). As land is cleared and converted to impervious surfaces to accommodate for such an increased population, the amount of non-point source pollutants that drain into the lake also increase.

With such a dramatic increase in population, Utah struggles to restore the quality of water in Utah Lake. Solutions such as a proposal to construct the first water reuse treatment satellite plant in Utah County have been considered. This plant would cost approximately five million dollars and the water from the plant would be used for irrigation (Nardi, 2005). This is one way of making better use of water resources in an area of increasing population. As more contaminants enter the water, more time, expenses, and equipment are needed to improve the water quality. Several steps have been taken to improve water quality. A Utah Lake Restoration Project has been enacted to take the necessary action needed to restore Utah Lake. This and similar projects are implemented as an effort to improve public opinion of the lake. If this project is successful in improving actual water quality and public opinion of water quality in surrounding communities, this could increase revenue brought in from the lake and the extra money would then be invested in further improving the health of the lake and could be used to offset the costs of lake improvements (Warnock, 2003). These projects require the lake to be monitored regularly and thoroughly. The restoration of Utah Lake will be costly regardless of the methods chosen. New techniques for water quality monitoring are being investigated and remote sensing may offer the ability to extend *in situ* data collection to the entire lake and throughout the year.

Thesis Contributions

The successful completion of this pilot study will produce a model using existing data that will serve as a guideline to determine what data to collect and how many samples are necessary to achieve valid results for a full-scale research of Utah Lake. This is the initial step to producing a method of extracting key water quality parameters on a lake-wide basis from remotely sensed imagery and making advances towards more efficient methods of dealing with water quality degradation. By validating the effectiveness of water quality measurement through remote sensing, remote sensing can be thoroughly implemented within the agencies that monitor Utah Lake. Through remote sensing, the entire water body can be monitored regularly, quickly, and in a cost-effective manner.

The development of a model that predicts the amount of chlorophyll *a* present in a water source can pave the way for other water contaminants to also be predicted successfully through remote sensing. This process requires only minor adjustments and redundant steps do not need to be repeated when additional parameters are measured through remote sensing. This allows time to be focused on the analysis of the data and not as much spent on the extraction of the data.

With the creation of these models, measurements can then be calculated and maps generated that show the distribution of the parameter across the lake. The data are extremely valuable as they allow for comparison across days, months, and years. The status of the water quality of Utah Lake can be monitored for present day and compared to past days in search of areas of improvement as well as areas that require more attention. Possible insights might be gained as to unknown sources of pollution or increased pollution at identified sources which would aid in the steps towards restoring the water quality of Utah Lake.

Structure of Thesis

Following this introduction, Chapter 2 contains a review of the literature available about the subject, background information on the Clean Water Act and remote sensing as well as specifics on water quality parameters including chlorophyll *a*. Chapter 3 outlines the methodology used for the research and finally Chapter 4 presents the conclusions, the limitations of the study, and a brief discussion of possible future research.

CHAPTER 2

Literature Review

Water quality has been an environmental concern for several decades. The regulation and monitoring of water in the United States was made official in 1972 with the passing of the Clean Water Act. As a result, field samples are regularly collected to assess the water quality in surface waters across the country. With over 469,495 square kilometers of surface water in the US it is not possible to collect samples at a fine spatial or temporal scale (CIA, 2005). Utah alone has over 11,000 miles of streams and 147,000 acres of lake and reservoirs (The Columbia Electronic Encyclopedia, 2005). New technologies have been developed to compensate for these limitations. Remote sensing is one of the most successful techniques that has been used. Several water parameters have been successfully quantified using remote sensing.

Clean Water Act of 1972

The Clean Water Act was enacted in 1972 in response to a growing public awareness and concern for water pollution control. This regulation was not the result of a single act but several occurrences over a long period of time. This Act established the standard of regulation regarding the discharge of pollutants into waters bodies in the United States. It further gave the U.S. Environmental Protection Agency the authority to implement pollution control programs and establish water quality standards. These standards define the acceptable characteristics of U.S. waters and are used as reference values for judging the quality of the water. Most of these measurements are quantitative, however, some are qualitative. Most states have water quality standards for the following characteristics: temperature, dissolved oxygen, turbidity, bacteria, solids, and toxic substances (Vigil, 2003). Under the Clean Water Act, states are required to identify water bodies not currently meeting water quality standards. It requires that all lakes and streams be regularly monitored to assure that standards are met. Congress requires a water quality report from each state every two years (UDEQ, 2000). These regulations were enacted to identify sources of pollution and improve water quality.

National Pollution Discharge Elimination System (NPDES). As part of the Clean Water Act, the NPDES is a permit program that is required for any pollution discharge from a point source into natural waters. These permits consist of four main parts. The first outlines concentration limits aimed at preventing further degradation. This defines the level of treatment necessary before discharging the wastewater. The second part defines the monitoring requirements which states how often samples are to be taken and what type of analysis is necessary. The third part contains the schedules and compliance conditions that must be met by the permit holder. The final part defines any special conditions that must be met. It often contains features to prevent or reduce pollution discharge in the first place (Vigil, 2003). NPDES permits are designed to protect the environment and public health and in the United States they are quite strict compared to those standards of other countries (Boyd, 2000). The permits are issued on a 5-year basis and an evaluation is conducted at the end of each 5-year period. Permit holders that do not comply with the regulations of the permit are subject to fines and penalties based on the frequency and extent of the damage. These permits can also be revoked if they are being violated (Vigil, 2003).

Each state has the option of creating a state program to meet the Clean Water Act requirements. When this state program is approved by the EPA, the state has primacy over the Clean Water Act. This is called "State Primacy." Their program must meet or exceed the water quality requirements stated by the EPA federal requirements. These are then enforced by the state. Utah chose to exercise its primacy and the National Pollution Discharge Elimination System is known as the Utah Pollution Discharge Elimination System under the rules established by the State of Utah.

Success of Act. The Clean Water Act is a tremendous task and responsibility and does not yet cover all activities that affect water quality. Enforcement of the Act is continuing and the United States has seen a high degree of success as a result of the Act. Water quality has greatly improved and continues to improve as the Act is further implemented (Boyd, 2000).

Measuring Water Quality

Traditional Methods. Several different methods have been used to obtain water samples to monitor our lakes, rivers, and streams. Conventional methods of obtaining water samples tend to be limited because they require high-tech devices, large amounts of field time, different sampling techniques, calibration of equipment, and extensive laboratory analyses to determine the water constituents. Even when carried out successfully, field sampling often fails to accurately represent all areas of the water body being sampled due to the heterogeneous and rapidly changing nature of water. A water sample is usually a grab sample - a sample collected at a random time and date. The results of this sample are only truly accurate for this particular time, date, and location. Pollution is dynamic and difficult to catch in water samples if they are taken at low temporal resolution. Field methods limit the number of samples collected because of costs and time requirements (Harrington and Schiebe, 1992; Lavery *et al.,* 1993; Liu *et al*., 2003; Lyon *et al*., 1988; Novo *et al*., 1991; Ritchie *et al*., 1987).

Emerging Methods. Technology has provided alternatives to the limitations of traditional methods. Satellites continuously circle the globe gathering enormous amounts of information which can be used to remotely sense water quality. Remote sensing does not replace traditional methods of gathering samples as it is reliant on the samples for modeling and calibration, however, it does provide a synoptic view which improves estimations over large areas. Remotely sensed data acquired from an aircraft or a satellite are much quicker and cost efficient than traditional methods. Satellites especially make it possible for repetitive data to be obtained and compared, and for dynamic and surface water maps to be created (Carpenter and Carpenter, 1983; Liu *et al*., 2003; Novo *et al.*, 1991; Harrington and Schiebe, 1992; Lavery *et al*., 1993; Ritchie *et al*., 1987; Rimmer *et al*., 1987; Zilioli and Brivio, 1997).

Remote Sensing

History. Remote sensing is not a new method in water resource management. Aerial photographs have been used for decades as a means of locating water bodies and identifying possible impairments. Satellite images of the Earth have been used since the early 1970s (Remote Sensing Water Resources Management, 1973). This form of measuring improves as technology advances and images are more easily analyzed.

Sensors. Satellites circle the Earth today which provide an array of images and data that are available for analysis. Satellite data are becoming so popular and widespread that – depending on the resolution - much of it is available to the public free of charge.

There are programs such as Google Earth that allow people access to satellite imagery worldwide (Porteus, 2005). Although remote sensing offers a solution to the limitations of conventional sampling methods, it does not eliminate *in situ* sampling; rather, it reduces the amount of samples required and allows the entire water body to be quantified. Samples are vital when creating and testing models to quantify water parameters. If properly calibrated to the *in situ* samples, remotely sensed imagery offers a solution for determining amounts of certain water characteristics and pollutants over the entire surface of the water body.

Successful quantification of water parameters through remote sensing is affected by the type of sensor used for the research. A range of different sensors have been tested for their ability to accurately measure water quality. While there are satellites that are designed specifically for water monitoring, these are made for ocean remote sensing. Due to the spatial resolution of these sensors they are not generally suitable for monitoring smaller, inland water bodies or rivers. Research of these waters often relies on meterological and Earth resources satellites with a higher spatial resolution. Some of these include but are not limited to Landsat, Spot, AVHRR, and IRS. Chlorophyll concentrations have also been measured using a variety of images including those produced by Landsat. Strong correlations have been produced using MSS, however TM is used more frequently (Liu *et. al*., 2003). Lyon *et al*. (1988) determined suspended sediment concentrations from Landsat and AVHRR imagery. This study used several different dates of imagery from each platform. Hoogenboom *et al*. (1998) used AVIRIS to detect chlorophyll concentrations of Dutch coastal and inland waters. Lui *et al.* (2003)

provides an overview of additional sensors that are available for use in quantifying water quality parameters.

Water Quality Characteristics well suited for Remote Sensing

Types of water parameters that can be successfully measured through remote sensing may range from pH level and minerals to in-water constituents and finally to physical characteristics such as temperature and bathymetry (Liu *et al*., 2003). Harrington and Schiebe (1992) used Landsat MSS to monitor suspended sediments, turbidity, and secchi depth of Lake Chicot, Arkansas and determined that remote sensing can provide meaningful information on water quality. Successful models were created for quantifying these parameters remotely. Lui *et al.* (2003) summarized that suspended sediment and turbidity, chlorophyll concentrations, Coloured Dissolved Organic Material (CDOM) - algae, yellow substance, and organic plumes, and Secchi disk depth (SDD) - water clarity or transparency have been successfully measured.

Chlorophyll a

Chlorophyll *a* (Chl *a*) is a green pigment found in plants and is used in photosynthesis to absorb sunlight and convert it to sugar. It is a phytopigment present in all algae groups in inland waters (Thiemann and Kaufmann, 2000). The amount of Chl *a* contained in a water body is a good indicator of phytoplankton biomass and is therefore commonly used as a water quality indicator. Phytoplankton are microscopic planktonic algae that are suspended in water. They remain near the surface because sunlight is more abundant for their growth. Growing phytoplankton are usually found only in areas where the water is illuminated and there is a source of nutrients for photosynthesis.

Effects on water quality. The most obvious affect of phytoplankton on water quality is the affect it has on the color of the water. The algae cause it to appear turbid and green because they contain chlorophyll (Vigil, 2003). In addition to this murky water looking undesirable, there are other unseen affects that have a much greater impact on the water. The three main effects are:

1. its influence on pH levels,
2. fluctuation of concentrations of dissolved oxygen,
3. and fluctuation of carbon dioxide amounts in the water (Boyd, 2000).

During photosynthesis of phytoplankton, carbon dioxide is used and dissolved oxygen is released which causes a decrease in carbon dioxide levels and an increase in dissolved oxygen. The decrease in carbon dioxide causes the pH level of the water to increase. One of the most important variables related to the well-being of aquatic ecosystems is dissolved oxygen concentrations (Vigil, 2003). During the day the dissolved oxygen concentrations are high but as the temperatures cool, they are lowered significantly. This fluctuation is exacerbated as the amount of phytoplankton production increases.

In addition to the affects on water chemistry, many nutrients are removed from the water by phytoplankton and it serves as a source of organic matter. If the concentration of algae is too high, it results in high amounts of organic matter and extreme fluctuations of dissolved oxygen. If dissolved oxygen levels decrease too much, many aquatic species are stressed and only those with a high tolerance for low amounts of dissolved oxygen are able to survive (Boyd, 2000). High levels of Chl *a* usually indicate low water quality and low levels normally indicate high water quality.

Measuring Chlorophyll a. The most common method of collecting Chl *a* is through water samples of a known volume that is then filtered through fine mesh filter paper. The sample is later analyzed to determine the content of Chl *a*. Three standard methods for determining the amount of chlorophyll are: spectrophotometry, fluorometry, and high performance liquid chromatography (HPLC). In general, chlorophyll concentrations are measured in micrograms per liter (*ug/l*).

