SPATIAL ANALYSIS OF WATER QUALITY AND EUTROPHICATION CONTROLS IN LAKE NAIVASHA KENYA

PATRICK MCLEAN

MARCH 2001

SPATIAL ANALYSIS OF WATER QUALITY AND EUTROPHICATION CONTROLS IN LAKE NAIVASHA KENYA

By

PATRICK MCLEAN

Thesis submitted to the International Institute for Aerospace Survey and Earth Sciences in partial fulfillment of the requirements for the degree of Master of Science in Water Resources and Environmental Management.

Degree Assessment Board

Chairman:	Prof. Dr. A.M.J Meijerink (Head WREM, ITC)
External Examiner:	Dr. ir. E.Seyan (Free University Amsterdam, (UrA)
Primary Supervisor:	Dr. ir. Chris Mannaerts (WRES, ITC)
Member:	Dr. Zoltan Verkedy (WRES, ITC)

ITC

INTERNATIONAL INSTITUTE FOR AEROSPACE SURVEY AND EARTH SCIENCES

ENSCHEDE, THE NETHERLANDS

Disclaimer

ABSTRACT

The design of a sampling scheme consists of defining the number, location and sample pattern of sampling sites. When temporal variations in the parameters being measured are relevant, the sampling scheme also must specify sampling frequency. Therefore in its most general form a sampling scheme determines where and how often to sample. Proper design of a sampling scheme will depend in part, on the objectives of the study. An objective common to many limnological surveys is the characterization of the water quality of an entire water body. However, it is frequently assumed that horizontal heterogenieties are insignificant relative to those occurring vertically. As a result, a single deep station is often established. In shallow lakes such as Naivasha which exhibit a high degree of patchiness, such an approach to sampling scheme design might not be appropriate. In this study spatial heterogenieties in a number of water quality variables in Lake Naivasha were examined to determine the optimal sampling locations and to estimate a necessary frequency and density of sampling for long term monitoring.

The method used to determine optimal sampling locations is based on the theory of regionalized variables, and assumes that spatial dependence is expressed quantitatively in the form of the semi-variogram. It assumes also that the maximum standard error of a kriged estimate is a reasonable measure of the goodness of a sampling scheme.

In second instance, the spatial water quality data were then used to evaluate the hydrochemical, nutrient and eutrophication status of the lake. Controls of the eutrophication process in Lake Naivasha were investigated using a dynamic water quality modelling approach.

ACKNOWLEDGEMENTS

To Dr. Chris Mannaerts for his supervision of this study, and his insight to the world of Environmental Hydrology (Maximum Respect due).

To Mr. Remco Dorst without your help the hippos would certainly have had a field day (Yu Respect Concrete).

To all those people in Naivasha who helped during work and play, especially Walter Tanui and the Scatman crew (big it up).

To Steve of La Belle Inn for sample storage.

To Kizito Oponda of the Ol-Karia Geothermal Laboratory for sample analysis.

I would like to extend my thanks to the following people and organizations in Kenya

- WRAP teams both in Nakuru and Nairobi for their support in terms of vehicles, personnel and other logistics during and after the fieldwork.
- Richard Okello of Panafcon for his assistance
- LNRA, KenGen, KWS, Kenya Meteorological Department and Ministry of environment and Natural resources for providing the data.
- Shell, Sulmac, Oserian, Kijabe, Sher agencies, and WWF for their financial support.
- Staff of Wambuku Hotel, Dancing spoons and La Belle in who provided us with the comfort during the fieldwork.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

LIST OF FIGURES

CHAPTER ONE

1. GENERAL INTRODUCTION

Eutrophication-the enrichment of water bodies with plant nutrients, typically nitrogen and phosphorous, and the subsequent effects on water quality and biological structure and function is a process, rather than a state. It represents the aging process of lakes, whereby external sources of nutrients and organic matter accumulate in a lake basin; gradually decreasing the depth of the water body, and increasing it's biological productivity to the point that the lake begins to take on a marsh-like character (Thornton, 1996). Under natural conditions, this process typically takes place over geological time. However, human influences in a drainage basin can greatly accelerate this enrichment process, rapidly diminishing the utility of a water body, sometimes within decades. This latter process, termed cultural eutrophication, can be distinguished from natural eutrophication by the fact that the former is a consequence of natural lake aging, whereas the latter results from human induced imbalances in the biogeochemical cycling of nutritive elements, such as nitrogen and phosphorous (Rast, 1994).

Many Lake Ecosystems are endangered by man's activities. Situations from improper land use, waste heat from power stations, and pollution by oil, pesticides, and herbicides are all causing serious alterations of the natural environment (Park et al, 1984). Probably nutrient enrichment from improper use of fertilizer and from the discharge of sewage is the most critical for lake ecosystems. Eutrophication resulting from nutrient enrichment may improve fishing, but it can also increase the population of undesirable species of fish, and of foul-tasting and odor-producing algae, and produce unsightly algal scum. In severe cases, eutrophication can deplete dissolved oxygen to the point of killing large numbers of fish (Thomann and Mueller, 1987). Unfortunately, total removal of nutrients from sewage is costly; further, as recreational developments spread along lakeshores there is more sewage to treat.

The pressing need for enlightened management of our lake ecosystems is made more critical by the difficulty of predicting the consequences of human induced

perturbations. The lake is a complex system with large numbers of biological, chemical, and physical components. Interactions among these components are characteristically nonlinear and involve intricate feedback loops. Desirable changes in one component of the system may trigger a complex chain of effects, which results in deterioration of the total system (Park et al, 1989).

The complexity of the lake ecosystem and the growing need for proper management of this natural resource has placed a high priority on dynamic water quality modelling approach to the problem. This study presents one such a model developed for a shallow tropical African lake.

1.1 Background and problem statement

Lake Naivasha is shallow fresh Water Lake located some 80 km of the equator at an altitude of 1890 m above sea level in the rift valley of Kenya. The area is generally well accessible (approximately a 100 km north west of the capital city Nairobi) and has become an important center for many activities including tourism and agriculture. Long famous for its bio-diversity, the lake was officially recognized under the Convention on Wetlands of International Importance (Ramsar) in 1985.

In the last 10-15 years the Lake has come under increasing pressure from a variety of water uses. The most notable of these has been agriculture primarily in the form of horticultural production for export markets in Europe. Fresh water drawn from the lake is used extensively to irrigate produce grown around the lake. It is generally well known that fertilizers and synthetic pesticides have been used extensively to ensure 'quality' produce. As a consequence, nutrient enrichment of the lake resulting from the use of these chemicals is a potential danger (Harper, 1996). The threat of nutrient enrichment is further reinforced by continuously declining Lake Levels, which accentuates the problem owing to the shallow nature of the lake. Due to the shallow nature of the lake, a relatively small drop in Lake Levels to produces large changes in area and volume; with the result that even a small amount of pollution gains enormously in significance (Harper et al., 1993).

On this basis, it seems reasonable to assume that major impairments to the quality of the lake water can occur if efforts to control or prevent nutrient enrichment are not taken. There is therefore, a definite need for a mathematical model, which could be used as a diagnostic tool, to study the effects of nutrient enrichment and other human induced perturbations on the lake ecosystem.

1.2 Relevance of research

Many studies have been conducted to assess the trophic conditions of the lake and its inflowing rivers. For example Harper (1991) established that the trophic state of the lake is seriously affected by human intervention in the form of lakeside agriculture through papyrus clearance and the application of fertilizer. Le Cren et al, (1980) found that most of lake Naivasha is moderately eutrophic with a significant portion of the nutrients coming from river inflows. Mavuti (1996) found that the levels of phytoplankton productivity within the lake were increasing. While all previous studies attest to an increase in nutrient levels of the lake, there have been very few attempts at the development of a dynamic water quality model of the lake

1.3 Objectives

The main objective of this study is to develop a dynamic model that will predict trophic state changes in Lake Naivasha. More specific objectives are:

1. To characterize the spatial variations in the water quality parameters in the lake.

2. To design a sampling scheme for lake water quality monitoring

3. To investigate eutrophication controls i.e., determine whether nitrogen is more limiting than phosphorous in controlling the lake's biomass levels.

4. To explain and quantify the influence of external nutrient inputs, light regime and water levels changes on the trophic state of the lake.

1.4 Research questions

The present study attempts to answer the questions listed below;

1. Is there significant horizontal variation in the water quality parameters of Lake Naivasha?

2. How many locations should be sampled to be able to evaluate the overall water quality status of a certain variable?

3. Is nitrogen a more limiting nutrient than phosphorus in Lake Naivasha?

4. Do seasonal water level changes; light regime and external nutrient input impose significant control on the trophic state of Lake Naivasha?

1.5 Research Hypothesis

It is proposed to test the null hypothesis listed below:

1. The horizontal variations in the water quality parameters of Lake Naivasha are insignificant.

2. Phosphorus and nitrogen are equally limiting nutrients in Lake Naivasha.

3. Changes in Lake Water Levels, light regime and external nutrient input do not impose significant control on the trophic state of Lake Naivasha.

1.6 Organization of thesis

The thesis is comprised of six chapters, including an introduction (chapter 1), in chapter (ii) the theory underlying the geostatistical estimation of a regional variable is reviewed. Also included in this chapter, is a review of the theory governing the random sampling formulation.

The methods used to test the proposed hypothesis are outlined in chapter (iii). A description of the study area, and the method used in data collection is also presented in chapter (iii).

The results of the (geo) statistical analysis carried out on the water quality parameters measured in the lake are presented chapter (iv). The sample pattern developed for various eutrophication-related water quality parameters is also presented in chapter (iv).

A description of the dynamic water quality model used to simulate the concentration of chlorophyll-a and nutrients in the lake is given in chapter (v). Also presented in this chapter, are the results of the calibration process carried out on the water quality model.

The results of the scenarios formulated to investigate factors controlling eutrophication are presented, and discussed in chapter (vi). The conclusion and recommendation from the study are also presented in this chapter.

CHAPTER TWO

2. THEORETICAL BACKGROUND

The design of a sampling scheme consists of defining the number, location and sample pattern of sampling sites. When temporal variations in the parameters being measured are relevant, the sampling scheme must also specify the sampling frequency. Therefore in its most general form a sampling scheme determines where and how often to sample (Loaiciga, 1989).

The manner in which a sampling scheme is designed depends in part, on the objective of the study. An objective common to many limnological surveys is the characterization of the water quality of an entire water body. However, it is frequently assumed that horizontal heterogenieties are insignificant relative to those occurring vertically. As a result, a single deep station is often established (Thornton et al, 1982). In lakes such as Naivasha, which exhibit a high degree of patchiness (Harper et al, 1982), such an approach to sampling design would be inappropriate. For instance, if the single sample station was located in a region with significant horizontal variation, then it would be difficult to obtain representative samples, and as a consequence, the lake might be incorrectly characterized. The occurrence of significant horizontal heterogeneity is therefore a major problem in designing a water quality sampling scheme, since it is not possible to sample at a single location and adequately characterize the entire water body (Rast, 1986). The first section of this chapter presents the theoretical basis underlying the approach used in this study to quantify the horizontal variations of water quality parameters in lake Naivasha, and hence determine the sampling density needed to completely characterize its water quality.

2.1 Sampling Density

Depending on the magnitude of the horizontal heterogenieties several sample stations might be needed to completely characterize the water quality of the lake. One approach that could be used to quantify these heterogenieties is Geostatistics. If the magnitude of these variations warranted the use of multiple stations, this method

would also enable the stations to be placed in such a way, that the spacing between them, would relate to the scale of variation of the water quality parameters been monitored. This is achieved by the following rationale; a regional variable is autocorrelated in time or space or both. Thus, a particular water quality parameter can be considered a regional variable if the values of spatially nearby samples are numerically closer than values of samples from remote areas. Geostatistics examines this correlation structure through the use of the semivariogram, which is defined as.

$$2\gamma(h) = E\{[Z(x) - Z(x+h)]^2\} \qquad \text{.....(2.1)}$$

When n observation of a property Z with values z_1, z_2 .....z_n are located in an area the expectation is replaced by an average value, and the semivariogram $2\gamma(h)$ is calculated as

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \{[Z(x) - Z(x+h)]^2\} \qquad \text{.....(2.2)}$$

Where $Z(x)$ and $Z(x+h)$ represents the value of the variable at two different points, the vector h is the lag, representing the distance and direction separating pairs of places in this expression, n represents the number of points pairs. The quantity $\gamma(h)$ is the semivariance at that lag.

To appreciate the concept underlying this process, one may begin by considering a kriged estimate $\hat{Z}(x_0)_i$, which is a weighted linear combination of n values of the regionalized variable given by:

$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i Z_i \qquad \text{.....(2.3)}$$

Where Z_i is the observed value of the variable at the sample point i and λ_i is the weight attached to the value at sample point i. To ensure that the estimate is unbiased, the weights are made to sum to 1.

$$\sum_{i=1}^{n} \lambda_i = 1 \qquad \text{.....(2.4)}$$

And the expected error is

$$E\left[\hat{Z}(x_o) - Z(x_o)\right] = 0$$

The estimation variance is

$$\text{var}\left[\hat{Z}(x_o)\right] = E\left[\left\{\hat{Z}(x_o) - Z(x_o)\right\}^2\right] = 2\sum_{i=1}^{N} \lambda_i \gamma(x_i, x_o) - \sum_{i=1}^{N} \sum_{i=j}^{N} \lambda_i \lambda_j (x_i, x_j) \qquad \text{.....(2.5)}$$

Where $\gamma(x_i, x_j)$ is the semivariance of Z between the data points x_i and x_j and $\gamma(x_i, x_o)$ is the semivariance between the ith data point and the target point x_o. For each kriged estimate there is an associated kriging variance, which can be denoted by $\sigma^2(x_o)$ and which is defined by equation (2.4). The next step in the process is to find the weights that minimize this variance, subject to the condition that they sum to one. This condition obtains when

$$\sum_{i=1}^{N} \lambda_i \gamma(x_i, x_j) + \varphi(x_o) = \gamma(x_i, x_o) \text{ for all } i \qquad \text{.....(2.6)}$$

where φ is the Lagrange parameter associated with the minimization. Equation (2.5) thus become

$$\sigma_k^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_o) + \varphi(x_o) \qquad \text{.....(2.7)}$$

The coefficients, are obtained by solving the set of n+1 simultaneous, equation (2.6) written in matrix form as

$$A\lambda = b \qquad \text{.....(2.8)}$$

where

$$A = \begin{bmatrix} \gamma(x_1,x_1) & \gamma(x_1,x_2) & \dots & \gamma(x_1,x_N) & 1 \\ \gamma(x_2,x_1 & \gamma(x_2,x_2) & \vdots & \gamma(x_2,x_N) & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \gamma(x_N,x_1) & \gamma(x_N,x_2) & \vdots & \gamma(x_N,x_N) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix},$$

$$\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \varphi(x_o) \end{bmatrix},$$

$$b = \begin{bmatrix} \gamma(x_{1,} x_0) \\ \gamma(x_2,x_0) \\ \vdots \\ \gamma(x_N,x_0) \\ 1 \end{bmatrix}.$$

The matrix A is inverted and the weights and the Lagrange multiplier are obtained as

$$\lambda = A^{-1}b \qquad \text{.....(2.9)}$$

The kriging variance is given by

$$\hat{\sigma}^2(x_o) = b^T\lambda \qquad \text{.....(2.10)}$$

An evaluation of equation (2.7) reveals an important feature of the kriging process, the variance depends on the semi-variogram and through it, on the configuration of observation points in relation to the point to be estimated and on the distances among them. It does not depend on the observed values themselves. Consequently, if the

semivariogram is known, the variances for any sampling scheme can be determined before the actual sampling. There are two common approach used to determine sampling density from the semivariogram. Firstly, one could just take 2/3 of the effective range and this would be within correlation scale of the variable. Secondly, following the approach of Burgess *et al* (1981) and McBratney *et al* (1981), equation (2.6) can be solved and the kriging variances and errors determined at the centers of grid cells for a range of sampling intervals by the solution of equation (2.7). A graph of variance against sampling interval can then be plotted and the required sampling interval read from this graph. This would then define the sampling density involving the least effort or cost for a given precision and is thus optimal in this sense. In any case a prerequisite, is a reconnaissance survey from which a semivariogram can be estimated.