Predicting Chl a content using Landsat TM Imagery

Predicting Chl *a* using a regression model derived from remotely sensed images has a marginal success history. Unique models are required for each study site. There are various factors that can contribute to or hinder the success of the model that may be present in some areas and completely absent in others. The size of the water body, depth, region, and surroundings can all have an impact on the success of the prediction equation. The purpose of the study may also influence the equation. Some studies may be conducted to predict Chl *a* values for a certain month of the year, others for only certain seasons, and others attempt to produce a single model to predict chlorophyll content throughout the year.

Allee and Johnson (1999) did research on Bull Shoals Reservoir located on the Salem Plateau of the Ozark uplift between Arkansas and Missouri using Landsat TM data and samples taken for four different seasons. Two significant models resulted from the seasonal regression analyses. One model found bands 2 and 3 to be good predictors of chlorophyll and the other used bands 1,2,3 and 5 for prediction. They also pooled the data together in an attempt to produce a single model that could be used to predict Chl *a* content throughout the year, however, no significant model was successfully created with the combined data.

Another study of Lake Kinneret – a freshwater body in Israel – conducted by Mayo *et al.* (1995) discovered that the effect of suspended matter on reflectance in the blue region of the spectrum must be taken into consideration because results using the ratio of band1/band2 were not successful. They found that the reflectance in band 3 depended primarily on suspended matter concentration; therefore, using the ratio algorithm (band1-band3)/band2 predictions were much more successful with an R^2 of .71. The regression equation was formulated from Landsat TM bands simulated using *in situ* reflectance measurements. When the equation was tested using reflectance extracted from atmospherically corrected TM data, the R^2 value dropped to .49. They suggested the reason for the decrease in accuracy was due to a three day time interval between the field data and the satellite overpass.

Research on a sewage outfall site off the North Head of Sydney Harbour, Australia conducted by Forster *et al.* (1993) used Landsat TM imagery to predict Chl *a* content. The multiple regression analysis produced $R=0.9$ which yield an $R^2 = 0.81$, however, due to the small number of ocean samples, the results were not significant.

Remotely Sensing Utah Lake

Many successful models have been developed for lakes, rivers, and coastlines. Unfortunately, these models are not universal and are unique only to the water body from which they were created. It has been shown that many parameters can be remotely sensed using satellite imagery, but unique models for each lake must be developed (Fraser, 1998; Lui *et al.*, 2003). As one of the impaired water bodies of Utah, Utah Lake is of great concern and the water quality must be monitored and improved. The distribution of Chl *a* in the lake is irregular, both spatially and temporally, and this causes severe sampling problems. Figure 3.2 shows how different the makeup of the water is across the lake. An attempt to accurately determine Chl *a* concentrations across the lake using *in situ* methods would require hundreds of samples at different locations on the lake in a very short amount of time. A successful model to predict and quantify levels of Chl *a* across the lake could be very beneficial and aid in the processes of restoring the lake water to meet current water quality standards.

Summary

Water quality regulations and standards were developed out of a growing public concern for cleaner water. As a result, samples are regularly taken, filed, analyzed, and reported. These samples are taken in the field and require large amounts of time as well as equipment and expertise to be collected correctly. There are some limitations to these methods and as a result newer techniques are being tested in an attempt to improve water quality monitoring as well as save time and money. Many technological advances have also allowed for entire water surfaces to be monitored as opposed to only those direct areas from which samples are pulled.

Remote sensing is one technology that has been used to help monitor water quality. Many different parameters, including Chl *a* have been successfully quantified for a variety of water bodies around the globe. This is a characteristic that indicates the amount of phytoplankton in a water body. Algal blooms are a perennial problem at Uath Lake and are one of the reasons why the lake has been classified as impaired. The successful measurement of such parameters in Utah Lake could be of great benefit when attempting to monitor the lake and cleanse it so it meets water quality standards.

CHAPTER 3

Methodology

Site

Utah Lake, located just south of Salt Lake City, Utah, (Longitude -111.7925, and Latitude 40.1958) is large, shallow, and open to wind. This presents hazards to recreational users and, therefore, it is often viewed as undesirable for recreation. People must be in the center of the lake to acquire adequate depths for swimming and boating (Shiozawa, 1977). Its flat profile and shallow waters makes it more vulnerable to water impairment and high turbidity. Covering about 96,900 acres, the deepest area of the lake is only 14 feet (DWQ, 2002). The Provo, Spanish Fork and American Fork Rivers are the primary inflows, and the Jordan River drains the lake north to the Great Salt Lake. Figure 3.1 shows Utah Lake including the primary rivers that flow in and out of the lake. The Division of Water Quality of Utah classifies Utah Lake with beneficial uses of: boating and similar recreation, protecting warm water species, protecting waterfowl, and agricultural use (DOQ, 2004).

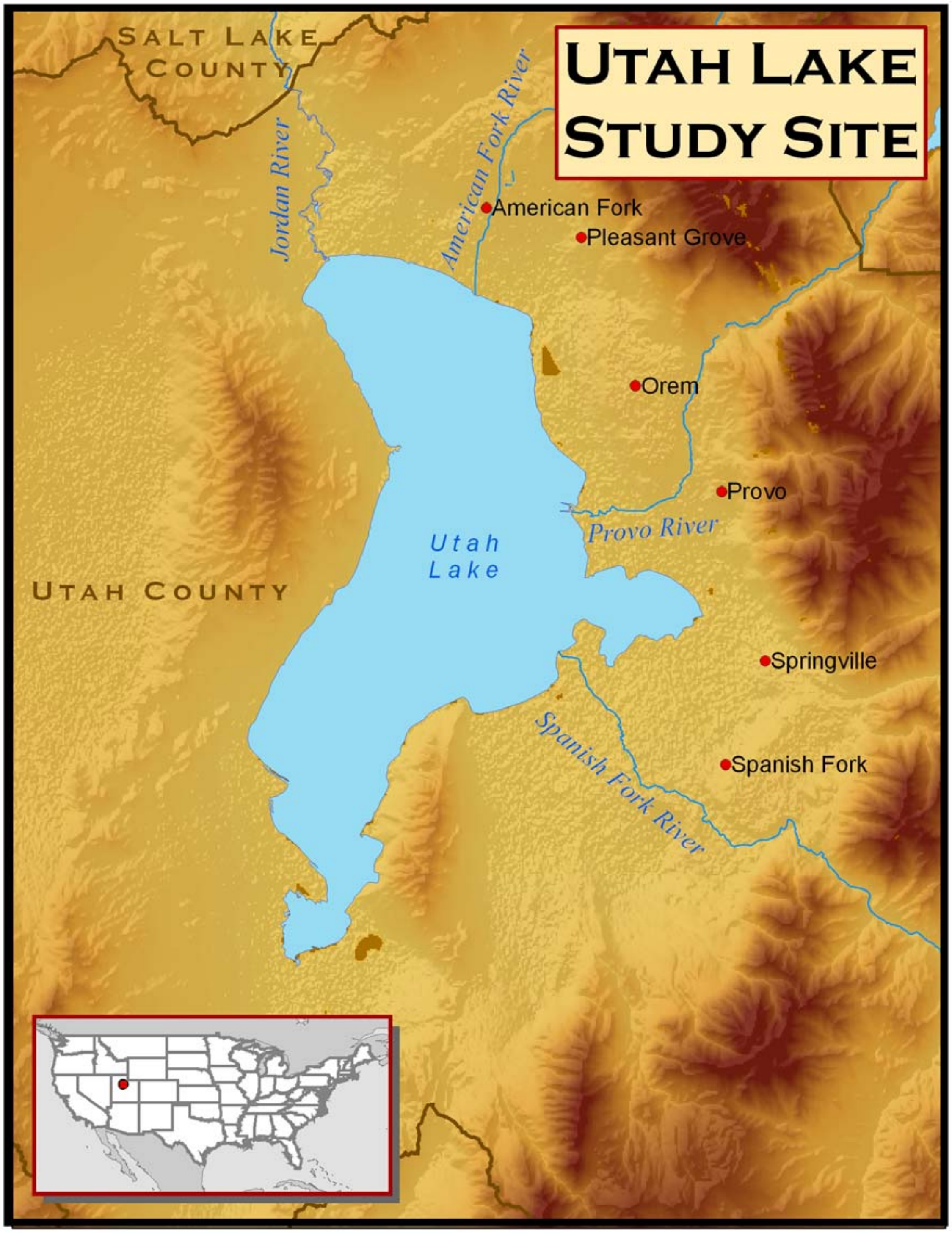


Figure 3.1 Utah Lake Study Site

As mentioned earlier, the population of the Provo-Orem area increased by nearly 40% between 1990 and 2000 (Perry and Mackun, 2001). Several cities in Utah County more than doubled in population within these ten years and the county population expanded by over 100,000 people from 1990 to 2000 (Utah County, 2004). Such a high increase in population has many consequences, particularly to land use. Residences and jobs are needed, demanding an increase in commercial, residential, and industrial areas. This ultimately increases the amount of impervious surface in an area. Water runoff is directly proportional to the amount of impervious surface. As the amount of impervious surface increases, floods, property damage, and non-point source (NPS) pollutants increase as well. In the EPA's 1992 Report to Congress, non-point source pollution was cited as the dominant cause of impairment for streams and rivers (EPA, 1994).

Table 3.1 below shows how population changed in Utah County over the decade from 1990 to 2000. Currently the population is located primarily in the Provo-Orem area, on the eastern side of Utah Lake. The proximity of the mountains to the east has slowed the growth in these cities, resulting in a shift in development to the north and south. The northern area of the county in particular has seen rapid growth as it serves both Utah and Salt Lake Counties (e.g., Lehi had an annual average growth rate of 8.4 %).

	1990	1994	1996	1998	2000	AARC (%)*
Alpine	3492	4634	5161	5418	7146	7.4
American Fork	15696	18222	19451	19215	21941	3.4
Highland	5002	5336	5939	6315	8172	5
Lehi	8475	11069	13810	15297	19028	8.4
Lindon	3818	4890	5941	6380	8363	8.2
Orem	67,561	76,987	79,736	78,937	84,324	2.2
Pleasant Grove	13,476	16,381	19,357	20,491	23,468	5.7
Provo	86,835	98,224	99,606	110,419	105,166	1.9
TOTAL COUNTY	263590	302052	319694	335635	368536	3.4

* Average Annual Rate of Change

Table 3.1 Utah County Population Increase (1990 - 2000)

Retrieved from: Governor's Office of Planning and Budget

With increasing population, concern over the health of Utah Lake is growing. Accurate and time-effective techniques for monitoring the water must be implemented before the water quality of the lake is exacerbated.

Utah Lake was chosen for this research on account of its unique characteristics. It is a large, shallow, freshwater lake in an arid environment. Spatially, the pollutants across Utah Lake are not evenly distributed. Due to the shallow nature of the lake, the distribution of pollutants across the lake can change quickly if there are high winds in the area. The area surrounding the lake has also changed with time. In the past, Utah Lake was an agricultural and industrial waste receiver and today it is surrounded mostly by residential areas. These variables affect the concentrations of pollution which makes this lake a great candidate for research on the effects of remotely sensing water quality.

Structure of Methodology

Several steps were required to achieve an equation that can accurately predict the chlorophyll *a* content in Utah Lake. These steps include:

- Gathering data
- Preprocessing
- Analysis

Each of these steps will be further simplified and explained in detail in the following paragraphs.

Gathering Data

The initial step for nearly all research is to gather the data necessary for the analysis. The data required for this research are:

- Water samples from Utah Lake

- Satellite Imagery

Gathering Data - Water Samples

The Environmental Protection Agency is a federal agency that ensures that states are meeting the water standards stated in the Clean Water Act and regularly sampling their waters. They gather the results of all samples that are taken by the states and post them on the World Wide Web. The sample results are available to the public through the Utah Division of Water Quality web site (http://www.waterquality.utah.gov/monitoring/data.htm) and are stored in a database called Storet. This data can be accessed by hydrologic unit, county, sampling station, date, and characteristic type. There are 29 sampling stations on Utah Lake that are monitored regularly for chlorophyll *a* content.