In practice, the estimation of the semivariogram involves three basic steps; data gathering, univariate statistical analysis, and semivariogram modeling. Univariate statistical analysis is necessary to determine whether the data set meets the assumptions of the kriging estimator. The semivariogram modeling involves choosing a model type (e.g. linear, spherical exponential etc.) based on the experimental semivariance calculated from the data set and, fitting the model curve to the experimental semivariance values calculated for the data set for different distances between points ('lags'). The shape of the model curve can be modified by changing the values of the nugget, sill and range so that the best fit with the experimental semivariogram is found, resulting in a curve describing the spatial variability of the investigated regional variable. The sill is the flat region of the semivariogram at longer lag distances, which is equivalent to the variance of the data set (figure 2.1). The distance at which the model curve approaches this sill value is known as the range and defines the neighborhood within which all locations are related to each other. In some cases the observed semivariogram does not approach zero with decreasing lag distance and show a non-zero y-intercept known as the nugget effect. This implies a source of variability at a scale smaller than the minimum lag distance, and hence, it cannot be modeled with the present sampling scheme. It is also an indication of the experimental error or data precision.

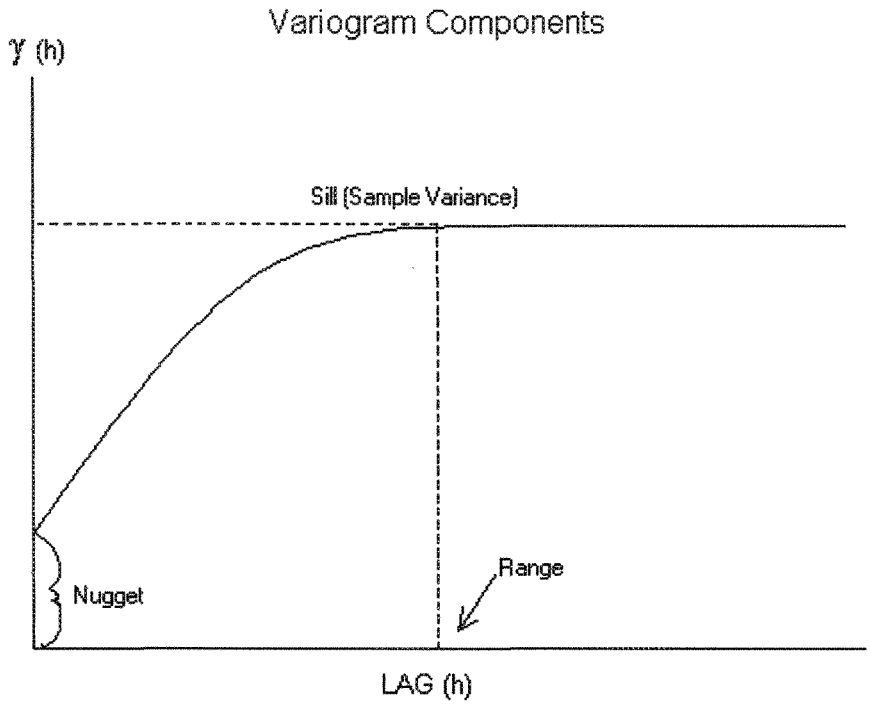


Figure 2.1: shows a typical variogram

2.2 Sampling Pattern

In addition to the number and location of the sampling sites, consideration must also be given to the sample pattern (i.e., the number of samples that should be collected in order to achieve a given precision) at each of these sampling sites. A formula that gives an initial estimates of this sample pattern can be derived by assuming that the observations (e.g. z_1 z_2 z_3 z_4 .....z_n) of a particular water quality parameter are independent and identically distributed observations, drawn from a normal distribution with mean μ and variance σ^2 (Cochran, 1977). The mean of this sample which is given by:

$$\bar{z} = \sum_i^n Z_i / n \qquad \text{.....(2.11)}$$

Would also be normally distributed with expected value μ and variance, which can be estimated using $s^2/\sqrt{n}$. By normalizing one obtains:

$$Z = \frac{z-\mu}{\frac{s}{\sqrt{n}}} \qquad \text{.....(2.12)}$$

Where Z is a standard normal variable with zero mean and unit variance from equation (2.10) the following probability statement holds:

$$P\left(-t_{\alpha/2} \le \frac{\bar{Z}-\mu}{s/\sqrt{n}} \le t\alpha/2\right) = 1-\alpha \qquad \text{.....(2.13)}$$

From which it follows that:

$$p\left(\bar{Z} - t_{\alpha/2}\left(s/\sqrt{n}\right) \le \mu \le t_{\alpha/2}\left(s/\sqrt{n}\right)\right) = 1-\alpha \qquad \text{.....(2.14)}$$

Indicating that a $100(1-\alpha)\%$ confidence interval for the mean is given by $\bar{Z} \pm t_{\alpha/2}\left(s/\sqrt{n}\right)$ where

the error, E in calculating μ by $\overline{Z}$ is:

$$E = Z_{\alpha/2}\left(s/\sqrt{n}\right) \qquad \text{.....(2.15)}$$

From which:

$$n = \left(t_{\alpha/2}\right)^2 s^2 / E^2 \qquad \text{.....(2.16)}$$

From equation (2.15) it can be seen that the number of samples to be collected at each station will depend on the variability of the water quality parameter of interest, and the desired precision of the estimate. Therefore by accepting a lower level of precision the actual number of samples can be reduced.

CHAPTER THREE

3. STUDY AREA, MATERIALS AND METHOD

3.1 Study Area

Lake Naivasha is shallow fresh Water Lake located approximately 100 km Northwest from the capital city of Nairobi in the eastern portion of the rift valley Kenya, (Longitude 0 45' S; latitude 36 20'E) at an altitude of approximately 1890m above mean sea level (figure 3.1). The lake can be classified as continuously warm polymictic (Lewis, 1983) and, is characterized by daily stratification cycle in which a thermal gradient develops during the calm morning hours but is later destroyed by strong afternoon winds and nocturnal cooling (Verschuren, 1982).

Physical Characteristics

3.1.1 Climate

Climatic conditions in the Naivasha basin are quit diverse due to considerable differences in altitude and landform. Although the lake is located within 80 km of the equator and is thus tropical it generally experiences relatively cool conditions (Richardson et al, 1976). Air temperatures are moderate with monthly means varying from 15.9 to 18.5 ^{o}C. The area is semi-arid, receiving on average 640 mm of rain each year, while annual evaporation is approximately 1360 mm. Evaporation exceeds precipitation throughout the year except at peak rainfall. The rainfall trend near the lake has a muted bimodality with a main pulse in April and May and a minor pulse in November (Melack, 1972). Only breezes are common in the morning but afternoon winds in the region of 15 km/h are typical, and often produce violent storms on the lake. The wind direction is usually Southwesterly to easterly throughout the year, becoming more southeasterly during period of rain (Hubble, 1999).

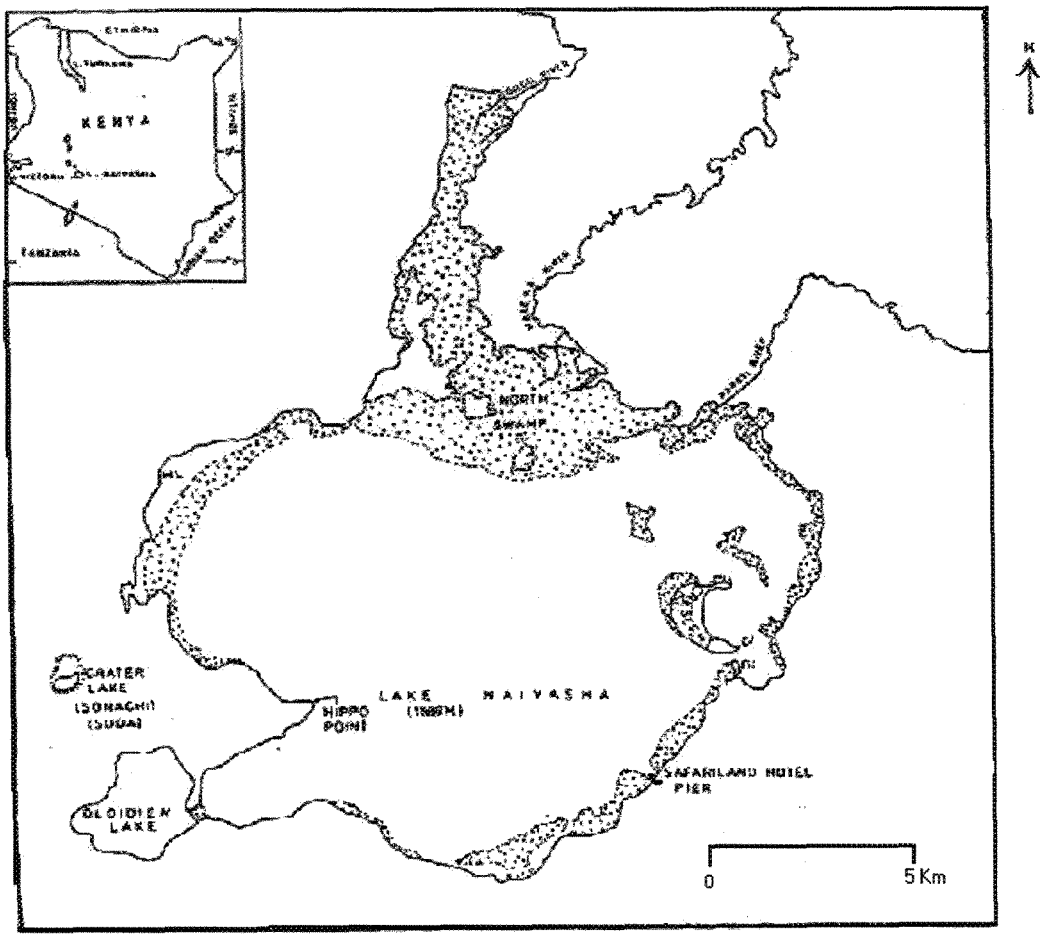


Figure 3.1: Map of Lake Naivasha

3.1.2 Morphometry

Lake Naivasha is a large (surface area = 150 km^2), shallow (mean depth = 4 m) water body composed of three principal water masses (figure 3.1), namely, Main lake (130 km^2), Crescent island bay (2.1 km^2), and Oloiden bay (5.5 km^2). Table 3.1 gives additional morphometric data for the main lake and figure 3.2 presents the bathymetry of the water body. From this figure it can be seen that the lake has a very flat bottom with major decrease in depth only close to the shores. The deepest part of the lake is located near Hippo point on the southwestern section of the lake. A profile of the lake shows that the main lake is flat while the two deepest sections displays a crater like morphology. Melack, (1972) attributes the flatness of the main part of the lake to large amounts of sediment build-up which has resulted in even bottom topography. The two deepest sections, which shows a crater like morphology indicates a volcanic origin.

Table 3.1: Shows morphometric data for the main lake of Naivasha.

Characteristic	Magnitude
Volume	680×10^{6} m^3
Surface area (SA)	130 km^2
Catchment area (CA)	3400 km^2
CA/SA	26.0
Maximum depth	9.5 m
Mean depth	4 m

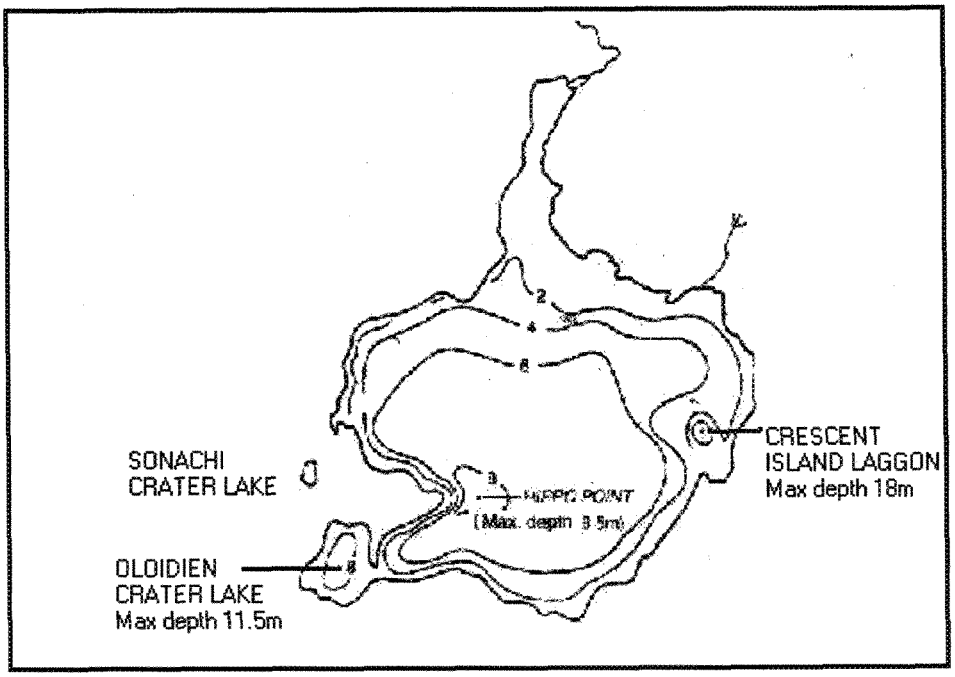


Figure 3.2: Bathymetric map of Lake Naivasha Source: Harper et al (1990).

Catchment Description

3.1.3 Geography

The catchment of the lake encompasses approximately 3,400 km^2 and is defined on the eastern portion by the Aberdare mountains and the Kinangop plateau while to the southwest it is bounded by the Mau encarpment and a combination of the Kiambago and Eburu hills (figure 3.3). The overall nature of nature of the catchment is rural with a variety of ecological zones including a lacustrine wetland.

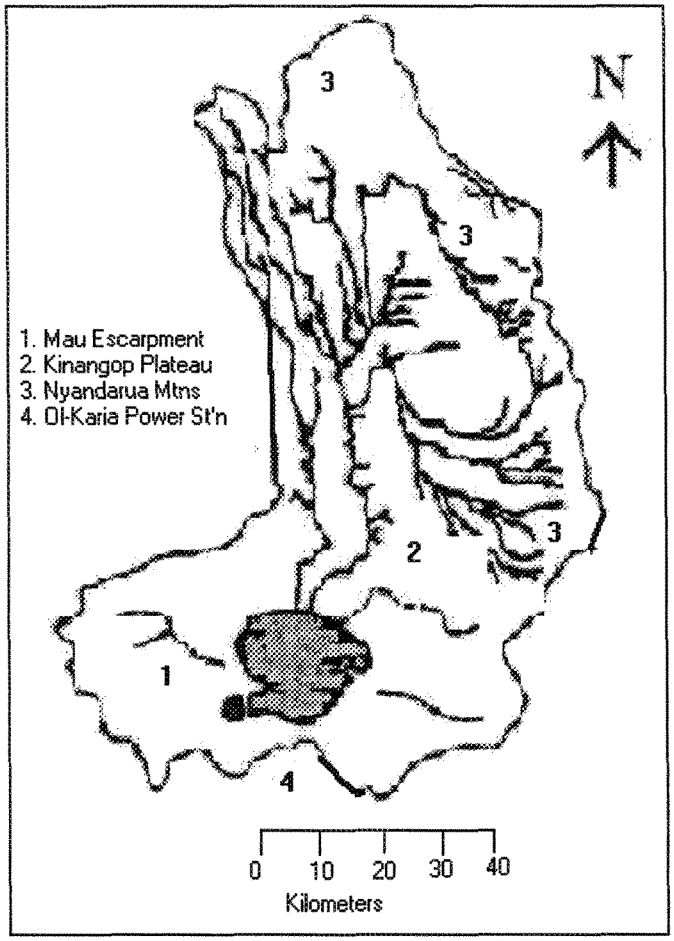


Figure 3.3: Map of Naivasha catchment

3.1.4 Land Usage and Demography

The Naivasha catchment can be divided into six major landuse units namely; agriculture, forest, natural vegetation, grassland, settlement, and open spaces, the last including areas of Naivasha's main lake and other water bodies in addition to bare soil, larva flow and scrub. Agriculture, principally horticulture and vegetable growing accounts for a large percentage of the land usage in the catchment as does dairy farming. Human settlement form a minor class of land usage, but the fact that the Naivasha sewage treatment plant releases untreated effluents directly into the lake weight heavily in terms of the potential impact of sewage discharge on the lake system.

The official human population of Naivasha town in 1995 was 50349, including the associated rural population. Along with 35337 in Gilgil, which gives a total official population of 85686 in the catchment, although this figure will certainly be an underestimate as there is migration into this relatively affluent area, general population increase in Kenya and the problem of accurately performing a census in a developing nation. The 1999 census figures indicates a total population of 112649, plus approximately 100000 uncounted and the population of Gilgil which is currently unavailable (Hubble, 1999).

3.1.5 Hydrology

Much of the rainfall destined for the rift valley is intercepted by the surrounding highlands, resulting in a strongly negative hydrological balance near the lake. The main basin of Lake Naivasha and its Satellite Basins Lake Oloidien and Crescent Island are maintained by river input primarily from the Malewa river, which drains the Kinangop plateau and wet highlands in the Nyandarua range. The remaining input comes from seasonal streams, direct precipitation and ground water seepage. Despite the fact that the drainage basin of Lake Naivasha is topographically closed the main lake itself is hydrologically open. Water lost by ground water seepage is estimated to average 12% of total water lost, the major route being lake-surface evaporation with 80% (Gaudet and Melack, 1981; Darling et al., 1990).

In contrast with the main lake, Oloidien is hydrologically closed. Water-budget calculations indicate that groundwater output is negligible and water losses are due

entirely to evaporation (Gaudet and Melack, 1981). The strongly negative local balance is compensated for by seepage from the main basin into Lake Oloidien through the sill, which separate the two basins (figure 3.1). Crescent Island Crater on the other hand is hydrologically open due to ground water throughflow, with its hydrology being dictated primarily by Ground water seepage and direct exchange of lakewater with the main basin (Verschuren, 1982).

3.2 Materials And Method

In order to test the proposed hypothesis, the following approaches were considered;

3.2.1 Spatial variations in water quality

To test the hypothesis on the spatial variations in the water quality of the lake, a field study was designed to produces a three-dimensional water quality database. Figure (3.4) give a schematic overview of steps involved in the creation of this database, the method used to characterize the spatial variations in a number of water quality parameters, and to determine the optimal sampling density for lake water quality monitoring. These steps are explained below.

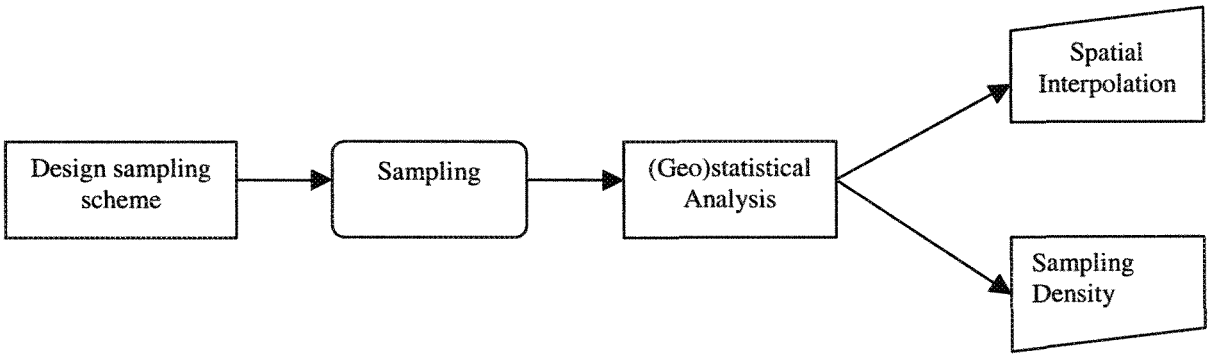


Figure 3.4: Schematic overview of method used to test hypothesis on the spatial variation in water quality parameters.

(a) Sampling scheme design

Owing to time available for fieldwork phase of the study and financial constrains, it was necessary to limit the number of sample points involved in the reconnaissance survey of the lake to 60. As a consequence of this small sample size, a careful trade-off had to be made between two partly conflicting demands:

I. The sampling points should cover the surface area of the lake as evenly as possible, in order to capture the main features of the spatial variations.

II. The accuracy of the experimental variogram depends in part on the number of point pairs per lag class. In general, a sampling scheme that covers the area evenly has very few to zero point pairs at short distances and as a result, yields a poor experimental variogram for these distances.

Due to these conflicting demands a systematic sampling scheme (e.g. a regular square grid) would not have been appropriate. Hence an optimal spatial sampling scheme was designed using spatial simulated annealing (SSA). This algorithm which was developed for the optimization of spatial sampling schemes, includes features such as, the incorporation of preliminary observations, and taking into account sampling constraints and boundaries. Furthermore, SSA allows the use of several quantitative optimization criteria, among them minimization of the maximum kriging variance (MK) and the minimization of the mean shortest distance criterion (MMSD) which aims at an even spreading of the observations over the area of interest of (Van Groeningen and Stein, 1999). For completeness, a description of the algorithm is included in appendix (1).

In this study, use was made of the above optimization criterions to design the sampling scheme in two phases.

I. The first 30 observations were selected according to the MMSD criterion, which aim at the minimization of the mean shortest distance between an arbitrarily chosen point within the study area and its nearest sampling point. In this way, a good coverage of the area with these observations was ensured.

II. Subsequently, the additional 30 observations were selected using the MK criterion in conjunction with a variogram inferred from a previous study (Donia, 1998). The scheme was designed for the estimation of the experimental variogram using the total 60 points.

The SSA algorithm was implemented using the software sanol for windows (Van Groeningen, 2000). A prerequisite for the use of software however, is the presence of a grid based GIS file of the sampling area. Figure (3.5) gives an overview of the steps involved in the development of this GIS file.

Image processing, color composite formation, and geo-referencing were carried out on a recent thematic mapper image of the Naivasha area using standard GIS functions available in Ilwis. The georeferenced color composite was then used as a background map for on screen digitizing of the lake area. The resulting segment map was converted to a polygon map of the lake area. Owing to the fact that Sano1 required a grid based GIS file in text format the polygon map in Ilwis was exported to surfer and them imported to Sano1.

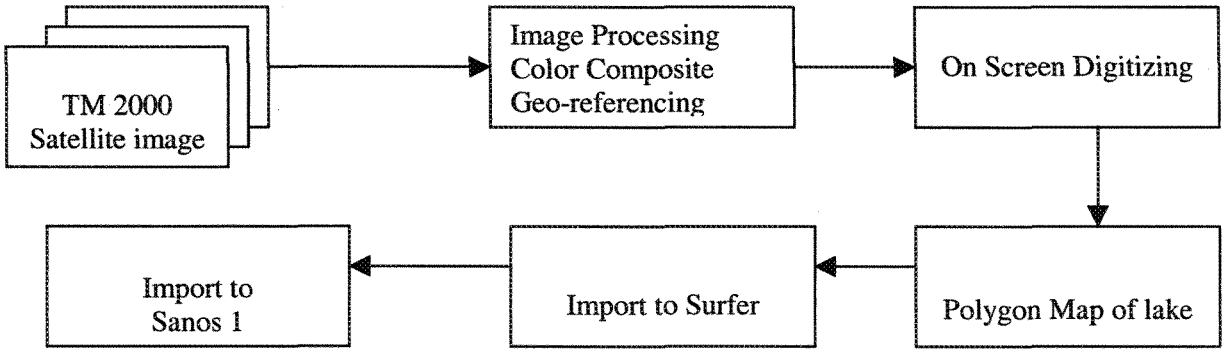


Figure 3.5: Schematic overview of the steps involved in the creation of a GIS over study area.

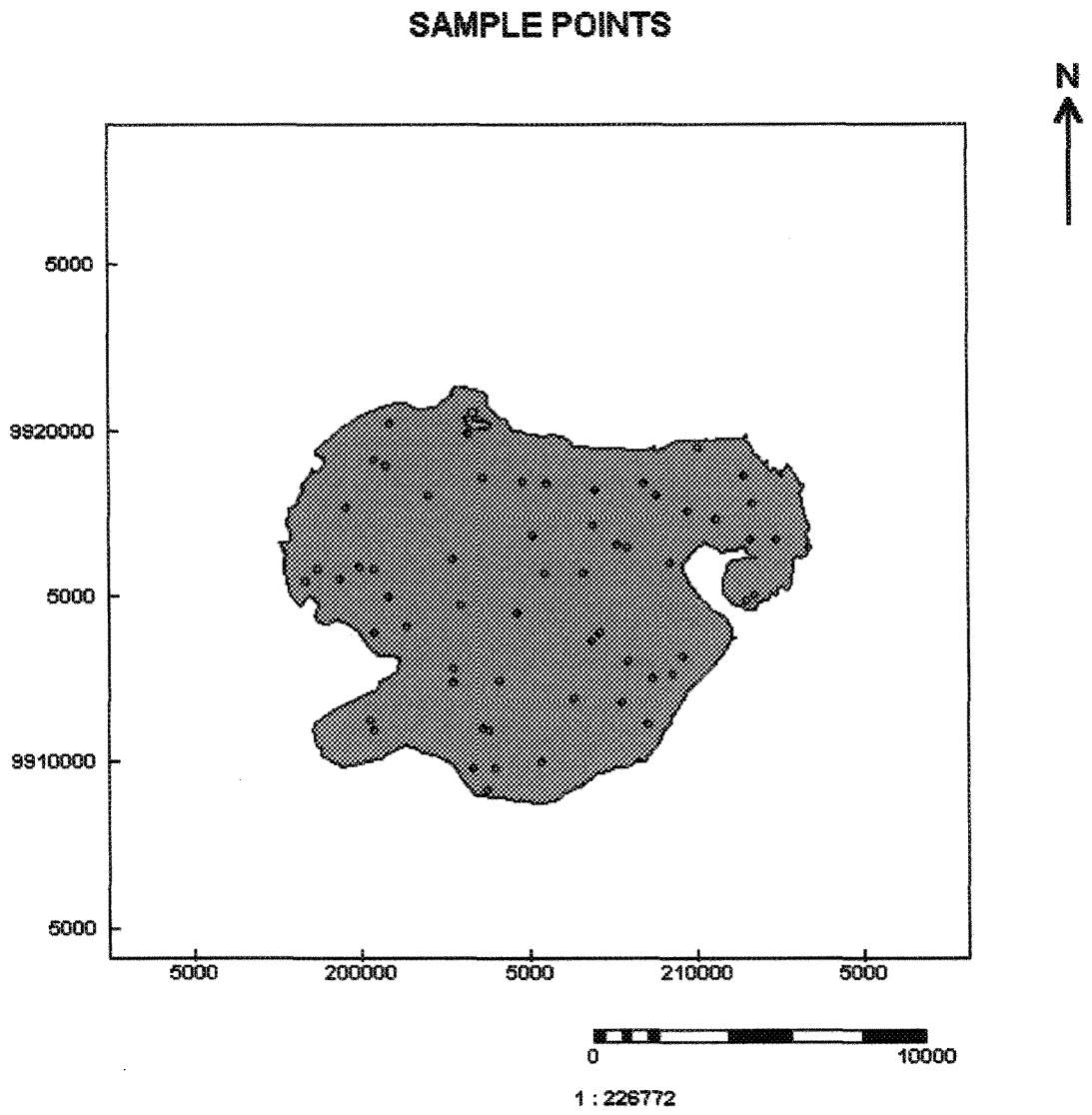


Figure 3.6: Final 60 points used to sample the lake

Figure 3.6 shows the final 60 points selected using SSA algorithm. The coordinates from the above map was then uploaded to a global positioning system unit (GPS), which was used to locate these points on the lake to within ± 7 – 10 meters.

(b) Sampling

The lake was sampled over a three-day period ranging from the 18 – 21 of September 2000. Despite the fact that the majority of sampling was performed on the main lake itself, it was necessary that the inflowing river Malewa be sampled to determine its contribution to the nutrient load of the main lake. The other inflowing rivers, Gilgil and Karati were dry at the time of sampling. The parameters listed below were measured *in situ* at depths of 0.2, 0.5, 1.0, 2.0, 3.0 m, until bottom for each site shown in figure 3.4, and at a depth of 0.5 m for three sites at the entrances of the Malewa into the lake, using a portable multi-electrode-probe (U10 Horiba Instrument Co., USA);

(a) Water temperature (^{0}C)
(b) Dissolved Oxygen (*mg/l*)
(c) pH
(d) Conductivity ($\mu S / cm$)
(e) Turbidity (*mg/l*)
(f) Salinity %
(g) Water depth (m)

The above parameters were also measured at a single site at haft hour intervals, for a fourteen-hour period, in order to obtain a "sense" of their diurnal variations. In addition, depth integrated water samples were also collected at each sample site by the direct dipping of a plastic bucket into the lake. The samples were then transferred to polyethylene bottles and stored cool (4 ^{O}C) in a water cooler and returned to the lab. They were then divided into four sub-samples. The first sub-samples were preserved by acidifying with concentrated nitric acid and were subsequently analyzed for Na, K, Ca, Mg and SiO2 at the nearby geothermal power laboratory. Since the second sub-samples would be analyzed for NO_3, NH_4 and organic nitrogen acidifying with concentrated sulfuric acid (1.8ml/l) and storing at 4° C in a refrigerator preserved them. The third sub-samples were preserved by the addition of $HgCl_2$ (2.8 ml/l) and freezing since they would be analyzed for total inorganic and total organic phosphorus. The final sub-samples were filtered using a vacuum pump assemble and membrane filters. The filters were then wrapped in aluminum foil and placed in seal zip locks plastic bags and stored frozen to be analyzed for chlorophyll-a.

Sediment samples were collected from the southern section of the lake, and the entrance of the Malewa (figure 3.1), using Perspex tubes of length 60 cm and diameter 7 cm, which were attached to PVC pipes and inserted into the lakebed. These samples were analyzed for total inorganic phosphorus, total organic phosphorus and total organic nitrogen.

The transparency of the lake water was determined by the extinction of a 20 cm diameter secchi disc. This is the depth where the disc vanishes, and occurs when irradiance is approximately 5 % of the incident level, from this two calculations were made;

The background light extinction coefficient (1.7/d)
The euphotic depth (2.35d) where 1% incident energy remains,

Where d is the extinction (Secci) depth

(c) Geostatistical Analysis

Descriptive statistics, histograms and normal probability plots were generated for the parameters measured at the surface of the lake. For completeness, chi-square goodness of fit test for normality was also conducted, so as to ascertain if a scale transformation (e.g., the taking of logarithms) would be necessary. In addition, test for stationary of the mean were carried out, by estimating the drift (Shafer et al, 1990) in each data set using:

$$\text{Drift (h)} = \frac{1}{2n(h)}\sum_{i=1}^{n(h)} Z(x+h) - Z(x) \qquad \text{.....(3.1)}$$

Omnidirectional variogram were estimated for each of the above parameters, using all 60 observations of that parameter. Pairs of observations were grouped into distance classes with a lag length of 800 m, yielding 8 distance classes. Equation (2.2) was then used to calculate variogram values. Theoretical variogram models were then fitted to the resulting experimental variogram. Owing to the predominantly Southwesterly to easterly direction of water movement within the lake, there was concern that significant anisotropy might exist in the spatial correlation. Therefore

directional variogram were also computed for each of the above variables. In order to determine the existence or absence of anisotropy within the data sets. All geostatistical analysis was carried out using Ilwis software package.

The total horizontal variations of water quality parameters in the lake were characterized by dividing the observation into groups depending of their depth of measurement. Four such groups were formed, measurement made at 0.2, 1, 2, and 3m. Descriptive statistics including, the coefficient of variation, and range were then generated for each parameter within a group. If the assumption were made that the parameter values approached a normal distribution then, plus or minus two standard deviations would result in 95 percent of the total distribution. Calculating the average standard deviation for each parameter and multiplying it by two resulted in the expected variation for each parameter. Comparing this value to the sensitivity of the sensor used to measure the parameters (Horiba U-10) provided an indication of whether the total variation was significant.

The vertical variations of the water quality parameters, were characterized, by first normalizing the values obtained at each depth for all sampling points, using the corresponding surface values. Descriptive statistics including the coefficient of variation, and range were then generated for these normalized values. The expected variation in each parameter was then calculated, as described above. To ascertain if the above parameters showed significant variation with depth i.e., if the lake was vertically homogenous; a regression of the pooled normalized values by depth was also attempted.