Water samples of chlorophyll *a* were downloaded from 1972 to present. The samples used in this study were collected using the fluorometric method and are uncorrected for pheophytin - the pigment fraction which is not active in photosynthesis. This data was then queried for samples taken within 5 days of a Landsat TM overpass. May 28, 1991 imagery was used with samples taken on May 23, 1991, the samples and imagery for July were the same date – July 15,1997, and finally samples take August 19, 1999 were used for the August 14, 1999 image. Twenty-nine samples were extracted that spanned across three different years of imagery from 1991 to 1999. Only 27 of these samples were used in the analysis, however. Two outliers were discarded from the sample set leaving the middle 95% of the data for analysis. These outliers were removed because of their influence on the statistics of the data. The values were 151 and 165 ug/l. Without these samples, the highest value drops to 62 ug/l. The average for the samples including these two outliers is 30.4 ug/l but excluding them leaves and average of 20.9 ug/l. The Division of Water Quality was contacted and declared that the samples were valid values. The locations of the samples were then analyzed for possible closeness to the shoreline. No explanation could be found except a high amount of chlorophyll in those areas; however, due to the influence on the statistics of the data, the outliers had to be discarded to assure a more accurate model. The 1999 image had four samples, six were taken from the 1997 image, and the 1991 image contained 17 samples. The 27 samples used were taken from 17 unique stations on Utah Lake. Precise geographic locations were extracted from the Storet site for each of the lake stations and were plotted in space using ArcGIS software. Figure 3.2 shows the distribution of the sample locations used in this research. The marked locations pertaining to the samples taken for each year of imagery were saved in separate arc coverage files.

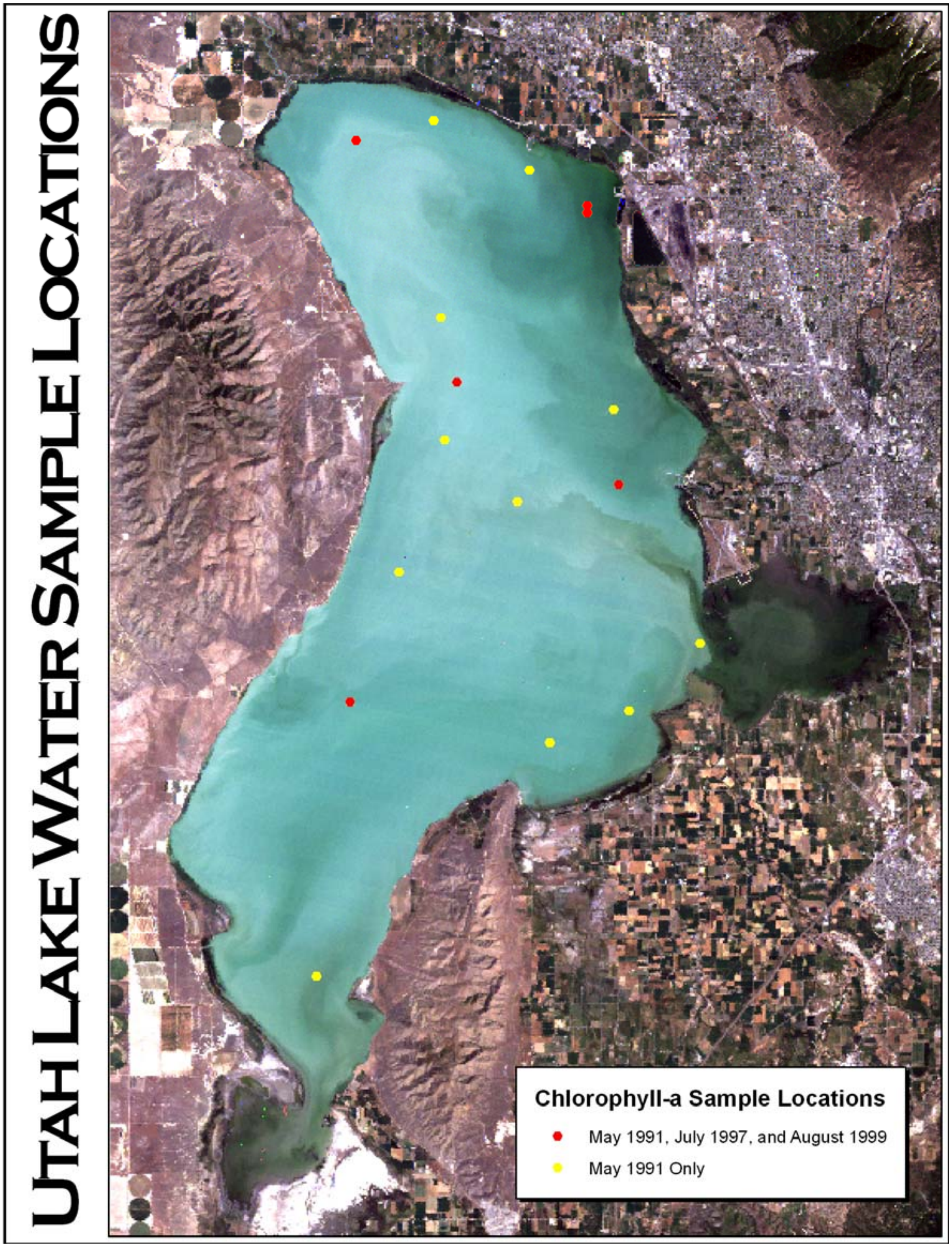


Figure 3.2 Utah Lake water sample locations

Note: Background image is the July 1997 Landsat TM image used in the research.

Gathering Data - Landsat Thematic Mapper (TM) Satellite Imagery

The first Landsat satellite was launched July 23, 1972 and the most recent one was launched April 15, 1999, making the Landsat Program the longest running program for acquiring earth images from space (Sheffner, 1999; Jensen, 2004; Jensen, 2000). They have launched six successful satellites during those years. Enhancements in the sensors have allowed for an improvement in the imagery as new satellites are launched. Landsat Thematic Mapper (TM) has an improved spatial resolution of 30 meters compared to the previous Landsat MSS which only had 60 meter resolution. Table 3.2 below shows the number of Landsat TM bands, general characteristics, and the spectral resolution of each.

Landsat Thematic Mapper		
Band	**Spectral Resolution (µm)**	**Characteristics**
1	0.45-0.52	blue
2	0.52-0.60	green
3	0.63-0.69	red
4	0.76-0.90	reflective infrared
5	1.55-1.75	mid-infrared
6	10.40-12.5	thermal infrared
7	2.08-2.35	mid-infrared

Table 3.2 Landsat TM bands

Source: Jensen (2000) pg. 186

For this research, geometrically corrected digital Landsat data (path 38 and row 32) for six spectral bands was obtained for May 1991, July 1997, and August 1999. These are the three images that coincided with the chlorophyll samples. These images were then re-projected the WGS 84 to comply with the water sample locations. The Landsat satellite gathers data in seven different bands; however, the 1999 image was missing band 6 (thermal band) therefore only the first five bands and band 7 were included in the analysis.

Preprocessing

After the necessary data was gathered, it required some preprocessing before the regression analysis was ran. These steps included:

- Normalizing the images to correct for atmospheric differences
- Subsetting the images to the extent of Utah Lake
- Extracting brightness values from the images at the locations of the water samples
- Masking out only the lake from the images

Preprocessing - Atmospheric Correction

When comparing several different images across time, changes in sun angle, atmospheric conditions, and sensor calibration cause the images to differ in brightness values in similar locations. These influences need to be controlled to attain an accurate comparison between the images. The images must appear as if obtained under the same atmospheric conditions and with the same sensor.

Several techniques exist using the angle of the sun and large amounts of data about atmospheric conditions which are sometimes not readily available. A simpler method, Multiple-date Normalization or Relative corrections, which successfully corrects for atmospheric differences, involves choosing a base image and adjusting the remaining images to the one that was chosen (Jensen, 2004; Hadjimitsis *et al.*, 2004). Song *et al.* (2001) discovered that the more complicated algorithms do not necessarily yield improved results and they recommend the more simple methods. This method was chosen for adjustment of the images.

The July 1997 image was chosen as the reference scene to which the other images would be normalized. Although not mandatory, it is common for the most recent image to serve as the base image. In this case the August 1999 image would have been chosen for the base, it being the most recent, however, it was the most dissimilar image while the other three were fairly similar in reflectance. Choosing a base image that was similar to the others would result in a smaller adjustment of the images and generally a smaller window of error. The July 1997 image served as the dependent variable and the other images were then normalized to it. Several areas where no brightness value changes (pseudo-invariant targets) should occur are then identified for each image (Hadjimitsis *et al.,* 2004). These areas are assumed to have a constant reflectance, therefore, any changes in brightness values between images were assumed to be due to detector calibration, atmospheric differences, and phase angle differences. Areas such as large rooftops, dried lake beds, clear water, and dry soil are good examples of constant reflectance locations (Ahern *et al*., 1977). The best results were attained when unique points were identified in each image and compared to the 1997 image. Twenty-one sites that suited these requirements were chosen for the 1999 image and 17 sites were chosen for the 1991 image. These sites were then overlaid on the images and the brightness values for each band were extracted for each point.

A regression analysis was performed for the brightness values of each band against those of the corresponding band of the 1997 image to predict what a given brightness value would be if the image had been acquired at the same time as the 1997 image. The resultant regression equations were then used to adjust each image for further analysis. The equations used to adjust each image are found in table 3.3 below and the R^2 values from the regression analysis are found in table 3.4.

Image	Atmospheric Correct Regression Equation
May 1991	$1991 = -4.135 + 1.030(B1) + 0.066 + 0.974(B2) + 4.288 + 0.938(B3) + 8.162 + 0.882(B4) + 11.172 + 0.880(B5) + 18.537 + 0.740(B7)$
August 1999	$1999 = 62.185 + 2.082(B1) + 018.921 + 1.139(B2) + 16.514 + 1.417(B3) + 11.202 + 0.911(B4) + 16.908 + 1.309(B5) + 1.153 + 1.150(B7)$

Table 3.3 Regression equations used to normalize the images

Aug 1999 Image		May 1991 Image	
Band	**R^2**	**Band**	**R^2**
1	0.925	1	0.966
2	0.912	2	0.955
3	0.923	3	0.964
4	0.911	4	0.943
5	0.908	5	0.894
7	0.773	7	0.678

Table 3.4 R^2 values from the atmospheric correction regression analysis

Spatial Modeler, a tool in Erdas Imagine, was used to create a model that automated the atmospheric correction. A unique model containing the regression equations from the comparison was created for each image. The script and overview for the model created for the May 1991 image are found in the attached Appendix A. The May 1991 model was created in a similar fashion.

This method has been found to be a simple, successful means of reducing differences across images due to atmospheric noise and interference. Once the variances in the images were removed, any change in brightness value can be assumed to be related to changes in surface conditions. The successful atmospheric correction allows the resulting Chlorophyll model to span time and space.

Preprocessing – Subsetting the Images

Erdas Imagine 8.7 produced by Leica Geosystems Software is designed specifically to process imagery. The entire satellite image was required for choosing pseudo-invariant targets for the atmospheric correction, however, once that was completed, this software was used to subset the images to the general extent of Utah Lake. This was done to dispose of unneeded data, to increase processing time, and to improve the comparability of the images.

Preprocessing - Extracting Brightness Values

Once the images were atmospherically corrected and subset, the lake locations of sample measurements were overlaid on each image. The average brightness values of each band of the surrounding 5x5 matrix of pixels of the sample locations were extracted. An average of the values was computed to compensate for possible location inaccuracies and the scatter of the sample pulled (Lavery, 1993). The resultant averages per band were added as a new field in the arc coverage file. The arc coverage file was an easy way to organize the data so that the original water samples and the corresponding brightness values for each band were placed orderly in columns in preparation for the regression analysis. This process was repeated for each image. A model was created to automate this process and reduce repetitive tasks. It also provides a skeleton for running similar analysis on different water characteristics in the future. An overview of the model and script are found in Appendix B.

Once the brightness values for each image were extracted, the tables from the arc coverage files were exported into .dbase files and opened in Microsoft Excel. The information for all three years were combined and arranged into seven columns and 27 rows. Six of the columns were the brightness values from the different bands used and the last one represented the amount of chlorophyll *a* sampled for each location. The 27 rows represented each of the water samples used in the research with the original measurement as well as the corresponding brightness values extracted from the images. This table combined and organized all the necessary information to run the regression analysis.

Preprocessing – Masking out the Lake

A digital elevation model (DEM) was combined with an algorithm utilizing band four to extract only the lake from the images. This would allow for easy overlay when mapping the predicted values as well as quicker processing time. This study is only concerned with the water of Utah Lake. A conditional statement was used for the mask, keeping only those areas of the image that had a brightness value in band 4 that was greater than 40 and an elevation less than 1,375 feet. These criteria were determined through examination of the images as well as the DEM and were successful at extracting only the lake. This mask was added to the final model of the regression equation as the initialstep before the image is run through the equation (see attached Appendix C). This was done to reduce the number of steps required to obtain chlorophyll *a* prediction values of Utah Lake.