(e) Sampling Density

For water quality parameters that showed a clear indication of spatial structure based on the semivariograms. The optimal distance between sampling points needed to completely characterize the water quality was taken as 2/3 of the maximum range of spatial correlation. This distance was then used to construct square grid, since it was assumed that sampling would take place at the center of these grids.

(d) Interpolation (Kriging)

Kriged maps along with error maps of prediction were generated for all water quality parameters whose variogram showed a clear indication of spatial structure. The maps were created using ordinary block kriging, with the size of the blocks been dictated by the range of spatial variation. All maps were created using the standard GIS functionality of Ilwis.

3.2.2 Surface temperature of lake using TM band 6

As a consistency check on the interpolated temperature map, the surface temperature of the lake was also determined using band 6 of the Landsat TM 2000 image. This band operates between wavelength ranges of 10.4 μm and 12.5 μm of the thermal infrared region. For an earth ambient temperature 300 k, the peak of emitted radiation occurs at about 9.6 μm, as such TM band 6 is ideally suited to map surface temperatures.

The radiant temperature of a given surface can be derived from the spectral data or digital number (DN). The DN values were first converted to spectral radiance using the equation developed by Markham and Baker (1996):

$$L_i = L_{\min,i} + \frac{L_{\max,i} - L_{\min,i}}{DN_{\max}} DN \qquad \text{.....(3.2)}$$

Where Lmax = spectral radiance in band i (mW/cm^2/Str/ μ m)

L_{max} = Spectral radiance at DN = 0 (mW/cm^2/Str/ μ m)

$L_{max,I}$ = Spectral radiance at DN =DN_{max} (mW/cm^2/Str/ μ m)

DN = digital number

DN_{max}= maximum digital number

These radiance values were then converted to the at satellite temperature of the earth-atmosphere system. Using the expression below:

$$T = \frac{K_2}{(K_1 / L_1 + 1)} \qquad \text{.....(3.3)}$$

Where T = at satellite temperature in Kelvin (K)

K_1 = calibration constant = 60.776 (mW/cm^2/Str/ μ m)

K_2 = calibration constant = 1260.5 (mW/cm^2/Str/ μ m)

3.3 Nutrient limitation

In order to test the hypothesis on whether phosphorus (P) was more limiting than nitrogen (N), the level of these two nutrients were determined from water sample collected at nine sites in the lake. These sites were chosen so as to provide a representative coverage of the different physical, and chemical conditions found throughout the lake. Nitrogen levels were measured as ammonium and nitrate ions, which together form the total dissolved nitrogen. Phosphorus levels were measured as orthophosphates with total phosphorus been measured separately. These values were then used to calculate lake wide average values for these two nutrients. The mass ratio of these averages, i.e., average total nitrogen to average orthophosphates (N: P) was used as an indicator of the nutrient that was limiting.

3.4 Eutrophication controls

To test the hypothesis on eutrophication controls the following procedure was undertaken; The trophic state of the lake was characterized by using three trophic states indicator, namely chlorophyll-a (chla), total phosphorus (TP) and secci depth (SD). These parameters were combine to produce a composite trophic state indicator by the following process. Firstly, A literature review was used to establish trophic state boundary for the three parameters. The values selected for chlorophyll were $4\,\mu g/l$ and $10\,\mu g/l$ as representative of the lower and upper mesotrophic boundaries respectively. The values selected for total phosphorus were $10\,\mu g/l$ and $20\,\mu g/l$ for the lower and upper boundaries respectively. For seeci depth the values selected were

2 and 1m for lower and upper boundaries respectively. In order for the trophic state boundary of each parameter to be comparable, it was necessary to normalize the three parameters to a common scale, thereby establishing a trophic index (TI). The upper mesotrophic TI boundary for each parameter was chosen arbitrarily as 10 TI units. The normalization based upon the upper mesotrophic boundary values of for the above parameters were then calculated as:

$TI_{Chla} = Chla$

$TI_{Tp} = 0.5\ TP$

$TI_{SD} = 10\ SD$

The composite index could then be calculated by:

$$TI_{Naiv} = (chla + 0.5PT + 10\ SD)/3 \qquad \text{.....(3.4)}$$

In a similar manner the lower mesotrophic boundary was established as 3 TI units. The dynamic behavior of the above parameters and hence the change in the trophic state was then simulated using the dynamic water quality modelling software Duflow. To test the proposed hypothesis three scenarios were formulated;

1. To simulate the effect of light regime changes, the light extinction depth was reduced (i.e., increase in the water transparency) from its present level in increments of 25 percent for each simulation run, keeping all other parameters and boundary conditions constant.

2. To simulate water level changes in the lake, the water depth at the start of each simulation run was increased in increments of 25 percent from its present value, keeping all other factors constant.

3. To simulate the manner in which nutrient enrichment would affect the trophic status of the lake. The concentration of nutrients (i.e., phosphorus nitrogen) at the model boundary was in creased in increments of 20 percent from the present value, keeping all other factors constant.

CHAPTER FOUR

4. Results of Spatial Analysis

4.1 Geostatistical analysis

Descriptive statistics for the water quality parameters that were measured at the surface of the lake are shown in table 4.1. The statistics for pH are based on the pH measurement without transforming them to H^+. The descriptive statistics for temperature, dissolved oxygen, EC, pH and secci depth indicates that these data sets are nearly symmetrical about their mean values (as is illustrated by the closeness of their mean and median values). The data sets for turbidity and silica on the other hand indicate that the concentration of suspended matter and dissolved silica at this level in the lake is not normally distributed. This conclusion is based on the relatively large differences between the mean and median values for these data sets, and is further supported by the large value of their standard deviations in comparison to the mean values. The coefficient of variation, used as a measure of relative dispersion also indicates that the variations in the turbidity and silica data sets are relatively high.

Table 4.1: Number of observations (N_{obs}) average values (m), standard deviation (s), median, minimum and maximum values for water quality parameters measured at the surface of lake Naivasha.

Variables	N_{obs}	M	Median	S	CV	Minimum	Maximum
pH	60	8.30	8.37	0.11	0.01	8.16	8.58
EC (µS/cm)	60	471.23	470.50	9.09	0.02	449.00	501.00
Turbidity (mg/l)	60	34.28	28.00	15.04	0.44	6.00	95.00
Dissolved oxygen (mg/l)	60	9.58	9. 80	2.97	0.31	3.90	14.26
Temperature (°C)	60	20.79	20.75	0.37	0.01	18.00	24.20
Salinity %	60	0.01	0.01	0.00	0.00	0.01	0.01
Secci depth (m)	60	0.60	0.60	0.17	0.28	0.20	1.40
Silica (mg/l)	54	15.24	14.10	7.24	0.47	4.70	39.20

CV = coefficient of variation

The above analysis is supported by the results of the Chi-square goodness of fit tests for normality, which indicated that with the exception of the silica, and turbidity data sets, the null hypothesis of normality could not be rejected at the 95 percent confidence level. Therefore scale transformations were unnecessary, and as a result these data sets were used for the determination of the spatial correlation. In the case of silica and turbidity data sets transformation and had to be applied, a log transformation for silica, and a power of two transform for turbidity.

4.1.1 Omnidirectional variogram

Table 4.2: Nugget (B), range (A), and sill (C_0), parameter values and variogram models for Water quality parameters measured at the surface of the lake.

Variables	Models	C_0	A	B
pH	Spherical	0.012	4000	0
EC (μS/cm)	Pure nugget	70.0	-	-
Transformed Turbidity	Spherical	1.8×10^{6}	3000	3×10^{5}
Dissolved oxygen (mg/l)	Spherical	7.4	3000	0
Temperature (°C)	Spherical	1.4	3000	0
Secci Depth (m)	Spherical	0.02	3000	0
Transformed Silica	Pure nugget	0.48	-	-

Table 4.2 shows the estimated semivariogram models and parameters for the data sets described in table 4.1. For pH, a spherical model showed the best fit with a nugget of 0 and sill of 0.012 increasing up to a range of 4000 m. Spherical models were also fitted to the sample variograms for temperature, dissolved oxygen, secci depth, and power transformed turbidity. The semivariogram for transformed turbidity showed a nugget effect of 3×10^{5} $(mg/l)^2$ increasing to a sill of 1.8×10^{6} $(mg/l)^2$, and reached a range of 3000 m. The semivariogram for temperature also showed a spherical model to a lag of 3000 m, with a sill of 1.4 $(^{o}\ C)^2$ and 0 nugget. The spatial variations in dissolved oxygen were also modeled with a spherical variogram with 0 nugget effect

and a sill value of 7.4 $(mg/l)^2$ to a range of 3000 m. All four semivariograms showed clear spatial dependence, although the models parameters were different. In contrast the semivariogram for EC and log transformed silica showed pure nugget effect to values of 70 $(\mu S/cm)^2$ and 0.48 respectively that the distribution of these variables were spatially independent. This analysis was not performed on the salinity data set, since it showed no variation i.e., the variance from table 4.1 is zero.

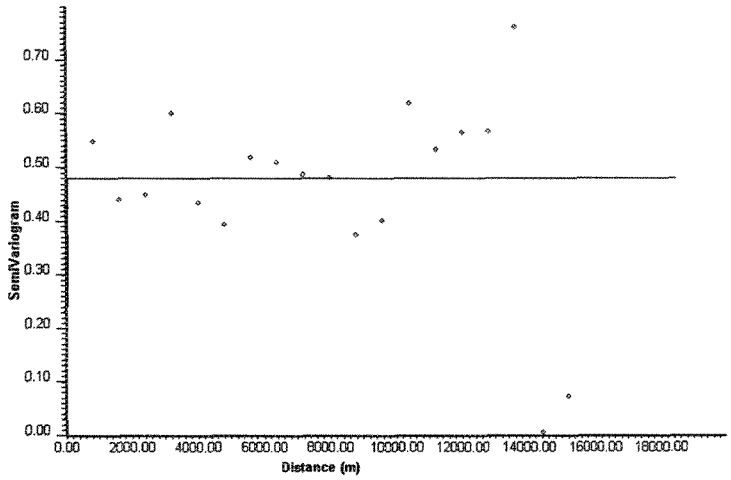


Figure 4.1: Omnidirectional variogram for dissolved silica, showing the lack of spatial structure i.e., a pure nugget effect.

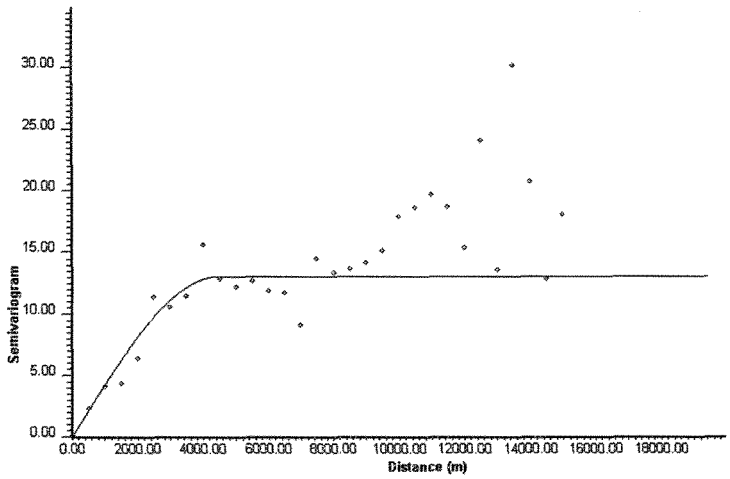


Figure 4.2: Omnidirectional variogram for pH, showing a good spatial structure and the hole effect

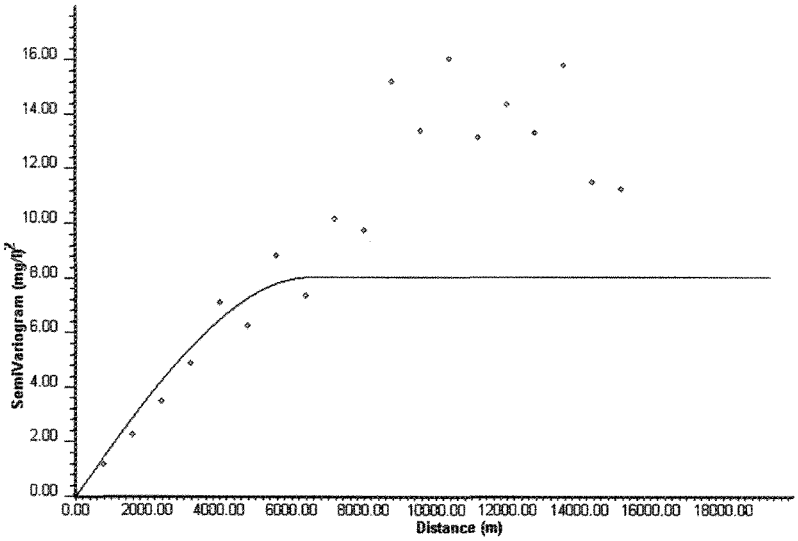


Figure 4.3: Omnidirectional variogram for dissolved oxygen, showing good spatial structure and the hole effect.

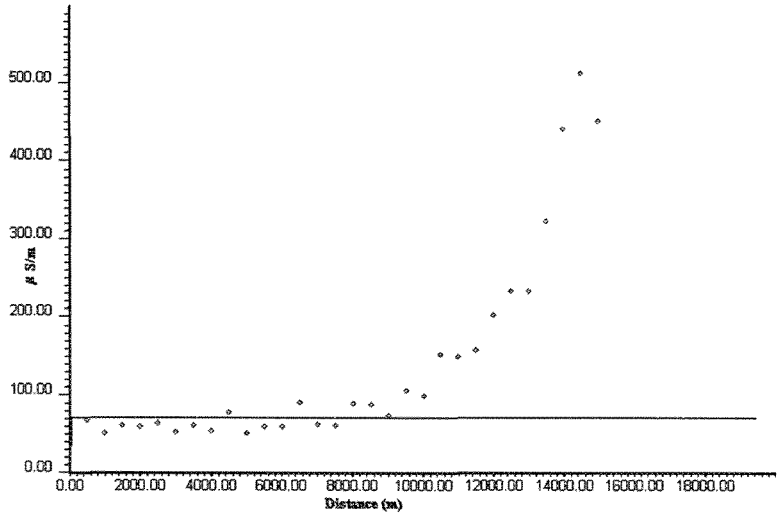


Figure 4.4: Omnidirectional variogram for EC, showing the lack of spatial structure i.e., pure nugget effect

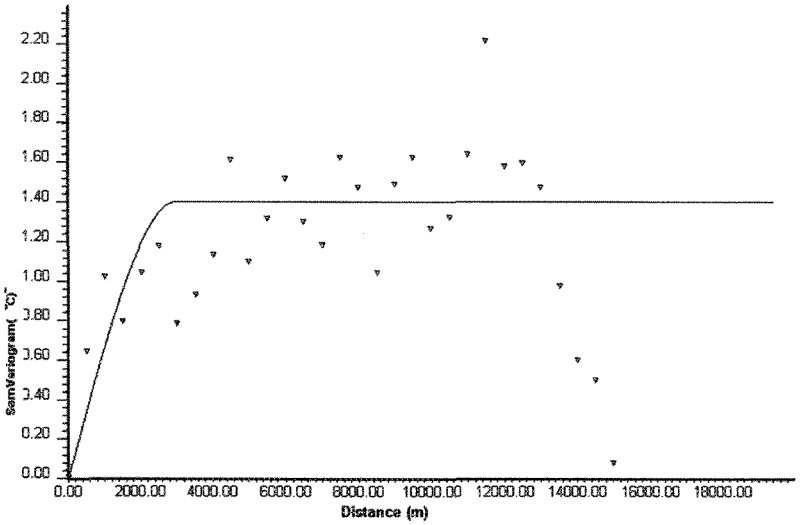


Figure 4.5: Omnidirectional variogram for temperature, showing the Of spatial structure and the hole effect.