Analysis

With the completion of the pre-processing, the data was ready for the central step of the research:

- The multiple regression analysis

- Power statistical analysis

Analysis - Multiple Regression Analysis

This research uses multiple regression for the analysis of the data. It is the most widely used dependence technique and is generally used when attempting to predict values for a dependent variable. Mayo *et al.* (1995) stated in their research that a statistical approach is the most common for determining a correlation between chlorophyll and satellite spectral bands. In a regression analysis, the dependent variable (chlorophyll *a*) is what is being predicted or explained by the independent variable(s). The independent variables (the brightness values of each band of the satellite imagery) are the variables being used to predict the dependent variable (Hair *et al.*, 1998).
A regression equation generally follows the following format:

$$Y = \alpha + \beta_i X_i + \epsilon$$

where Y = cholorphyll *a* (water quality parameter), α = the model intercept, β_i = regression coefficient for variable *i* (the slope), X_i = measured value for variable *i,* and ϵ = error term.

There were two assumptions upon which a multiple regression analysis is dependent:

1. For any value of X, Y is normally distributed.
2. The Y values are statistically independent of one another.

The data from the table created during the preprocessing steps was then entered into SPSS statistical software. This data were used to form linear models for chlorophyll *a* from the satellite data using a multiple regression procedure. The dependent variable was the measured values of chlorophyll *a* samples and the predictor variables were the spectral radiance values of each of the bands for the three years of imagery.

Initially all six bands were included in the analysis, but a more efficient model was obtained using fewer bands. Frequently used band ratio algorithms (e.g. 3/4, 2/3) were also tested, and eventually several more complex ratios were added to the regression (e.g. (3+4)/1, 2/(1+3)) including NDVI. Many of the bands and ratios were further transformed using square root, cubed root, squared, and cubed to make them have a more linear distribution. All regression models were initially evaluated based on the R^2 value and *p*-value. The R^2 value quantifies how much variance in the dependent variable is explained by the model. A value of one would indicate 100% of the variance explained by the model; therefore, higher R^2 values were indicative of better models (Allee, 1999; Berry and Feldman, 1985). The adjusted R^2 indicates how well the model predicts the dependent value based on the number of terms in the equation and the significance of each term. The *p*-value indicates the degree of significance (as a percentage) that each parameter contributes to the overall model. Each term was only considered if it resulted in a p-value < 0.05. A linear regression analysis was most successful and resulted in an equation which utilized three bands in different ratios and transformations. General statistics for the regression equation are summarized in table 3.5. The equation is as follows:

$$\mathit{Chlorophyll\ a} = 9.566 - \left(4.06 \times 10^{-05}\right)(B3)^3 + 2.141(B4) - 4.253(B7)$$

where chlorophyll *a* is measured in *ug/l*.

R	R^2	Adjusted R^2	Std. Error of the Estimate
0.75	0.562	0.505	9.302

Table 3.5 Regression equation general statistics

Once again, Spatial Modeler was used to implement this equation into a model. This would allow for quick analysis of each pixel in each of the three images. An overview of the model as well as function codes are found in the attached Appendix C. Each image was then entered into the model and the resultant images contained only the predicted chlorophyll *a* values for the entire Utah Lake for each corresponding year.

Analysis –Statistical Power Analysis

A pilot study is often conducted to determine necessary requirements for larger research. One of the objectives of this study is to determine an adequate sample size to achieve valid results for a full-scale research of the chlorophyll *a* content of Utah Lake. One method of determining necessary sample size is through a power analysis. Power is the probability that a significant relationship between the dependent and independent variables will be found if one actually exists (Hair *et al.*, 1998). The power of a test is defined as 1-beta, where beta is the probability of falsely accepting the null hypothesis when it is not true – stating there is no relationship when one actually exists. Given a measure of desired power, an alpha, and an effect size, an adequate sample size can be determined (Buchner, 1997). In a study using a single dependent variable, effect size refers to how much of the variance is guaranteed to be explained with the resultant equation. Effects that are smaller than the effect size are considered negligible. Alpha is the significance level desired for the results.

G*Power statistical software was used to run the power analysis and determine the necessary sample size for more in-depth research. It is a software designed to calculate power values, sample sizes, and alpha and beta values and is available free of charge on the World Wide Web (Erdfelder, 1996). An *a priori* analysis using an F-test

(MCR) was used because it was designed to be run using multiple regression analysis data before an experiment is conducted and it determines sample size. J. Cohen provides some guidelines for power analysis and suggests that a study should be designed to achieve significance levels of .05 with power levels of at least 80 percent (Hair *et al.*, 1998). He also provides effect size conventions for small, medium, and large effects for a multiple regression power analysis which were implemented into the G*POWER software (Buchnar, 1997). A medium effect size (.15) was utilized in the power analysis for this pilot study, assuring the explanation of at least 15 percent of the variance. The effect size, the power value, and alpha were entered into the software. It was determined that a minimum of 77 samples are necessary to achieve the assigned degree of significance and power for larger research on the chlorophyll *a* content of Utah Lake. Also, additional samples would be required to be used in the accuracy assessment at the multiple regression analysis would require the use of all 77 samples. As the sample size increases, the power of the equation increases along with the amount of control over a Type I error. If sufficient funds and time are available for a larger project, gathering more than the minimum of 77 samples would further increase success and power of the predictive equation.

Limitations of the Model

Although normalizing the images to the 1997 image does allow the model to span time and space, unfortunately, it also ties the model to that image. Anyone that wishes to use the model must first normalize their image to the same 1997 image before proceeding. This does not necessarily limit the uses of the model; however, it makes it more inconvenient. An atmospheric correction must be performed before running data through similar models regardless of the process chosen, however in this case they are limited and must use the 1997 image. There is also a possibility of decreasing the accuracy of the results if the image used is drastically different than the 1997 image used in this pilot research.

Extracting Predicted Values

Once the regression equation was applied to all three images, the predicted values for each of the sample sites were determined for comparison with actual sample values. A small model was created to automate this process. The prediction values were the average of the 5x5 matrix of surrounding pixels to correspond with the brightness values that were extracted for the regression analysis. An overview of this model is contained the attached Appendix D.

CHAPTER 4

Results and Conclusions

Accuracy of the Model

With a small sample size of only 27, every sample was required to do the analysis, leaving none for the accuracy assessment. In situations where this occurs, Lui *et al.* (2003) discovered that the validation of the model can be based on the R^2 value, the significance level of the model, and scatterplots of the predicted values against the original values. These will be the strategies used for validation of this research. The regression analysis produced an R^2 value of 0.56. Interpreted, this means that the results from the regression equation account for 56% of the variance in the cholorphyll *a* content. With this degree of accuracy, the data can still be used to estimate actual chlorophyll *a* content at any particular spot; however, it might better serve to identify general patterns and areas of concern. Chart 4.1 shows a graph comparing the predicted chlorophyll values with the actual sample values. The closer the values fall to the regression line, the better the model. While the points fall fairly close to the line on this scatterplot, it also shows that there are some outliers in the data. These are the areas that would not be predicted as accurately by the model as they represent cases that differ from normal chlorophyll measurements. Also, the model would tend to slightly over-predict the amount of cholorphyll *a* in an area. The higher measurements of real data have skewed the prediction line towards higher values and many of the points actually fall below the line.

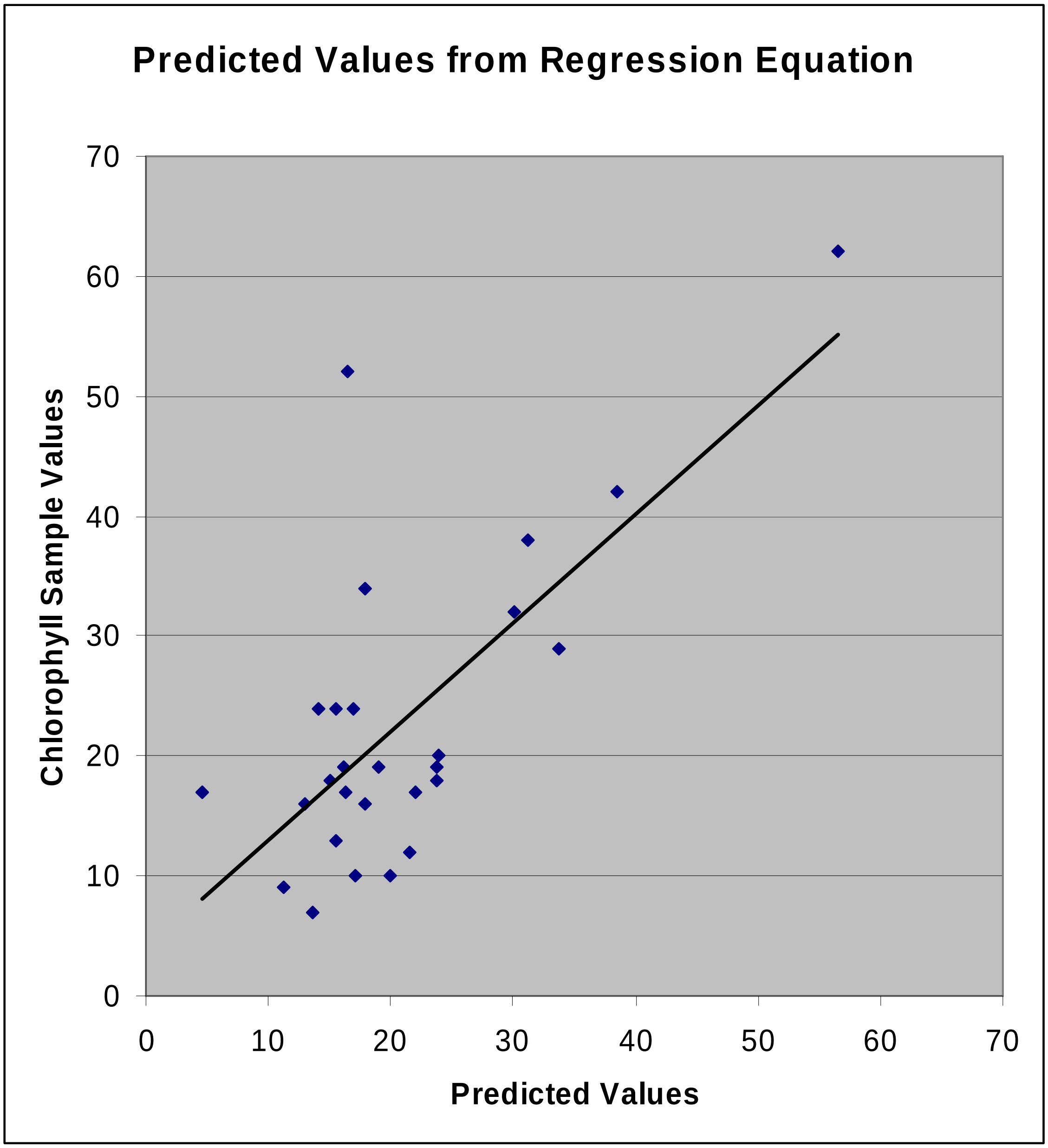


Chart 4.1 Scatterplot of chlorophyll-a sample values vs. predicted values

Results

Figures 4.1 – 4.3 show the distribution of chlorophyll *a* across Utah Lake as determined by the model. Many states suggest that, in general, a shallow lake should contain between 10-20 ug/l of cholorphyll *a* (ADEQ, 2005). These values were used in the maps to better show patterns and areas of concern. As seen, the majority of the chlorophyll *a* content in Utah Lake exceed these criteria. The values were mapped in such a way as to highlight problem areas as well as areas of slight concern through an increase of 5 ug/l increments. The areas with a value below 20 ug/l are considered no risk and are mapped as blue, healthy water. Sections of the lake that contain between 21 and 45 ug/l have exceeded recommended maximums and are considered to be of concern. They are mapped in shades between yellow and orange. Finally, the red areas are representative of those areas that greatly exceed the maximum suggested values of chlorophyll *a*. The slight striping that is evident in some of the images is a result of striping that occurred in the images that were processed. Landsat scans back and forth to collect data which results in a minor difference between passes due to the bidirectional reflectance distribution which causes a slight striping to appear in the images.

Obvious patterns of concern are evident. Many of the images show critical areas to the east and south of the lake. This would be expected as those areas are extremely shallow areas and often recede with the fluctuation of the level of the lake. The red areas along the shoreline are expected as those areas are also shallow and would therefore have different brightness values. The difference in brightness could be caused by either reflectance off the bottom of the lake in the shallow areas or warmer water along the shoreline which would have more nutrients and therefore more chlorophyll *a* due to rapid phytoplankton growth, or a combination of the two. The areas of concern are those red areas that are found towards the middle of the lake, especially evident in the 1999 image. These sites would be the targets for further testing and analysis if the chlorophyll *a* content of the lake is to be improved. It is also important to keep in mind the month of the imagery. It is expected for the May image to have lower levels due to the fact that the shallow water has not yet been exposed to the summer heat so the phytoplankton are not found in as much abundance. The August images would have much higher levels of chlorophyll *a* as algae tend to bloom and thrive with more intense sunlight.