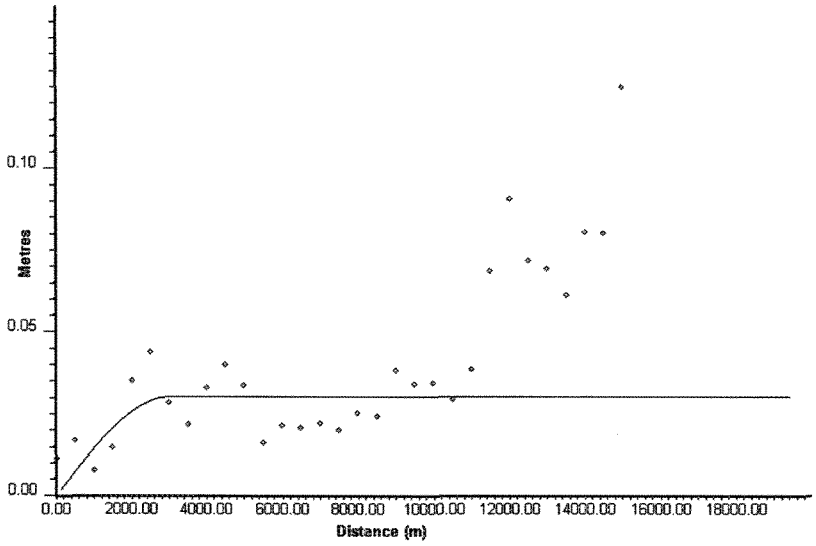


Figure 4.6: Omnidirectional variogram for secci depth showing the hole effect.

The variograms for silica and pH showed a hundred percent nugget variance. The implication of this is that the variation of these parameters within the Lake cannot be distinguished from that of non-correlated random variables. This might be attributed to the fact that these parameters are dissolved species and as such would be advectively dispersed throughout the water body, which would cause low spatial variability, and hence a lack of spatial correlation. In the case of silica however, this result was unexpected owing to the large diatom population of the lake. It was expected that silica concentrations would be lower where diatom productivity was high, thus introducing a certain amount of variability.

The variograms for the remaining water quality parameters gave clear indications of spatial structure. In most cases however, the variograms did not reach a constant sill value but displayed what is commonly called a hole effect. In this situation the variogram values oscillates about a constant value indicating an underlying periodicity in the spatial process as would be obtained if the dominant hydrogeochemical process controlling the distribution of these parameters were wave like. Statistical variation within the data sets could also generate these types of patterns within the variograms. Most of the variograms showed a nested structure, which implies that, spatial correlation probably, occur at two different scales within the lake. A local scale occurring at a range of about 3 km, and a more global process which involves the entire lake and processes occurring at the edge of the lake. Most of the variograms also showed the presents of a trend.

The variogram for transformed turbidity showed a large nugget effect, which indicates that spatially dependent variations occur over distances smaller than the lag distance of 800 m. Since the turbidity was measured to a precision of 3 mg/l which would have resulted in a nugget variance of 9 $(mg/l)^2$ if due only to measurement variability. The remaining variability can be attributed to short distance spatial variability. Also some amount of temporal variability might be present since the single sampling episode lasted for just under three days.

4.1.2 Directional variogram

Figure 4.4 gives an example of the directional variogram calculated for the parameters described above. Most of the data sets were anisotropic, indicating that the range of influence or the spatial correlation is influenced by direction. The semivariograms exhibits a well-defined spatial structure, which in the east-west direction is comparable to the omnidirectional variogram. Directional variogram in the north-south direction (perpendicular to the main direction of water movement within the lake) also showed well-defined spatial structure, the range in this direction however was some what less than that in east-west direction, but was comparable to that in the Omnidirectional variograms.

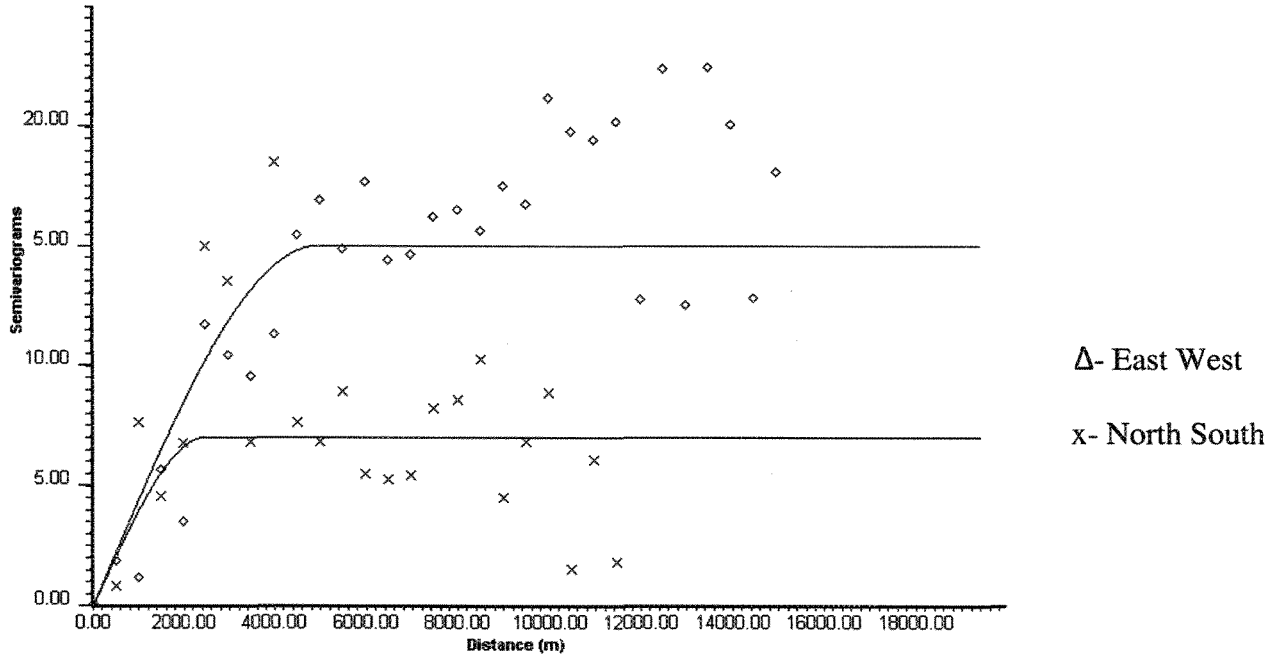


Figure 4.7: Directional variogram for pH, showing good spatial structure in both the east-west and north south direction. The range in the east-west direction is longer however.

4.1.3 Drift estimates

Examples of the drift estimates calculated for the parameters that showed a clear indication of spatial structure is presented in figure 4.5. For all lag distances less approximately 4500 m drift fluctuates about zero. Implying that these data set are second order stationary, at least for lag distances less than 4500 m which is well beyond any apparent correlation scale. The average drift estimated over all lag distances for temperature was 0.29 °C, in the case of pH, turbidity and dissolved oxygen it was –0.007, 2.41 mg/l, and 0.93 mg/l respectively.

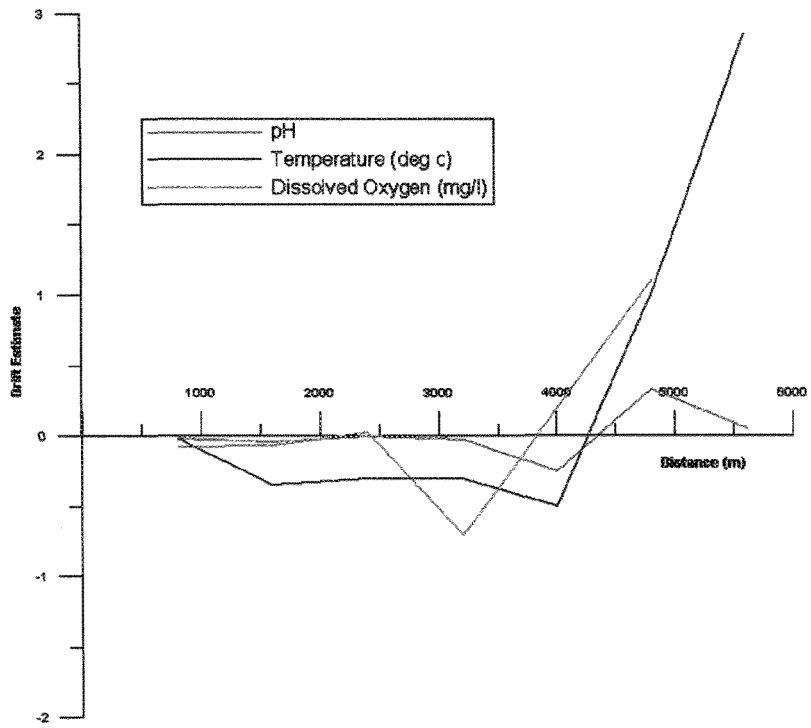


Figure 4.8: Examples of drift values calculated for data set

4.1.4 Kriged maps

Owing to the fact that the range of spatial correlation in the north-south direction for the above parameters was more or less equal to that in the Omnidirectional variogram, and they were second order stationary to lags of at least 4.5 km. Ordinary kriging using a 3 km search neighborhood, was applied for estimation of the these parameters over the entire lake, using grids of 100 by 100 m. Error maps of prediction were also generated (appendix 2). In the case of silica and EC a moving average method was used, since these parameters showed no sign of spatial correlation.

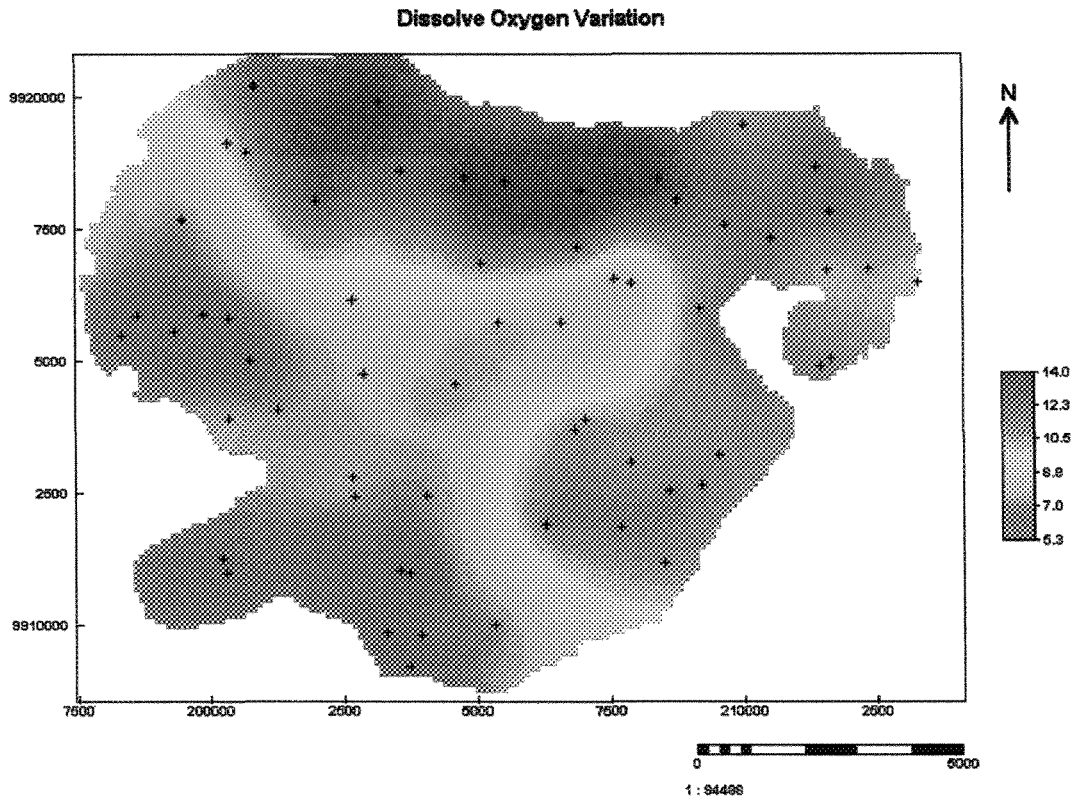


Figure 4.9: Distribution of dissolved oxygen within the lake, the actual sample points are denoted by crosses.