Based on the regression equation that resulted from this research and the power analysis run, a minimum of 77 *in situ* samples are necessary to achieve valid results when determining the chlorophyll *a* content of Utah Lake. Additional samples should also be gathered to perform any necessary accuracy assessments of the prediction equation. A known sample amount will be valuable in determining the cost and timeline of such extensive research.

An image acquired August 29, 1990 (Figure 4.4) was also normalized to the July 1997 to be compared to the August 1999 image. These images are of the same month and represent nearly the same decade of growth as the census statistics cited in this research. It is interesting to note that, although both images contain extreme amounts of chlorophyll, the August 1999 image appears to contain less. During the decade, populations in the area increased dramatically, demanding more land for residences. The amount of land used for agriculture was reduced during this decade; therefore, agricultural runoff into the lake was also reduced. This is a possible explanation for the decrease in chlorophyll *a* content.

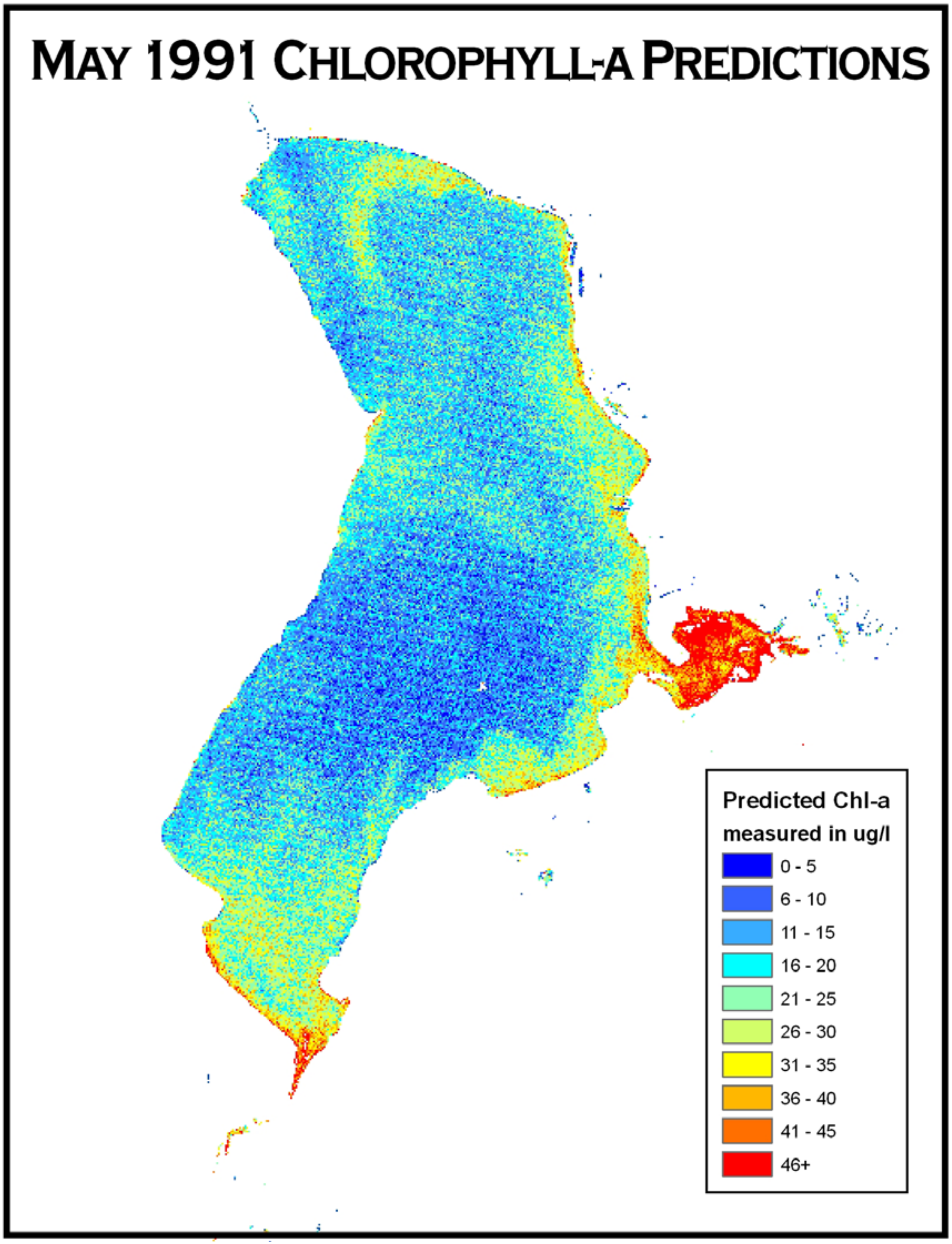


Figure 4.1 May 1991 Cholorphyll-a prediction results

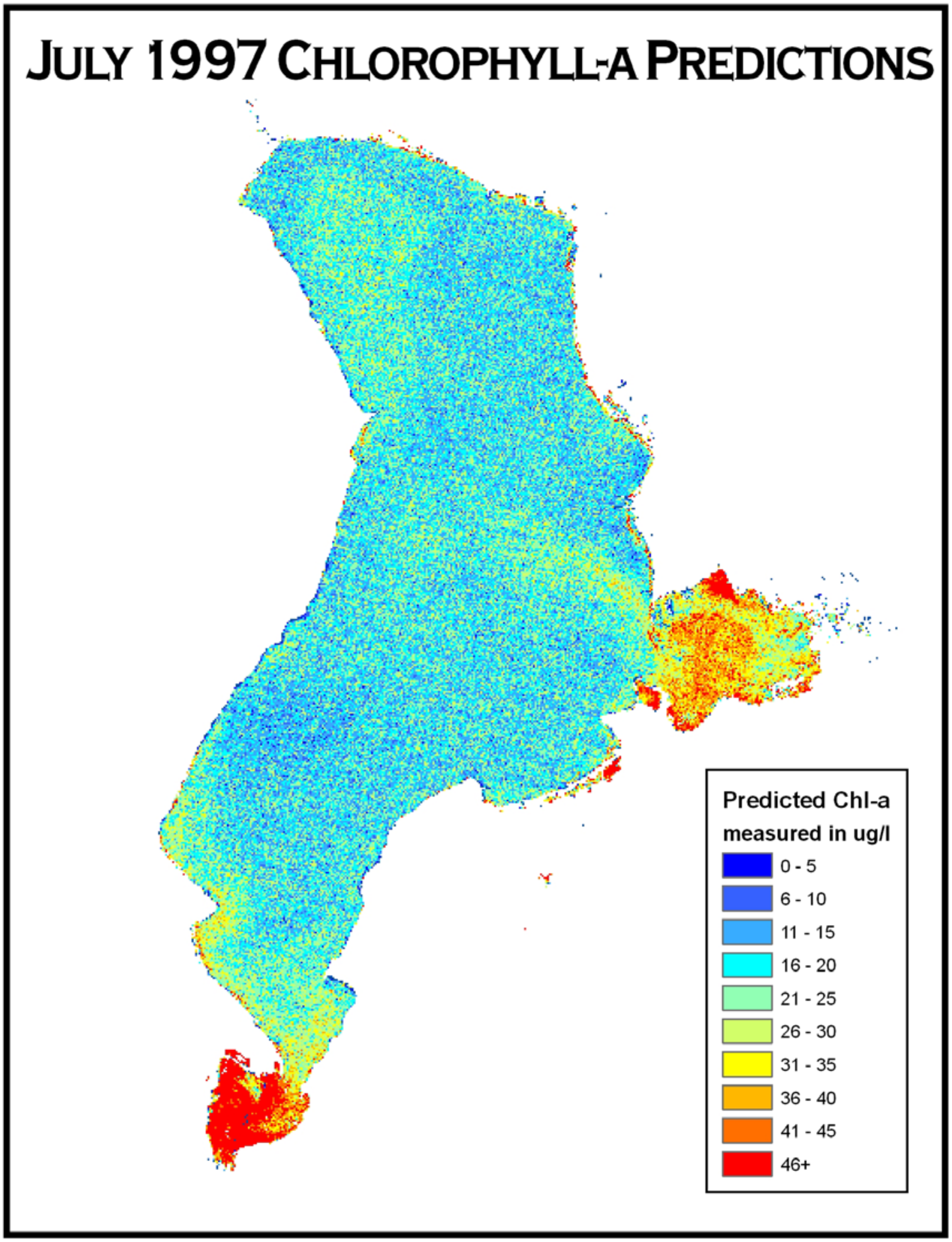


Figure 4.2 July 1997 Chlorophyll-a prediction results

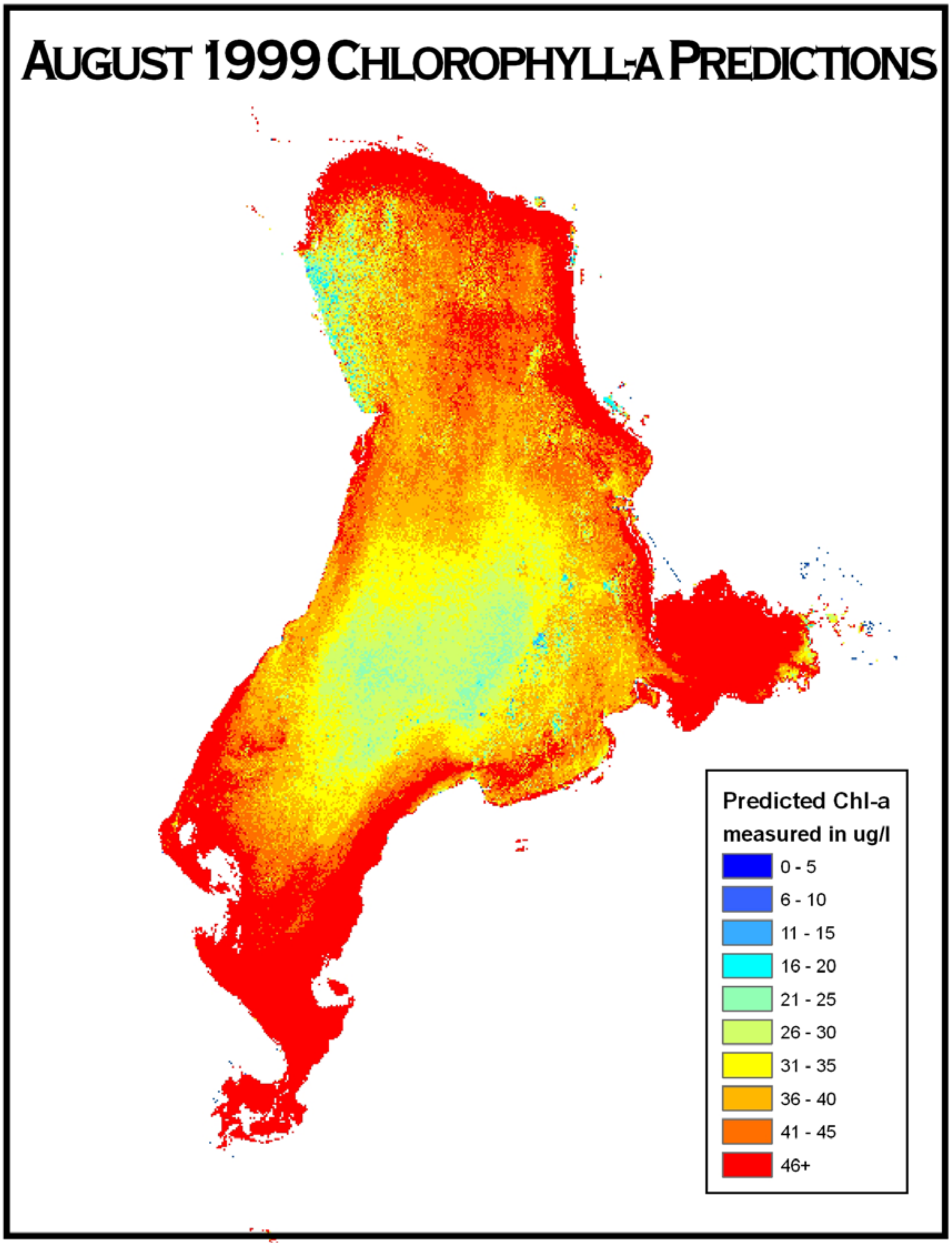


Figure 4.3 August 1999 Chlorophyll-a prediction results

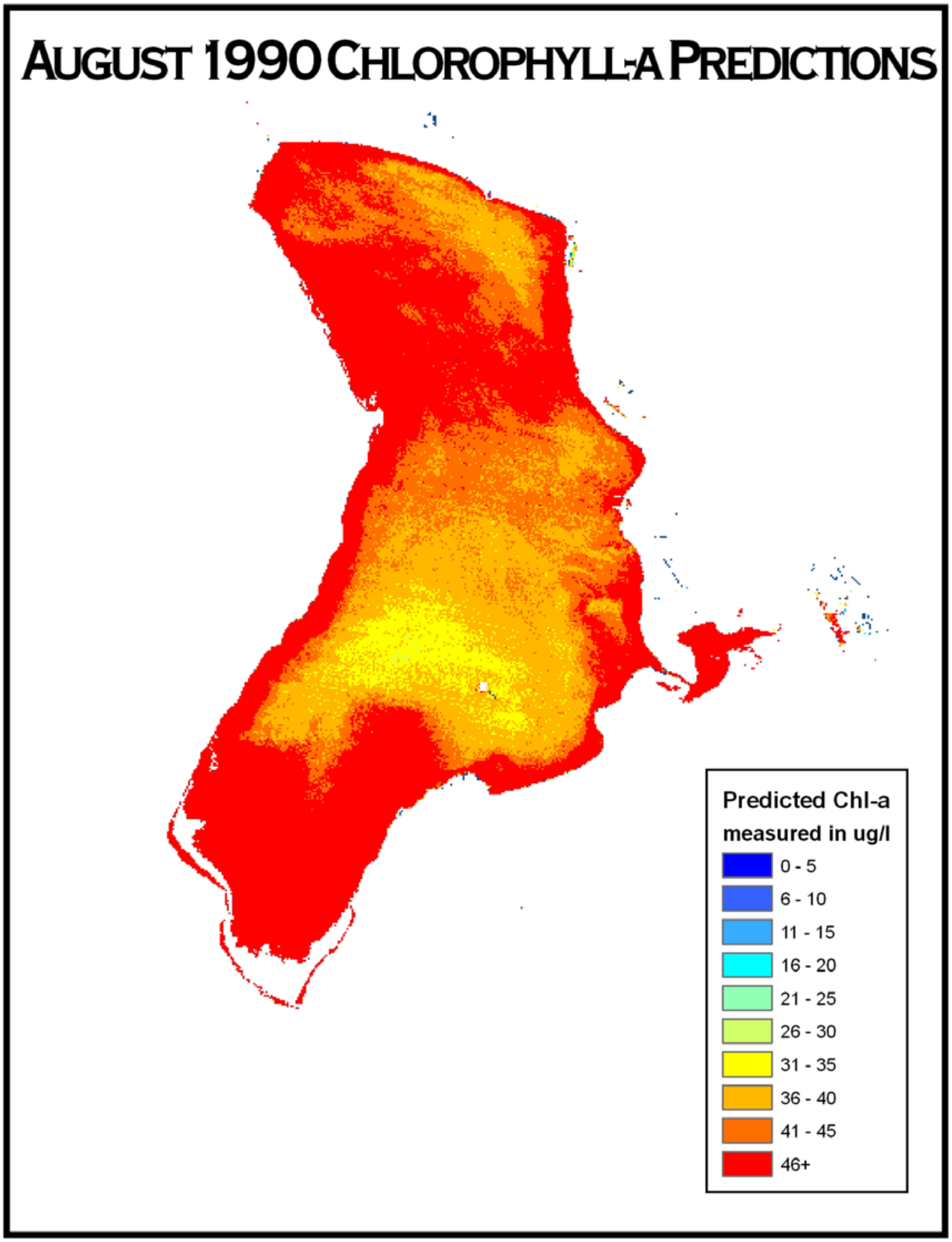


Figure 4.4 August 1990 Chlorophyll-a prediction results

Discussion

With the successful mapping of chlorophyll *a* across the Utah Lake, the objectives of this thesis have been successfully accomplished.

Chlorophyll strongly absorbs radiation at about 450 and 670 nm (Lui *et al.*, 2003). With a high amount of chlorophyll content, the reflectance in green wavelength is increased and reflectance in blue wavelength decreases. As the chlorophyll concentration increases, the peak reflectance shifts from about 680 to 715nm. (Gitelson in Lui *et al.*, 2003). The model created incorporated bands 3, 4, and 7 from the satellite imagery. Band 3 was cubed and therefore transformed to better fit a linear distribution. This band covers a range from 630 – 690nm and therefore the original peak of regular chlorophyll content falls in this range. Due to its shallowness and the relatively large amounts of suspended sediments and accompanying nutrients (i.e. phosphorous and nitrogen), Utah Lake is prone to extremely high levels of algal growth, evidenced by chlorophyll *a* concentrations. As these concentrations increase in some areas of the lake, the peak reflectance is then shifted from the range of band 3 and closer to that of band 4, thus allowing it to be an important factor in predicting chlorophyll *a* content. The importance of band 7 in the equation was surprising. No previous research reviewed has incorporated band 7 in the final equation to predict chlorophyll concentrations. Lyon *et al.* (1988) did a study to determine suspended sediment concentrations. They chose not to use chlorophyll *a* as a parameter based on the fact that their study area contained high concentrations of suspended sediment and it obscured any contribution of chlorophyll. Utah Lake is an extremely shallow lake which allows for large amount of turbidity to result from even small winds. It is possible that band 7 is representing other water parameters that are mixing with the chlorophyll *a*.

As the model separates the water from the land, additional information about the surface area of the lake for any given year can be determined as shown in Table 4.1. This allows easy and quick comparison of the level of the lake as it fluctuates from low levels due to the common droughts in Utah to high levels when water is more abundant. As a fairly shallow lake, the surface area can vary greatly from year to year.

Utah Lake Surface Area	
Image/Year	**Acres**
August 1990	75,044
May 1991	83,300
July 1997	93,065
August 1999	87,509

Table 4.1 Estimated Utah Lake surface area

Limitations

The largest limitation to this type of model, regardless of the water constituent that is being measured and monitored, is the lack of *in situ* data. The small sample size in this research is a limitation that cannot be ignored. Any attempt to represent such a large lake would require as much field data as possible to validate the results. A small sample size when used in a regression analysis makes it difficult to place significant confidence in the results. It is difficult to generalize data over a large population when it is derived from a very small sample of that population. Additionally, when only a few samples are available, all of the samples are needed to develop a regression equation, leaving none for an accuracy assessment.

The use of a 5-day window between the *in situ* samples and the satellite imagery may have introduced unwanted errors. Although it is a small window, weather changes within this time period could easily change the water conditions so they no longer reflect the same pollution that was gathered in field samples just days earlier. Optimally, a sufficient number of samples should be gathered on the same day – and possibly the same time – as the satellite flyover.

Bathymetry data would have also been useful in this study. Such data would better explain some of the different reflectance across the lake and would be a great variable for the multiple regression analysis.

Summary

This thesis proposed three research objectives concerning the use of remote sensing for measuring water quality. These research objectives are:

- Develop a universal model to measure certain characteristics (chlorophyll *a*) through remote sensing of the water of Utah Lake across time.