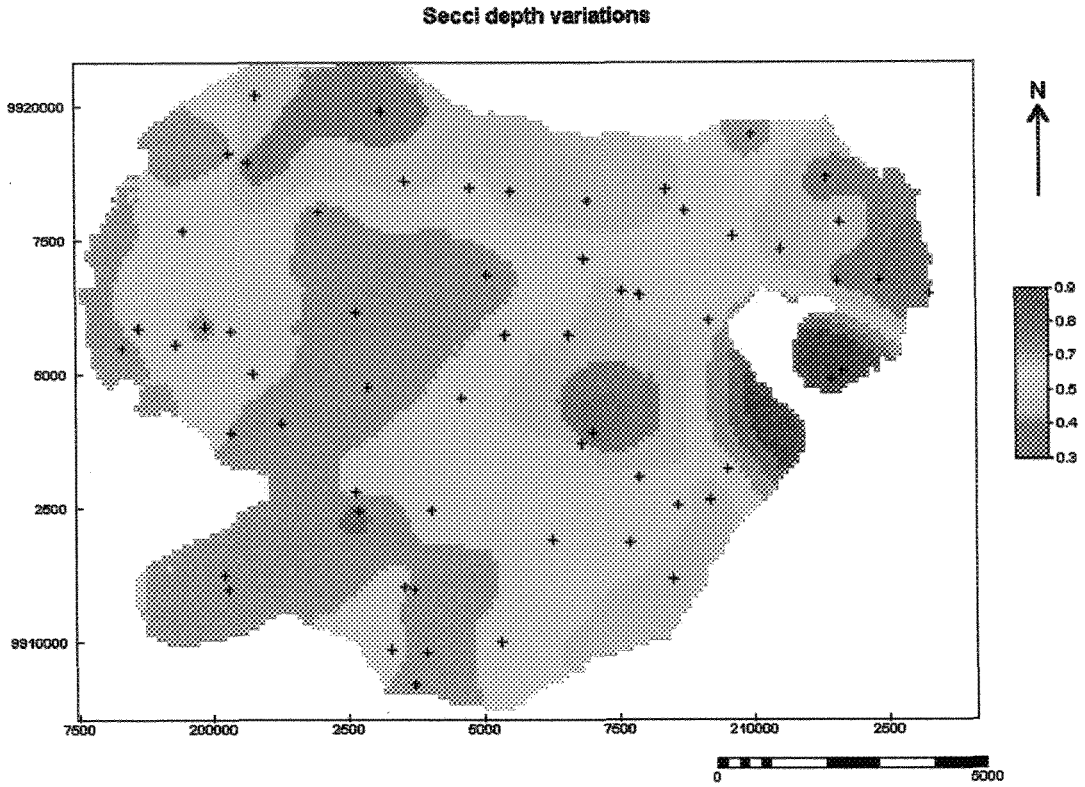


Figure 4.10: Distribution of secci depth within the lake

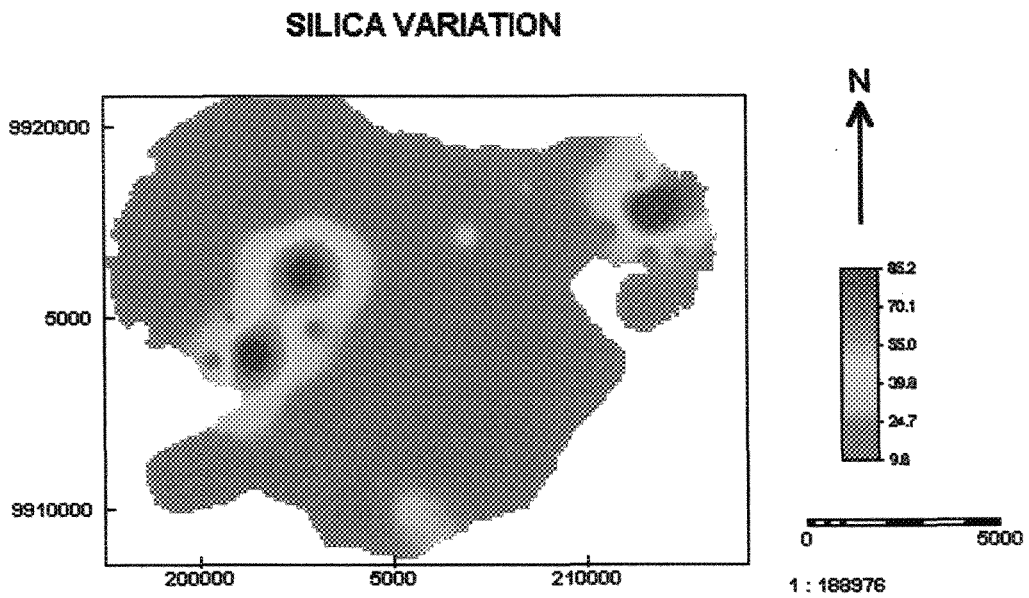


Figure 4.11: Map showing the distribution of silica within the lake

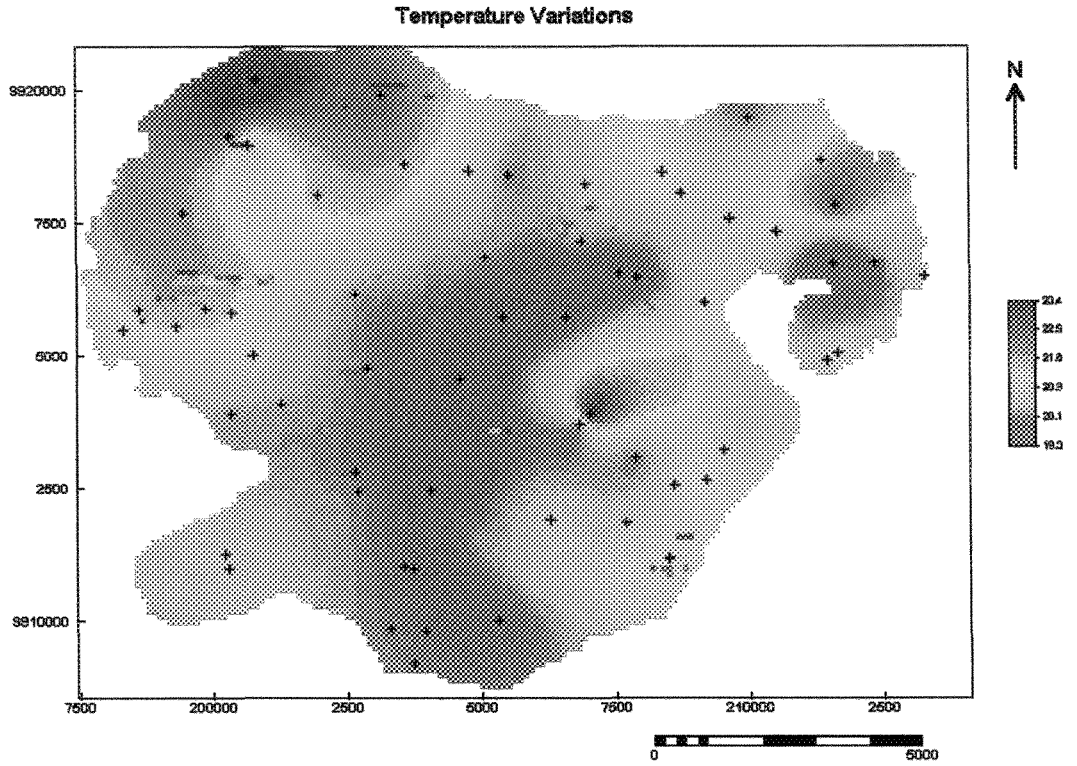


Figure 4.12: Map showing the distribution of temperature within the lake

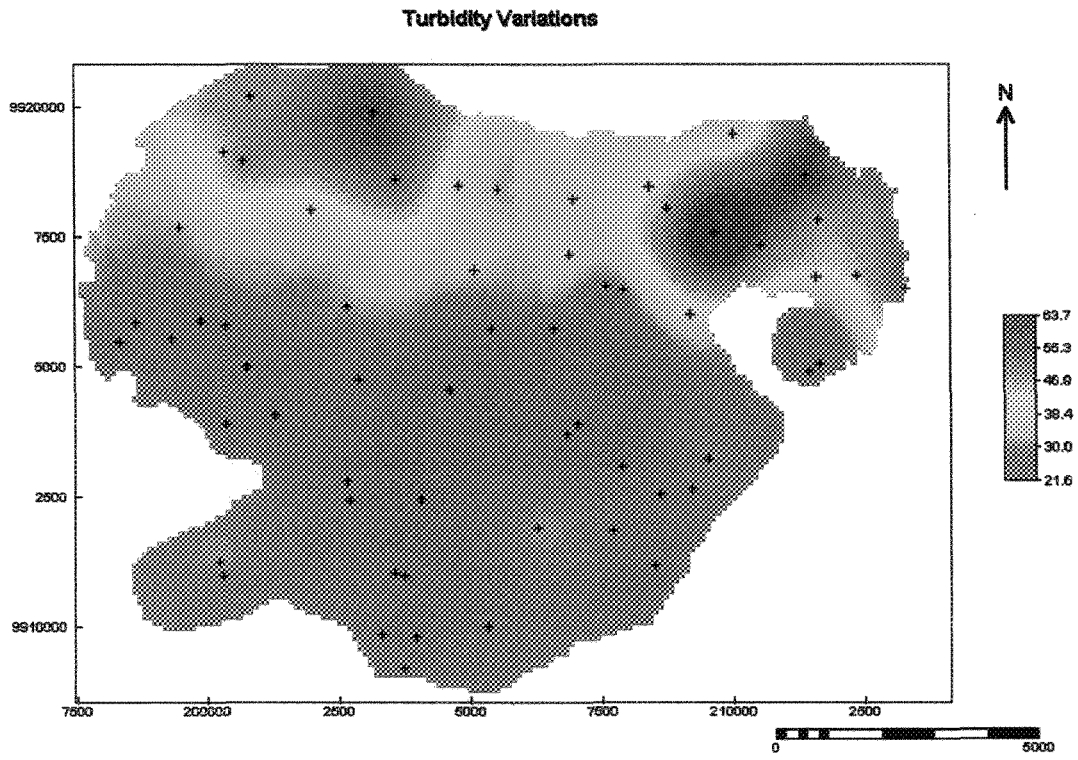


Figure 4.13: Map showing the distribution of suspended matter within the lake

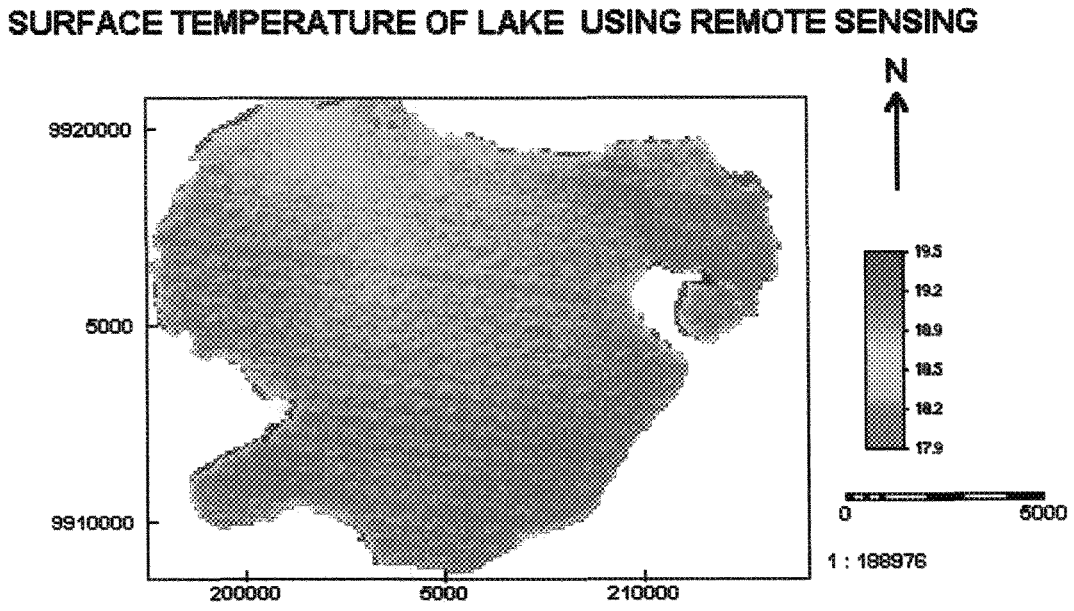


Figure 4.14: Surface temperature of the Lake derived using TM Image

The maps derived from the kriging process showed that a number of the water quality parameters exhibits what can be considered as hotspots. In the case of temperature, two such areas are located in the northeast and southwest sections of the lake respectively. It could not be determined whether they were the result of geothermal activities or local depth effect. The center of these areas however, coincided with just one sample point, and so could also have resulted from sampling variations. The fact similar areas are not present in the satellite derived temperature map is not surprising, and can be attributed to the different scales of measurements that were used to derive these maps. Also the fact that the satellite derived maps lacked the temporal variations present in the kriged maps, owing to a sampling episode of just under three days. In addition atmospheric corrections were not applied to the satellite image. Despite the fact that these corrections are small for the thermal bands of the TM image, they could still be masking temperature variations.

4.2 Verification of variogram s

Table 4.3 : Shows the descriptive statistics generated from the kriged maps

Variables	M	S	Minimum	Maximum
pH	8.38	0.094	8.6	8.2
Turbidity (mg/l)	34.01	11.69	91.6	7.7
Dissolved oxygen (mg/l)	9.15	3.074	14.4	3.9
Temperature (°C)	20.88	0.869	24.2	18.1
Secci depth (m)	0.615	0.14	1.38	0.21

Some measure of confidence in the variograms can be found by considering the following, the kriging estimates should give comparable summary statistics to the sample data. With this in mind, the descriptive statistics generated from the kriged maps are shown in table 4.3. From this, it can be seen that while the mean and the range of the estimates are more or less equal to those of the sample data, the standard deviations are higher in most cases. This result was unexpected, since kriging like most estimation processes has a smoothing effect on the variance, so it was expected that the variance would be lower for the kriged estimates. The higher values can probably be attributed to outliers in the sample data sets and also errors in the experimental variograms, both of which would cause the generation of outliers in the estimated maps. It should be noted however, that the standard deviations obtained are of the same order of magnitude as those in the raw data sets. The above analysis was confirmed by the results of the cross validation procedure, which indicated that the kriged estimates tended to be somewhat lower than the values in the raw data sets.

4.3 Total Horizontal variations

Table 4.4: Number of observations (N_{obs}) average values (m), standard deviation (s), median, and range values for temperature measured different depths in lake Naivasha.

Depth M	N_{obs}	Mean °C	Median °C	S °C	CV	Range °C
0	60	20.79	20.75	0.37	0.05	5.8
1	60	20.67	20.67	0.94	0.04	4.5
2	52	19.90	19.90	0.78	0.04	2.8
3	51	19.60	19.60	0.60	0.03	1.8

CV = Coefficient of variation

Table 4.3 shows the descriptive statistics obtained from the analysis of temperature measured at different depth in the lake. An inspection of the coefficients of variation and means indicates that the distribution of temperature is approximately the same at each horizontal level i.e., the data sets are equivalent. This was confirmed by the results of the Kruskal-wallis test, which indicated that at the 95 percent confidence level the null hypothesis that these distributions are the same could not be rejected ($p = 0.342$). Since none of the groups tended to produce values larger than the other, an average standard deviation was calculated. Assuming that the distribution of the data approaches a normal distribution, then plus or minus two standard deviations would result in 95 percent of the total temperature variation over a horizontal layer. Since the average standard deviation in this case was 0.7 ° C, then the expected variation in temperature would be ± 1.4 ° C. Similar analysis carried out on the other water parameters measured in the lake resulted in the expected variations shown in Table 4.5.

Table 4.5: Shows expected variations of the water quality parameters in the horizontal direction

Variables	Expected Variation
pH	0.58
EC (μS/cm)	0.02
Turbidity (mg/l)	16.8
Dissolved oxygen (mg/l)	7.4
Temperature (°C)	1.4

Table 4.5 shows the expected variation in water quality parameters in horizontal direction. If these values are compared to the instrument sensitivity, it can be seen that the variation in dissolved oxygen, temperature and turbidity cannot be neglected owing to the fact that sensitivity of the U-10 for these parameters are ± 0.3 mg/l, ± 3 mg/l, and ± 0.03 °C respectively. As a consequence, these properties can be expected to show significant variation in the horizontal direction. Since the sensitivity of the U-10 pH probe was ± 0.1 pH units, the horizontal variations in pH can be neglected. The sensitivity of the U-10 conductivity probe was ± 1umhos/cm, as such, there were negligible horizontal variations in conductivity as well.

4.4 Total Vertical Variation

Table 4.6: Number of stations (N_{obs}) average values (m), standard deviation (s), median, and range values for water quality parameters normalized to determine the magnitude of the vertical variation

Variables	N_{obs}	M	Median	S	CV	Range
pH	60	0.24	0.25	0.46	1.46	0.12
EC (µS/cm)	60	-0.001	0.001	0.001	-1.00	0.003
Turbidity (mg/l)	60	-1.5	-1.25	1.3	-1.02	2.75
Dissolved oxygen (mg/l)	60	0.45	0. 50	0.65	1.42	1.40
Temperature (°C)	60	0.69	0.75	0.50	0.56	1.64

CV = coefficient of variation

Table 4.4 shows the descriptive statistics for the normalized data, from the range of the values it can be seen that the variation of the water quality parameters in the vertical direction is not very large. The fact coefficients of variation exceeds one in most cases can be attributed to skewness which resulted from small number of vertical samples used to estimate the average values. As a result of the sample size the standard deviations would be inflated, leading to the high CV values. Owing to this the expected variations in the vertical direction could not be calculated, as was done for the horizontal direction. Figure 4.15 shows the regression of normalized temperature on depth, the value of r^2 indicate that only 16 percent of the variation could be attributed to depth. Despite this, the results of the ANOVA indicated that at the 95 percent confidence level the null hypothesis that the slope was equal to zero had to be rejected (p= 0.00). Similar analysis of the dissolved oxygen and pH gave r^2 values of 10 and 20 percent respectively, in both cases the null hypothesis could not be accepted well (p= 0.00). In the case of EC and turbidity the r^2 values were 0.7 and 0.1 percent respectively and the null hypothesis could not rejected (p values of 0.170 for EC and 0.244 for turbidity). From the size of the regression coefficients it would seem that the lake is vertically well mixed since only a small part to the total variation in these properties could be attributed a change in depth.