- Develop a model to automate the measurement of these parameters for different years to be mapped and compared.
- Generate a map from the information derived from these models that would provide insight to the distribution of pollutants and contaminants in Utah Lake.
- Determine an adequate sample size from this pilot study for larger research on determining chlorophyll *a* content of Utah Lake.

Each of these research objectives were completed successfully. A model was created that automated the prediction of chlorophyll *a* across Utah Lake. The results from this research confirm that remote sensing can be used effectively to determine water characteristics of Utah Lake. As shown in the previous figures, successful maps were created to easily identify those areas of the lake where the chlorophyll *a* content does not meet good water quality standards. This pilot study was also used to successfully determine the number of samples necessary for more in depth research.

There are many water quality parameters that are tested each year in Utah Lake; unfortunately, each parameter requires separate research and measuring. Future additions to this thesis may include similar research of chlorophyll *a* with a much greater number of samples taken on the same day as the satellite imagery. Also, models could be developed for many of the other water quality parameters that are found to be unsatisfactory in Utah Lake.

This pilot study of Utah Lake was successful in using remote sensing to measure the amount of chlorophyll *a* in the lake. Such technology, as well as that gathered from future research, could be very useful in monitoring the water quality of Utah Lake and ultimately restoring lake to conditions necessary to meet water quality standards.

REFERENCES

Ahern, F., Goodneough, D., Jain, S., Rao, V., and Rochon, G. (1977). Use of clear lakes as a standard reflector for atmospheric measurements. *Proceedings of the Eleventh International Symposium of Remote Sensing of the Environment*, Vol II, 731-751.

Allee, R. and Johnson, J. (1999). Use of satellite imagery to estimate surface chlorophyll *a* and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA. *International Journal of Remote Sensing*, 29 (6), 1057-1072.

Arizona Department of Environmental Quality (ADEQ). (2005, August). *Water Quality Standards for Surface Waters.* Retrieved November 7, 2005, from the Arizona Department of Environmental Quality Web site: http://www.azdeq.gov/environ/water/standards/index.html

Boyd, C. (2000). *Water Quality: An Introduction.* Norwell, MA: Kluwer Academic Publishers.

Buchner, A., Erdfelder, E., & Faul, F. (1997). How to Use G*Power [WWW document]. URL http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/how_to_use_gpower.html

Carpenter, D. and Carpenter S. (1983). Modeling Inland Water Quality Using Landsat Data. *Remote Sensing of Environment*, 13, 345-352.

Christopherson, R. (2004). *Elemental Geosystems* (4th ed.). New Jersey: Pearson Education.

CIA. (2005). *The World Fact Book*. Retrieved Nov. 8, 2005 from www.cia.gov/cia/publications/factbook

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Erlbaun.

The Columbia Electronic Encyclopedia (6th ed.). Columbia University Press (2005). Retrieved Nov. 8, 2005 from www.infoplease.com.

Department of Environmental Quality Division of Water Quality (DWQ). (2002, September). *Utah Water Quality Assessment Report to Congress (305(b) Report).* Retrieved March 8, 2005, from the Water Quality Web site: http://www.waterquality.utah.gov

Department of Environmental Quality Division of Water Quality (DWQ). (2004, April). *Utah's Year 2004 303(b) List of Waters.* Retrieved March 8, 2005, from the Water Quality Web site: http://www.waterquality.utah.gov

EPA. (1994) *National water quality inventory: 1992 report to Congress.* EPA 841-R-94-001. U.S. Environmental Protection Agency. Office of Water. Washington D.C.

EPA. (2000, December). *Ambient Water Quality Criteria Reommendations: Lakes and Reservoirs in Nutrient Ecoregion II.* Retrieved October 22, 2005, from the EPA Web site: http://www.epa.gov/waterscience/criteria/nutrient/ecoregions/lakes/index.html

Erdfelder, E., Faul, F., & Buchner, A. (1996). G*Power: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*, 1-11.

Forster, B., Sha, X., and Xu, B. (1993). Remote sensing of sea water quality parameters using Landsat-TM. *International Journal of Remote Sensing*, 14 (15), 2759-71.

Fraser, R. (1998). Multispectral remote sensing of turbidity among Nebraska Sand Hills lakes. *International Journal of Remote Sensing*, 19 (15), 3011-3016.

Hadjimitsis, D., Clayton, C., and Hope, V. (2004). An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs. *Remote Sensing of Environment*, 25 (18), 3651-3674.

Hair, J. Jr., Anderson, R., Tatham, R., and Black W. (1998). *Multivariate Data Analysis* (5th ed.). New Jersey: Prentice Hall.

Harrington, J., Jr. and Schiebe, F. (1992). Remote Sensing of Lake Chicot, Arkansas: Monitoring Suspended Sediments, Turbidity, and Secchi Depth with Landsat MSS Data. *Remote Sensing of Environment*, 39, 15-27.

Hoogenboom, H., Dekker, A., and Althuis IJ. (1998). Simulation of AVIRIS Sensitivity for Detecting Chlorophyll over Coastal and Inland Waters. *Remote Sensing of Environment*, 65, 333-340.

Jensen, J. (2000). *Remote Sensing of the Environment: An Earth Resource Perspective*. Upper Saddle River, NJ: Prentice Hall.

Jensen, J. (2004). *Introductory Digital Image Processing: A Remote Sensing Perspective* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Kondratyev, K., Pozdnyakov, D. and Pettersson, L. (1998). Water quality remote sensing in the visible spectrum. *International Journal of Remote Sensing*, 19 (5), 957-979.

Lavery, P., Pattiaratchi, C., Wyllie, A. and Hick, P. (1993). Water Quality Monitoring in Estuarine Waters Using the Landsat Thematic Mapper. *Remote Sensing of Environment*, 46, 268-280.

Liu, Y., Islam, M. and Gao, J. (2003). Quantification of shallow water quality parameters by means of remote sensing. *Progress in Physical Geography*, 27 (1), 24-43.

Lyon, J., Bedford, K., Yen, C., Lee, D. and Mark, D. (1988). Determinations of Suspended Sediment Concentrations from Multiple Day Landsat and AVHRR Data. *Remote Sensing of Environment*, 25, 107-115.

Mayo, M., Gitelson, A., Yacobi, Y., Ben-Avraham, Z. (1995). Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data. *International Journal of Remote Sensing*, 16 (1), 175-182.

Nardi, E. (2005, March 26). Saratoga Springs approves plan for water treatment plant. *Daily Herald*, retrieved October 18, 2005 from the Daily Herald website: http://archive.harktheherald.com.

Novo, E., Steffen, C. and Braga, C. (1991). Results of a Laboratory Experiment Relating Spectral Reflectance to Total Suspended Solids. *Remote Sensing of Environment*, 36, 67-72.

Perry, M. and Mackun, P. (2001). Population Change and Distribution. *US Census Bureau April 2001,* retrieved March 14, 2004 from the US Census Bureau website: http://www.census.gov/prod/2001pubs/c2kbr01-2.pdf.

Porteus, L. (2005, September 29). Google Earth Rocks Satellite Imagery World. *Fox News*, retrieved October 17, 2005 from the Fox News website: http://www.foxnews.com.

Remote Sensing Water Resources Management. (1973).Urbana, IL: American Water Resources Association.

Rimmer, J.C., Collins, M.B. and Pattiaratchi, C. (1987). Remote sensing of water quality in active to inactive cooling water reservoirs. *International Journal of Remote Sensing* 8, 85-102.

Ritchie, J., Cooper, C. and Yongqing, J. (1987). Using Landsat Multispectral Scanner Data to Estimate Suspended Sediments in Moon Lake, Mississippi. *Remote Sensing of Environment*, 23, 65-81.

Sheffner, E. (1999). Welcome to the Landsat Program. Retrieved January 13, 2005, from the Nasa Web site: http://geo.arc.nasa.gov/sge/landsat/landsat.html.

Shiozawa, Dennis K. (1977). The Microdistribution and Population Trends of Larval Tanypus Stellatus Coquillett and Chrionomus Frommeri Atchley and Martin (Diptera: Chironomidae) in Utah Lake, Utah. *Ecology,* 58, 610-618.

Song, C., Woodcock, C., Seto, K., Lenney, M., and Macomber S. (2001). Classification and Change Detection Using Landsat TM Data. *Remote Sensing of Environment*, 75 (2), 230-244.

Thiemann, S. and Kaufmann, H. (2000). Determination of Chlorophyll Content and Trophic State of Lakes Using Field Spectrometer and IRS-1C Satellite Data in the Mecklenburg Lake District, Germany. *Remote Sensing of Environment*, 73, 227-235.

UDEQ (2000). Year 2000 305(b) Assessment. Retrieved January 9, 2005, from Utah Division of Water Quality Web site: http://waterquality.utah.gov/documents/DOC_RULE.HTM.

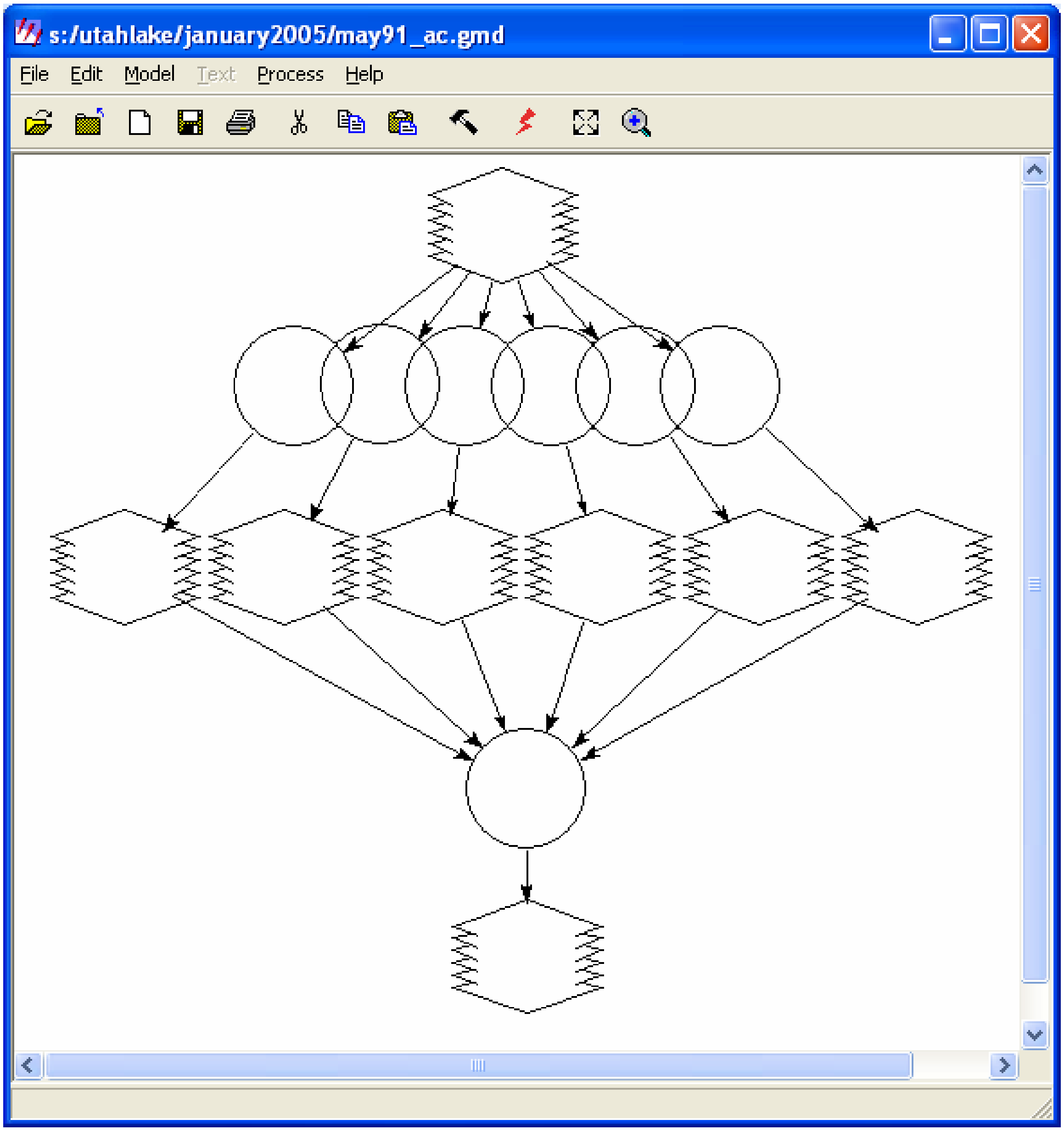
Utah County (2004) Utah County Demographics. Retrieved April 2, 2004 from the Utah County Web site: http://www.co.utah.ut.us.

Vigil, K. (2003). *Clean Water: An Introduction to Water Quality and Water Pollution Control* (2nd ed.). Corvallis: Oregon State University Press.

Warnock, C. (2003, April 17). Group pushes for a Utah Lake Authority. *Daily Herald*, retrieved October 18, 2005 from the Daily Herald website: http://archive.harktheherald.com.

Zilioli, E. and Brivio, P.A. (1997). The satellite derived optical information for the comparative assessment of lacustrine water quality. *Science of the Total Environment* 196, 229-45.

APPENDIX A: ATMOSPHERIC CORRECTION

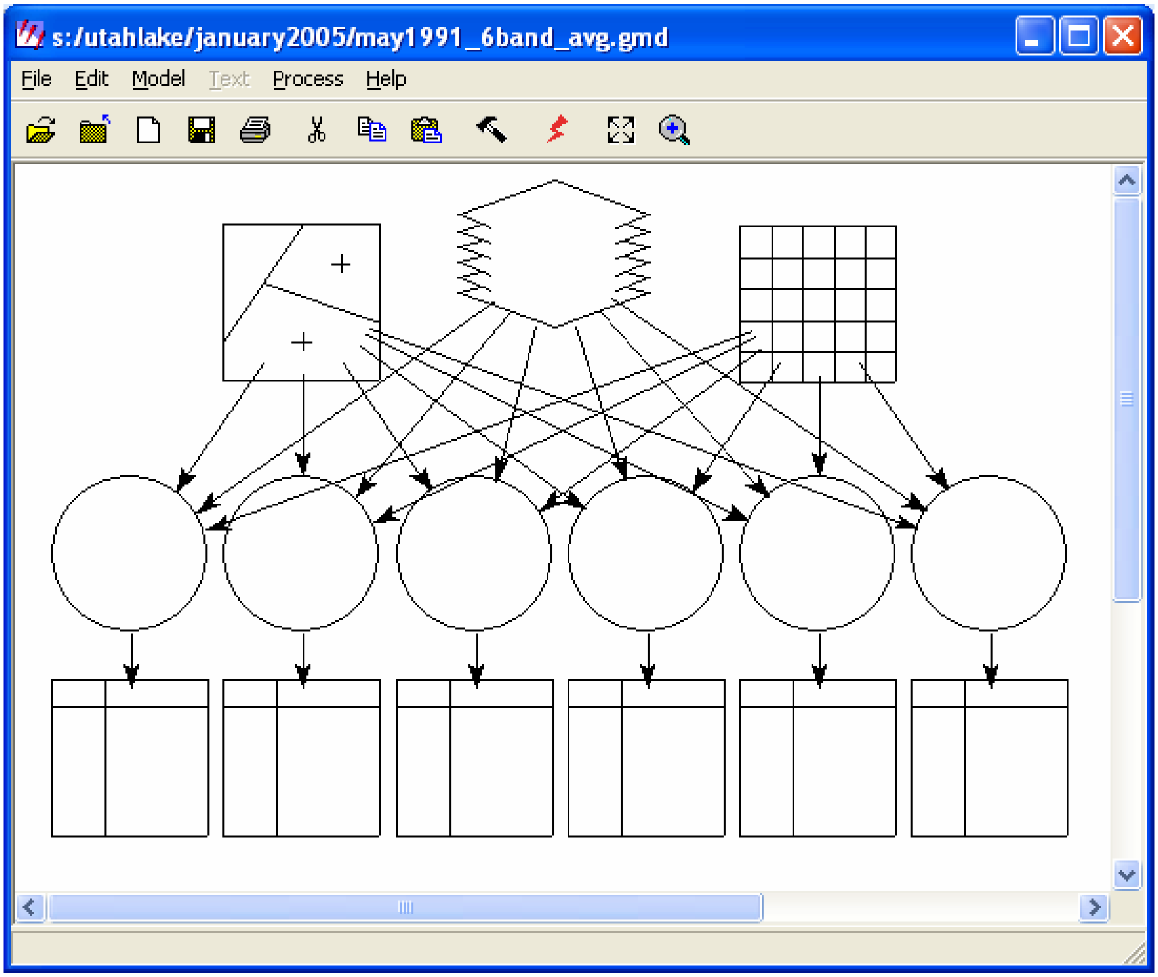


Model Script:

```
COMMENT "Generated from graphical model:
      s:/utahlake/january2005/may91_ac.gmd";
#
# set cell size for the model
#
SET CELLSIZE MIN;
#
# set window for the model
#
```

```
SET WINDOW UNION;
#
# set area of interest for the model
#
SET AOI NONE;
#
# declarations
#
Integer RASTER n1_28may91r FILE OLD NEAREST NEIGHBOR AOI NONE
      "q:/classdata/414/ut_cnty/28may91r.img";
Float RASTER n14_may91_atcorrection FILE DELETE_IF_EXISTING USEALL
      ATHEMATIC FLOAT DOUBLE
      "s:/utahlake/january2005/newones/may91_atcorrection.img";
#
# function definitions
#
#define n13_memory Float((1.004 * $n1_28may91r(1)) -2.573)
#define n12_memory Float((1.010 * $n1_28may91r(2)) -3.296)
#define n11_memory Float((0.963 * $n1_28may91r(3))  +0.065)
#define n10_memory Float((.897 * $n1_28may91r(4)) + 4.625)
#define n9_memory Float((0.964 * $n1_28may91r(5)) + 2.622)
#define n8_memory Float((0.981* $n1_28may91r(7)) + 3.554)
n14_may91_atcorrection = STACKLAYERS
      ($n13_memory,$n12_memory,$n11_memory,$n10_memory,$n9_memory,$n8_m
      emory ) ;
QUIT;
```