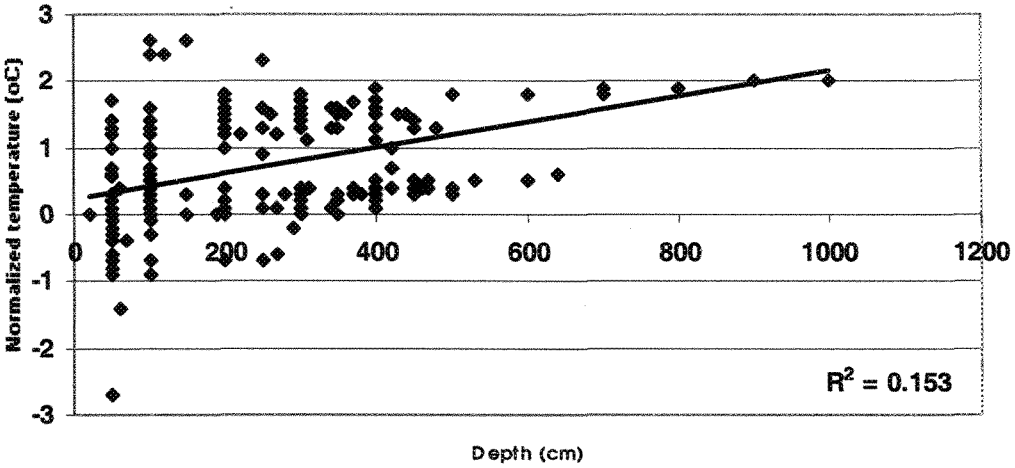


Figure 4.15: Regression of normalized temperature with depth

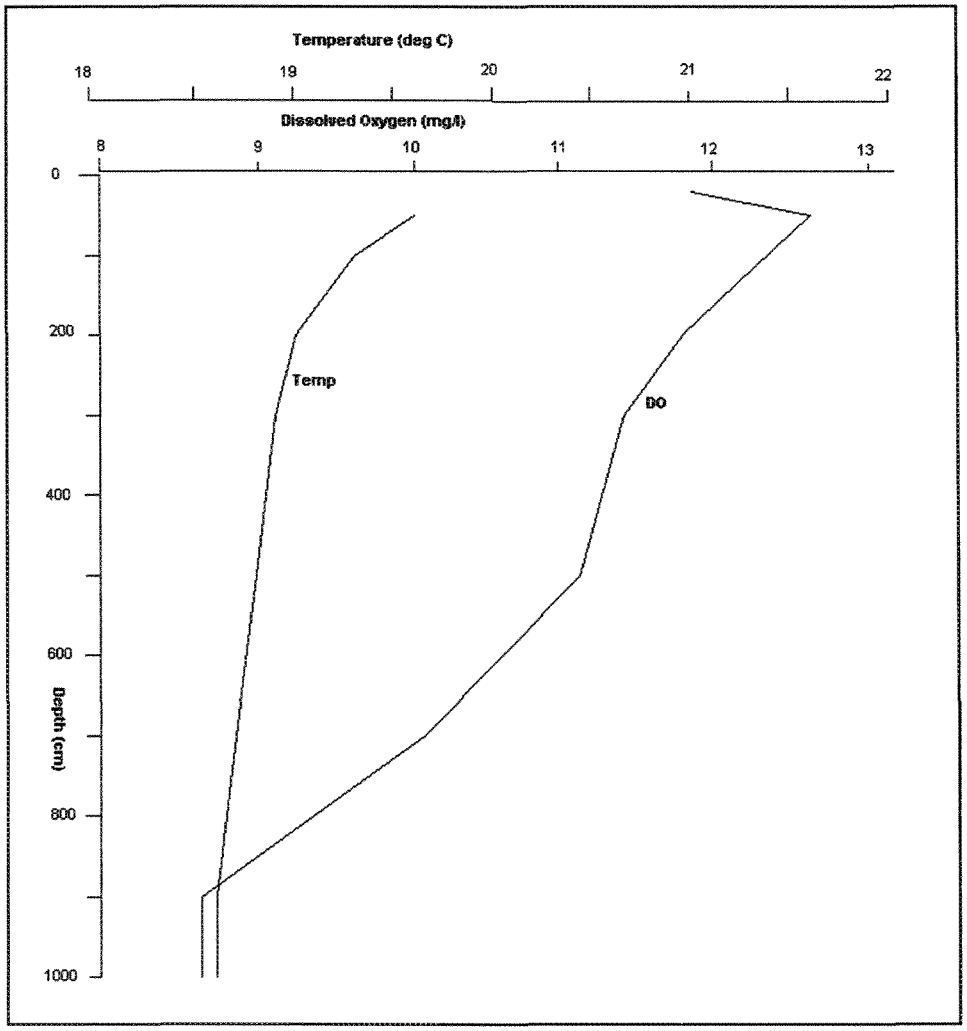


Figure 4.16: Shows Oxygen and Temperature profiles measured at the deepest section in the Lake, the almost vertical temperature profile supports the above assumption that the lake is probably vertically well mixed.

4.5 Sampling density

On the basis of the structural analyses, the omnidirectional variograms were assumed to effectively characterize the spatial correlation of the above water quality parameters. Given (1) the fact that the data sets were second order stationary, at least for lag distances less than 4.5 km which is beyond any apparent correlation scale; (2) despite the presents of anisotropy in most of the data sets, the range in the least correlated direction was about the same as for the omnidirectional variogram. Since the range of spatial correlation from the variogram were above 3 km in all cases where the parameters showed spatial structure, the optimal distance between sampling stations was estimated to be about 2.5 km. With an average surface area of approximately 130 km^2, it would require some 19 sampling points to characterize the water quality of the lake, assuming that these points would be at the center of regular square grids.

4.6 Sample pattern

Table 4.7: Shows the Number of Samples that should be collected at the above sites for a desired precision and probability level.

Total Phosphorus	Mean[1]	194 µgP/l					
	Precision[2]	±5 µgP/l			±10 µgP/l		
	Probability	95%	90%	80%	95%	90%	80%
	Sample No.	7	5	3	2	1	1
Turbidity	Mean	36 mg/l					
	Precision	±5 mg/l			±10 mg/l		
	Probability	95%	90%	80%	95%	90%	80%
	Sample No.	40	27	17	12	8	5
Chlorophyll-a	Mean	27 µg/l					
	Precision	±5 µg/l			±10 µg/l		
	Probability	95%	90%	80%	95%	90%	80%
	Sample No.	16	11	7	6	4	2
Total Nitrogen	Mean	1.76 mg/l					
	Precision	± 0.5 mg/l			± 1 mg/l		
	Probability	95%	90%	80%	95%	90%	80%
	Sample No.	8	6	4	2	1	1

Table 4.7 above gives an initial estimate of the number of samples that should be collected at each of the above sampling stations to be within a desired precision of the station mean at a particular probability level. The calculations underlying this matrix are based on the random sampling formula (eqn. 2.14), and include only four eutrophication-related parameters but can be upgraded to include others (See Appendix 2 for calculations). It is evident from table 4.7 that there is inconsistency in sample patterns. These range from the collection of 40 turbidity samples to obtain the desired precision, at the 95 percent level, to 8 for total phosphorus. It is also evident that small increases in precision require large increases in sampling effort as is exemplified by turbidity, where the number of samples changed from 40 to 27 for a five percent increase precision. These effect no doubt results from the assumptions underlying the random sampling formula, the major ones being that the observations are independent and normally distributed. If the parameters are not normally distributed then the variance will be inflated and as such the number of samples needed to estimate the mean will be large, the same applies if the observation are correlated in space and time. In this case, since both assumptions are violated, the best that could be hoped for is an initial estimate of the sample size needed to estimate the station mean. The sampling pattern could then be reviewed at least on an annual basis as data becomes available, the formula could then be used to reflect the increase or decrease variability found over the monitoring period.

Notwithstanding the fact that the spatial variation of important the parameters such as chlorophyll-a was not measured directly, it is expected that the phytoplankton would distribute themselves in response to physical factors such as temperature and turbidity. With the result that these parameters would show a similar spatial pattern to the parameters that was actually measured, thereby allowing the same sampling locations to be used for these parameters as well

CHAPTER FIVE

5. Eutrophication Modelling

A quantitative study of the eutrophication process has to take into account not only biological and chemical processes but also physical processes in the water system as well. Box models are useful and mathematically simple tools for this kind of investigations. The choice of the subsystems (box) should correspond to regions of homogenous chemical and biological conditions. The model then describes the temporal variations of the relevant concentration within each box depending on a set of model parameters which themselves can be different for each box (Gatcher, 1984). Despite the mathematical simplicity of this approach one of the strengths of this type of model is its ability to provide insight into the internal processes controlling eutrophication (e.g. why or how the process occurs). It also allows multiple changes to be considered simultaneously. Manipulation of the values of the state variables in the model can provide potentially valuable insight into what factors exert primary control in the eutrophication process (Imoden, 1981).

Against this backdrop, this chapter gives a description of the one box model used to simulate chlorophyll-a, phosphorous and secci depth in the lake. As there is no generally accepted unique definition of a lake trophic state, nor is there a quantitative measure the above water quality indicator are employed to provide a composite description of the lake trophic state.

5.1 Duflow

The dynamics of chlorophyll-a, dissolved nutrients, and secci depth in lake Naivasha were studied using DUFLOW 2 (EDS, 1998), a microcomputer package for the simulation of non-steady flow and water quality processes in one-dimensional channel systems. These channel systems are discretisized by defining a network consisting of nodes connected by sections.

DUFLOW allows user definition of water quality processes by means of a special modelling language, which permits users to create models that range from simple

mass balance models of conservative substances up to complex ecosystem models. In addition it has two predefined eutrophication models that can be modified based on the particular water body. The water quality model is solved either simultaneously with or subsequent to, the water flow. For a comprehensive description of DUFLOW see EDS (1998). The water quality model used in this study simulates the concentrations of chlorophyll-a, phosphorus, ammonium, nitrate, and secci depth dynamically. The differential equations used in the model, and the processes concerning the state variables are also explained in EDS (1998). While the parameter meanings and symbols are given in appendix 5.

5.2 General Method of Calibration

5.2.1 Discretization and boundary conditions

The dynamic water quality model described above was calibrated for Lake Naivasha for the period 1/1 -30/12/1997. For this, the lake area (Figure 3.1) was discretized into a single section of average length of about 15,000m and width 10,000m i.e., a box. The conditions for water composition at the upper boundaries and in the lake were obtained from field data used in a previous study (Hubble, 1999). Data on stream inflows and climatology (rainfall, evaporation and sunshine duration) were obtained from the extensive databases contained at the Institute of Aerospace Survey and Earth Sciences. Surface runoff at model boundaries was neglected owing to the semi-arid nature of the area, and the fact that the soils surrounding the lake is composed chiefly of volcanic and diatom earth, which gives them a high infiltration rate. Data on rainfall composition were lacking; the concentrations at the model boundary were therefore obtained from an earlier study by Mavuti (1996).

5.2.2 Calibration

The method used to calibrate the model is based on that developed by Van der Perk, (1997). The range of the parameters to be calibrated is converted to discrete values by division into a number of equal steps, as an example consider the parameter algal maximum growth rate (U_{max}) which ranges from 1 – 5 1/day, it would be divided into four equal steps of values 1, 2, 3,and 4 1/day. This process was repeated for all parameters, and all possible combination drawn from this discrete distribution. A disadvantage of this method however, is the considerable computational time

involved in the process for this reason, the calibration of the complete set of model parameter was done in three steps. First, the parameters for which only one value could be found in the literature were assume fixed. Secondly, a sensitivity analysis was performed on the remaining parameters to determine those parameters to which the model was least sensitive, these parameters were also fixed, thereby reducing the number of parameters to be calibrated. All model calculations were evaluated by means of an objective function; in this case the sum of squared differences between the observed and simulated concentrations. The parameter combination yielding the minimum value of this objective function was taken as the best parameter set. Thus, in this case the best parameter set is a least-squares estimate. To evaluate model performance the best parameter set, the correlation coefficient was used to make comparisons with the observed concentrations.

5.2.3 Fixed parameters for the water quality model

Several parameter values of the water quality model were fixed beforehand, because they were directly derived from field measurements or only one value could be obtained from the literature. Consequently, the following parameters were fixed; Tga, tra, POR HB, RHO, EAC, tnit, tdenB, tdaB, T, Kmn, E0, Ia, and L.

5.2.4 Prior sensitivity analysis of the water quality model

To reduce the number of model parameters that must be calibrated, the parameters to which the model is less sensitive were fixed in the subsequent calibration stages. To select these parameters, a priori sensitivity analysis of the water quality model was carried out using the default values given in the Duflow manual as a base set of parameters. Sensitivity analysis was then carried on this base set, by changing one parameter at a time, while keeping the rest constant. The parameter values were changed over a range (i.e., increased and decreased in increments of 25 percent of the base value) and relative sensitivity coefficients calculated, and ranked into high, medium and low. The parameters listed below were ranked as low, and were therefore fixed in the subsequent calibration stage; Edif, eSS, achlc, apc, ealg and Is, TminB.

The chlorophyll-a model was calibrated using 8 chlorophyll measurement obtained from a pervious study (Hubble, 1998). Unfortunately however, this model could not

be tested owing to the fact that the samples collected during the present study all gave values below the detection limit. This can be attributed to the use of membrane as opposed to glass fiber filters, in the filtering of the water samples taken from the lake during the present study.

The nitrogen, phosphorus secci depth models were also calibrated using data from the same source as that for the calibration of the chlorophyll-a model. 3 measured values were used for the calibration of both the nitrogen and phosphorus model. On the other hand 6 measurements were used to calibrate the secci depth model. The water flow model was calibrated using monthly average water depth measured from 1932- 1997.

5.2.5 Calibration Results

Figures 5.1- 5.3 show the observed and simulated concentrations as a function of time for the above parameters. The simulated chlorophyll-a concentration fit the measured values reasonable well. The correlation coefficient for the relationship between the measured and simulated values amounts to 0.842. In the case of the secci depth the correlation coefficient between the measured and simulated values was found to be approximately 0.62. Owing to the small number of measured data points, a correlation coefficient was not generated for either the phosphorus or nitrogen models. Figure 5.5 shows the measured and simulated water depth for the lake. From this figure it can be seen that the simulated values are fairly close to the measured values ($r^2 = 0.74$), the differences after 1962 can be attributed to an increase in abstraction of lake water for, agricultural purposes. Which was not properly accounted for by the model.

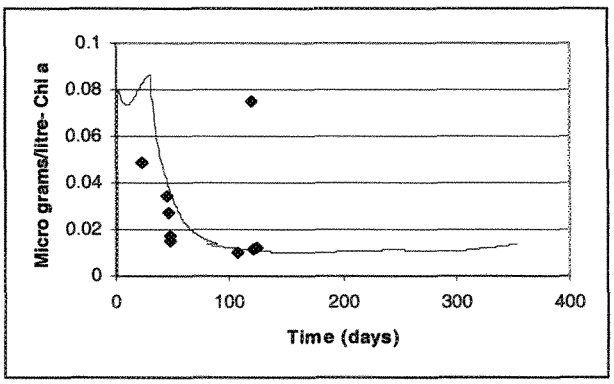


Figure 5.1: Shows simulated and measured chlorophyll-a concentrations for the Lake, the points represent the measured values and line the simulated.

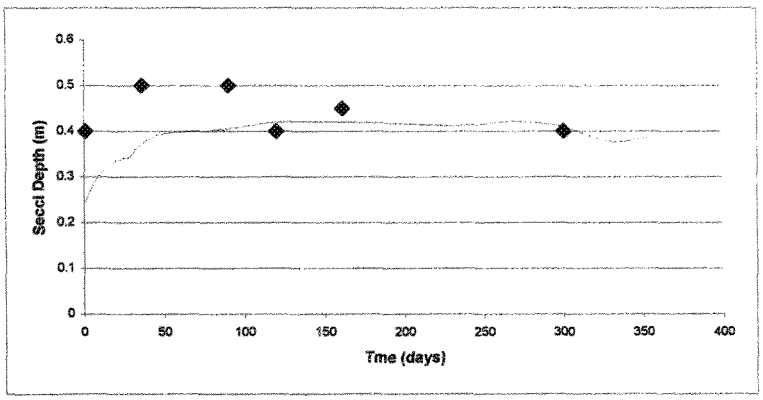


Figure 5.2: Shows simulated and measured secci depth

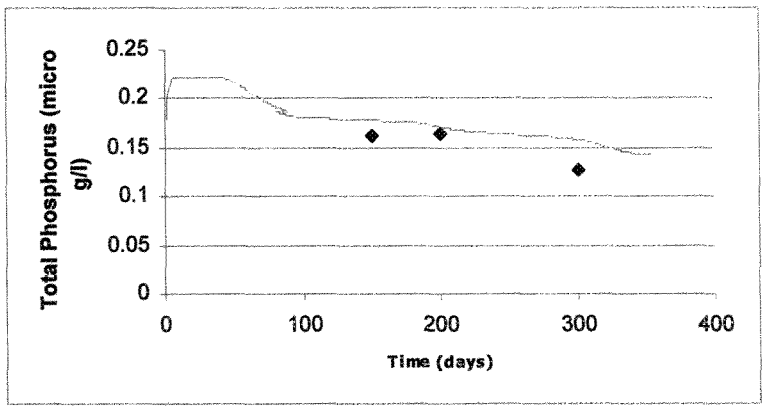
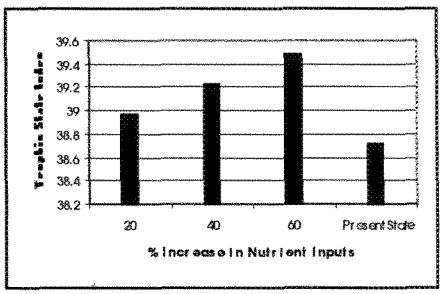


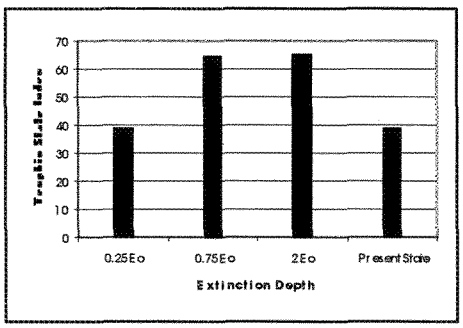
Figure 5.3: Shows simulated and measured total phosphorus for the Lake

5.3 Results on Eutrophication controls

The calibrated model was used to simulate the changes in the trophic state of the lake for the scenarios formulated in chapter three (section 3.22). Figure 5.4 shows the changes in the composite trophic state index predicted by the model. From this it can be seen that the as the water transparency increases, the eutrophication status of the lake also increases (i.e., the water quality become worse). This result is intuitive and can be attributed to the fact that the main energy source for primary production is sunlight. Since the extinction depth determines the light attenuation properties of the lake, it regulates the amount of solar energy reaching the phytoplankton. It can therefore be expected that as water become more transparent the chlorophyll-a

lake, it regulates the amount of solar energy reaching the phytoplankton. It can therefore be expected that as water become more transparent the chlorophyll-a concentration would also increase. This increase however, would not be indefinite as is shown by the relatively small difference between the trophic state for a 75 percent and 100 percent decrease in extinction depth. This small change might be attributed to photoinhibition of the phytoplankton, which would decrease their concentration, and as a consequence, the trophic state. The pattern observed for changing lake levels can be attributed to the fact that as the water depth increases the path length for the sunlight to travel also increases, as a result, light attenuation causes a reduction in phytoplankton growth and hence a reduction in the trophic state of the lake. The simulation results of increasing nutrient levels on the trophic state of the lake is shown below, from this it can be seen as the input of nutrients increases the lake also become more eutrophic as is expected.