APPENDIX B: EXTRACT BRIGHTNESS VALUES FOR SAMPLE LOCATIONS

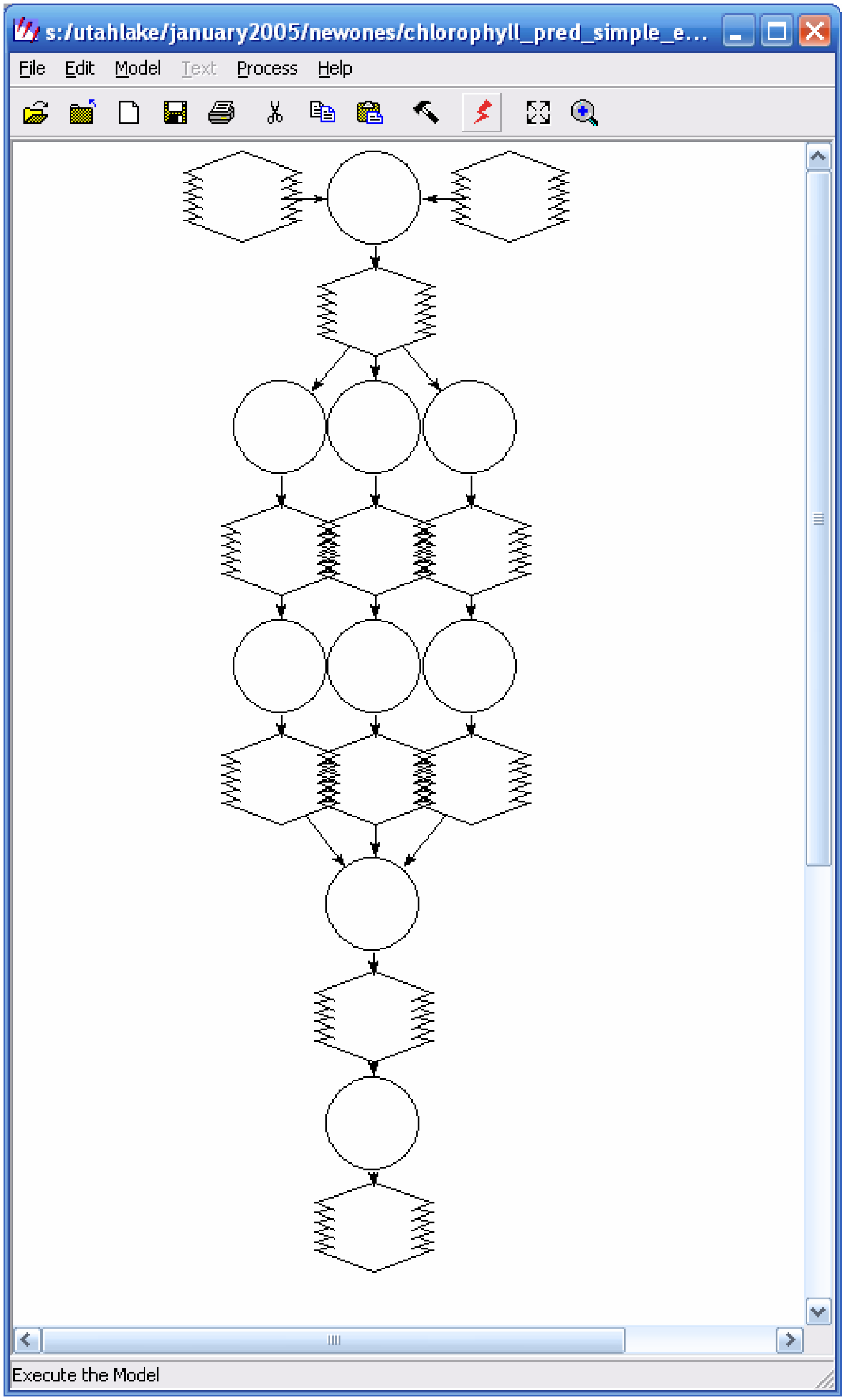


Model Script:

```
COMMENT "Generated from graphical model:
s:/utahlake/january2005/may1991_6band_avg.gmd";
#
# set cell size for the model
#
SET CELLSIZE MIN;
#
# set window for the model
#
SET WINDOW UNION;
#
# set area of interest for the model
#
SET AOI NONE;
#
# declarations
#
```

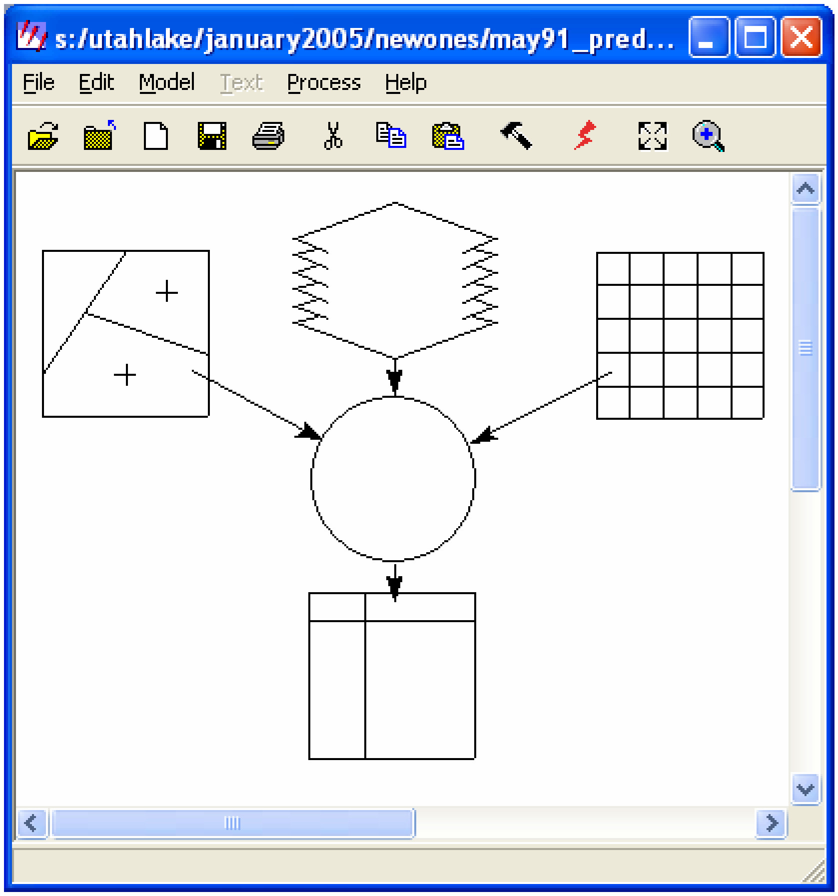
```
Integer RASTER n41_may91_atcorrection FILE OLD NEAREST NEIGHBOR AOI
NONE "y:/tiana davis/may91_atcorrection.img";
Integer VECTOR n44_may1991 COVER AOI NONE POINT RENDER TO MEMORY
"s:/utahlake/january2005/may1991";
FLOAT MATRIX n45_Custom_Float;
VECTOR n42_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n42_Output ATTRIBUTE $n42_layer :: "MAY91_B3";
VECTOR n52_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n52_Output ATTRIBUTE $n52_layer :: "MAY91_B4";
VECTOR n53_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n53_Output ATTRIBUTE $n53_layer :: "MAY91_B5";
VECTOR n54_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n54_Output ATTRIBUTE $n54_layer :: "MAY91_B7";
VECTOR n55_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n55_Output ATTRIBUTE $n55_layer :: "MAY91_B2";
VECTOR n56_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n56_Output ATTRIBUTE $n56_layer :: "MAY91_B1";
#
# load matrix n45_Custom_Float
#
n45_Custom_Float = MATRIX(5, 5:
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1);
#
# function definitions
#
n56_Output = ZONAL MEAN ( $n44_may1991,(FOCAL MEAN (
$n41_may91_atcorrection(1), $n45_Custom_Float)  ) );
n55_Output = ZONAL MEAN ( $n44_may1991,(FOCAL MEAN (
$n41_may91_atcorrection(2), $n45_Custom_Float)  ) );
n54_Output = ZONAL MEAN ( $n44_may1991,(FOCAL MEAN (
$n41_may91_atcorrection(6), $n45_Custom_Float)  ) );
n53_Output = ZONAL MEAN ( $n44_may1991,(FOCAL MEAN (
$n41_may91_atcorrection(5), $n45_Custom_Float)  ) );
n52_Output = ZONAL MEAN ( $n44_may1991,(FOCAL MEAN (
$n41_may91_atcorrection(4), $n45_Custom_Float)  ) );
n42_Output = ZONAL MEAN ( $n44_may1991,(FOCAL MEAN (
$n41_may91_atcorrection(3), $n45_Custom_Float)  ) );
QUIT;
```

APPENDIX C: CHLOROPHYLL PREDICTIONS



```
                COMMENT "Generated from graphical model:
    s:\utahlake\january2005\newones\chlorophyll_pred_simple_eq.gmd";
#
# set cell size for the model
#
SET CELLSIZE MIN;
#
# set window for the model
#
SET WINDOW UNION;
#
# set area of interest for the model
#
SET AOI NONE;
#
# declarations
#
Integer RASTER n1_15jul97lake FILE OLD NEAREST NEIGHBOR AOI NONE
"s:/utahlake/january2005/newones/15jul97lake.img";
Float RASTER n3_temp;
Integer RASTER n10_2002_dem_lake FILE OLD NEAREST NEIGHBOR AOI NONE
"s:/utahlake/january2005/newones/2002_dem_lake.img";
Float RASTER n84_jul97_simple_eq FILE DELETE_IF_EXISTING USEALL
ATHEMATIC FLOAT DOUBLE
"s:/utahlake/january2005/newones/jul97_simple_eq.img";
#
# function definitions
#
n3_temp = EITHER $n1_15jul97lake IF ( $n1_15jul97lake(4)<40 and
$n10_2002_dem_lake<1375) OR -999 OTHERWISE ;
#define n25_memory Float(EITHER -999 IF ( $n3_temp(4)==-999 ) OR
$n3_temp(4) OTHERWISE )
#define n64_memory Float(EITHER -999 IF ( $n25_memory==-999 ) OR
$n25_memory * 2.141 OTHERWISE )
#define n23_memory Float(EITHER -999 IF ( $n3_temp(3)==-999 ) OR
$n3_temp(3) ** 3 OTHERWISE )
#define n62_memory Float(EITHER -999 IF ( $n23_memory==-999 ) OR
$n23_memory * -0.0000406 OTHERWISE )
#define n19_memory Float(EITHER -999 IF ( $n3_temp(6)==-999 ) OR
$n3_temp(7) OTHERWISE )
#define n66_memory Float(EITHER -999 IF ( $n19_memory==-999 ) OR
$n19_memory * -4.253 OTHERWISE )
#define n82_memory Float(EITHER -999 IF ( $n66_memory==-999 ) OR
($n62_memory + $n64_memory + $n66_memory) OTHERWISE )
n84_jul97_simple_eq = EITHER -999 IF ( $n82_memory==-999 ) OR
$n82_memory + 9.566 OTHERWISE ;
QUIT;
```

APPENDIX D: EXTRACT PREDICTED VALUES



Model Script:

```
COMMENT "Generated from graphical model:
s:/utahlake/january2005/newones/may91_predictions.gmd";
#
# set cell size for the model
#
SET CELLSIZE MIN;
#
# set window for the model
#
SET WINDOW UNION;
#
# set area of interest for the model
#
SET AOI NONE;
#
# declarations
#
Float RASTER n41_may91_simple_eq FILE OLD NEAREST NEIGHBOR AOI NONE
"s:/utahlake/january2005/newones/may91_simple_eq.img";
Integer VECTOR n44_may1991 COVER AOI NONE POINT RENDER TO MEMORY
"s:/utahlake/january2005/may1991";
```

```
FLOAT MATRIX n45_Custom_Float;
VECTOR n52_layer COVER POINT "s:/utahlake/january2005/may1991";
FLOAT TABLE n52_Output ATTRIBUTE $n52_layer :: "AC_pred91_eq";
#
# load matrix n45_Custom_Float
#
n45_Custom_Float = MATRIX(5, 5:
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1,
      1, 1, 1, 1, 1);
#
# function definitions
#
n52_Output = ZONAL MEAN ( $n44_may91,(FOCAL MEAN (
$n41_may91_predictions, $n45_Custom_Float)  ) );
QUIT;
```