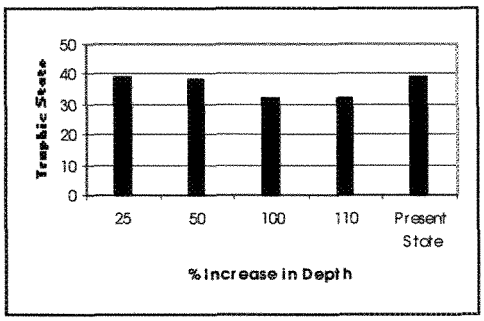


Figure 5.4: Shows simulated average yearly changes in the trophic index of the Lake for the different scenarios formulated in the study

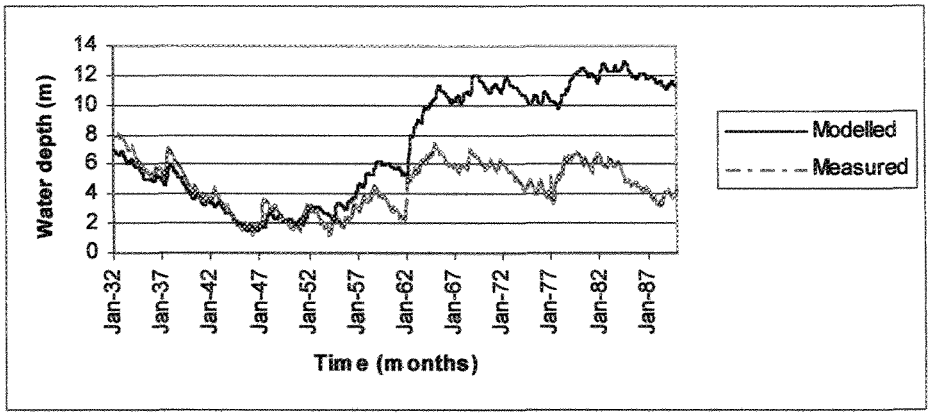


Figure 5.5:Showing measured and simulated water depths for Lake Naivasha

5.4 Nutrient limitation

Since the general ratio at which phytoplankton take up nutrient is assumed to be 7N: 1P in terms of mass, (i.e., 16N; 1P in atomic units) if the ratio of the measured concentration is less than this value then nitrogen is the potential limiting nutrient. Conversely if the ratio is greater that 7 then phosphorus is the limiting nutrient. In the case of lake Naivasha, the lake wide average N: P ratio calculated from (figure 5.6) the nine points sampled in the lake was approximately 2.1, indicating that the lake is nitrogen limited. As a consequence the level of biomass in the lake would be controlled chiefly by nitrogen.

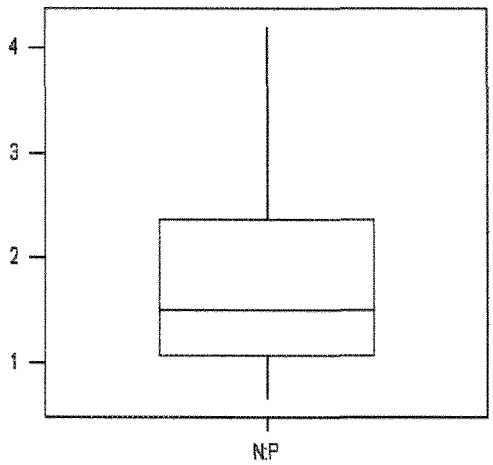


Figure 5.6: Shows box plot of N:P ratio data set for the Lake

CHAPTER SIX

6. DISCUSSION AND CONCLUSIONS

6.1 Discussion

It has been shown in this study that the method of Geostatistics can be a useful tool for the determination of spatial correlation, and hence the optimal location and number of sampling points needed to characterize water quality within the lake. It is noted however, that the use of this method is not without shortcomings. The main deficiency of this approach is that the underlying hydrogeochemical processes controlling the spatial distribution of these water quality parameters cannot be assumed to be time invariant; and as such it is intuitive to expected that the variograms might show some amount of temporal variations. With the consequence that the optimal number and locations of sampling points cannot be considered time-invariant, but is expected to be a function of the prevailing hydrological condition. The data collected in this study however, did not permit an analysis of the temporal variation in the variograms. The flow conditions the time of data collection were low inflow at the Malewa and correspondingly little or no rainfall. One way of determining if these temporal variations are significant, would be to compare variograms derived from the present study with other derived from data collected during high discharge condition of the Malewa. Such a comparison would give an indication of whether there were changes and if so just how significant these changes actually were.

Based on the criterion used in this study several important water quality parameters (example water temperature, dissolved oxygen and turbidity) were judged to show significant horizontal variations. A large part of this variation however, might not be natural variation in the properties themselves, but could be due to temporal and sampling variability. Had these been accounted for the observed variation might have been significantly less than the instrument variability. The data sets however, did not permit a partitioning of the variance; as a result that the percentage of the total variation that could be attributed to the natural variation in the properties themselves could not be obtained. A calculation of this might have resulted in most of the

property showing total horizontal variations, which could be considered insignificant. This is supported by the results the variogram analysis, which showed that for most of the properties, the spatial variation is long range i.e., over 3 km. Also the nature of the medium itself, which generally bring about advective dispersive transport and so tends to smooth out variations in a most cases.

The results of the simulation performed to test eutrophication control parameters indicated that all three parameters exert an influence on the trophic state of the lake. This result was expected based on the role played by each of these factors in the eutrophication process. Of the factors examined however, it is of note that the light regime seem to dominate the other controls, this might be attributed to the fact that solar radiation in conjunction with turbidity determines the potential level of primary production in the lake. The change in the trophic state predicted by these controls might have been higher had the simulation taken into the shallow nature of Lake. This shallowness implies that even small changes in depth causes correspondingly large changes in area and volume. If for example, the lake were display a reduction in level it can be expected that since volume decreases faster than area, there will be more energy input per unit volume. Leading to a faster heat transfer, coupled with a lower thermocline and decreased hypolimnetic volume. As a consequence, the level of nutrient re-supply from the sediment would increase with the result that overall nutrient level within the lake would tend to increase. The Duflow model however, did permit a simulation of this effect.

The trophic state change in the lake was characterized using a composite index. The rational for this is that any single parameter such as chlorophyll-a or total phosphorus might show variability both spatially and temporally which would have little significance in whole lake terms, and would therefore produce incorrect conclusion if a single parameter was used or even the average of that parameter was used. The use of this composite index was expected to have a buffering effect of sorts on these variabilities.

6.2 Conclusions

The results obtained from this study supports the conclusions listed below;

1. A statistical analysis of the data, using a comparison of the expected variation to the instrument sensitivity has resulted in the conclusion that the horizontal variation in a number of important water quality parameters in the lake is significant.

2. On the basis geostatistical analysis it was concluded that the number of sample points needed to completely characterize the water quality of the lake is approximately 20.

3. An analysis of the mass ratio of total nitrogen to total phosphorus for nine sampling locations within the lake resulted in the conclusion that nitrogen is a more limiting nutrient than phosphorus in controlling the biomass levels of the lake.

4. On the basis of the simulation results it was concluded that changes in Lake Water Levels, light regime and external nutrient input impose significant control on the trophic state of the Lake.

6.3 Recommendations

On the basis of the results obtained in this study the following recommendations are being made;

1. The temporal variations in the spatial structure of the water quality parameter of the lake should be investigated, by comparing the variograms from two different time periods.

2. Further investigations should be undertaken to determine the contributions of different types of variability (i.e., sampling, laboratory, and natural) to the overall variability of the water quality parameters in the lake, and to assess the implications of these variabilities to lake water quality modelling.

3. Owing to the shallow nature of the lake, small changes in depth result in correspondingly large changes in area and volume. The implication of this is that more energy per unit volume reaches the lake. This effect could not be modeled using Duflow. It is therefore been recommended that the modelling exercise be repeated using a model that takes such effects into account.

REFERENCES

Abiya, I. O. (1996). "Towards sustainable utilization of Lake Naivasha, Kenya." Lake & Reservoirs: Research Management **2**: 231-242.

David Harper, M., G. Phillips, et al. (1993). "Eutrophication prognosis for Lake Naivasha, Kenya." Verh. Internat. Verein. Limnol **23**.

Gaudet, J. J. and F. M. Muthuri (1981). "Nutrient regeneration in shallow tropical lake water." Verh. Internat. Verein. Limnol **21**: 725-729.

George, E. P. B., G. William Hunter, et al. (1978). Satistics for Experimenters: An Introduction to Design, Data Analysi and Model Building, John Wiley & Sons Inc.

Gieske, A. (1999). Geostatistics for Hydrologists. Advanced GIS.

Gregor, D. J. and W. Rast (1992). "Simple Trophic State Classification of the Canadian nearshore Waters of the great Lakes." Water Resources Bulletin **18**(4): 565-573.

Groenigen, J.-W. v. (1999). Constrained Optimisation of Spatial Sampling. Enschede, International Institute for Aerospace survey and Earth Sciences.

Helsel, D. R. and R. M. Hirsch (1992). Statistical Methods in Water Resources. Amsterdam, Elsevier.

Hosper, H. (1997). Clearing Lakes; an ecosystem approach to the restoration and management of shallow lakes in the Netherlands. Enschede, University of Twente.

Hubble, D. S. (2000). Control on Primary Production in Shallow tropical freshwaters. Biology, University of leicester.

Imboden, D. M. (1974). "Phosphorus Model of Lake Eutrophication." limnology and Oceanography **19**(2): 297-304.

Imboden, D. M. and R. Gachter (1978). "A Dynamic Lake Model for Trophic State Prediction." Ecological Modelling **4**: 77-98.

John Guadet, J. (1979). "Seasonal Changes in Nutrients in a Tropical Swamp: North Swamp, Lake Naivasha, Kenya." Journal of Ecology **67**: 953-981.

John Melack, M., Peter Kilham, et al. (1982). "Responses of Phytoplankton to Experimental Fertilization with Ammonium and Phosphate in an African Lake." Oecolagia **52**: 321-326.

Jongman, R. H. G., C. J. F. ter Braak, et al. (1987). Data analysis in community and landscape ecology. Den Haag, Pudoc Wageningen.

Keesman, K. and G. v. Straten (1990). "Set Membership to Identification and Prediction of Lake Eutrophication." Water Resources Research **26**(11): 2643-2652.

Kent Thorton, W., H. Robert Kennedy, et al. (1982). "Reservoir Water Quality Sampling Design." Water Resources Bulletin **18**(3): 471-480.

Lerman, A. (1974). "Eutrophication and Water Quality of Lakes: Control by Water Residence Time and Transport to Sediments." Hydrological Sciences **1**(3): 25-35.

Liebhold Andrew, M., E. Rossi Richard, et al. (1993). "Geostatistics and Geographic Information Systems in Applied Insect Ecology." Annual Entomol. **38**: 303-327.

Maitima, J. M. (1991). "Vegetation Response to Climatic Change in Central Rift Valley, Kenya." Quaternary Research **35**: 234-245.

Marani, A. (1988). Advances in Environmental Modelling. Amsterdam Oxford New York Tokyo, Elsevier.

Mark Varljen, D., J. Micheal Barcelona, et al. (1999). "A Jacknife Approach to Examine Uncertainty and Temporal Change in the Spatial Correlation of a VOC Plume." Environmental Monitoring and Assessment **59**: 31-46.

McBratney, A. B., R. Webster, et al. (1981). "The Design of Optimal Sampling Schemes for Local Estimation and Mapping of Regionalized Variables - 1." Computers and Geosciences **7**(4): 331-345.

Paul Whitfield, H. (1983). "Evaluation of Water Quality Sampling Locations on the Yukon River." Water Resources Bulletin **19**(1): 115-122.

Perk, M. v. d. (1998). "Calibration and identifiability analysis of a water quality model to evaluate the contribution of different processes to the short-term dynamics of suspended sediment and dissolved nutrients in the surface water of a rural catchment." Hydrological Processes **12**: 683-699.

Rast, W. and A. Jeffrey Thornthon (1996). "Trends in Eutrophication Research and Control." Hydrological Processes **10**: 295-313.

Ryding, S. O. and W. Rast (1989). The Control of Eutrophication of Lakes and Reservoirs, The Parthenon Publishing Group.

Shafer, J. M. and M. D. Varijen (1990). "Approximation of Confidence Limits on Sample Semivariograms From Single Realizations of Spatially Correlated Random Fields." Water Resources Research **26**(8): 1787-1802.

Stein, A. and G. Sterk (1999). "Modeling space and time dependence in environmental studies." International Journal of Applied Earth Observation and Geoinformation **1**(2): 109-121.

Stein, A., C. Varekamp, et al. (1995). "Zinc Concentrations in Groundwater at Different Scales." Journal of Environmental Quality **24**(6): 1205-1214.

Tegaye, T. A. (1988). The Hydrogeological System of The Lake District Basin, Cental Main Ethiopian Rift.

APPENDIX 1: SIMULATED ANNEALING

Spatial Simulated Annealing

The central concept behind SSA is the fitness function $\phi(S)$ that has to be optimized. The process starts with an initial sampling scheme S_o, consisting of randomly drawn locations over the sampling area. Subsequently, an alternative sampling scheme S_1 is derived from S_o by a transformation of one of the sampling locations over a random vector. The probability of S_1 being accepted as a basis for further optimization can be described as:

$$P_c(S_i \rightarrow S_{I+1}) = 1, \qquad \text{if } \phi(S_{I+1}) \leq \phi(S_I)$$

$$P_c(S_i \rightarrow S_{i+1}) = \exp\left(\frac{\phi(S_i) - \phi(S_{i+1})}{c}\right), \qquad \text{if } \phi(S_{I+1}) > \phi(S_I) \qquad \text{.....(A1)}$$

Where c represents a positive control parameter, which is lowered according to a cooling schedule as the process evolves, to find a global minimum. A transition takes place if S_{I+1} is accepted. Next a solution S_{I+2} is derived from S_{I+1} and the probability ($S_{I+1} \rightarrow S_{I+2}$) calculated with a similar acceptance criterion as equation (A1).

In this study two optimization criterions were used firstly, the minimization of the mean shortest distance criterion (MMSD). By minimizing the average distance between an arbitrarily chosen point in the area of interest, and its nearest observation point, this criterion allows for the spreading of the observations over the sampling area. The minimization function in this case is estimated by:

$$\phi_{MMSD}(S) = \frac{1}{n_e}\sum_{j=1}^{n_e}\left\| \bar{x}_e^{\,j} - V_s\left(\bar{x}_e^{\,j}\right) \right\| \qquad \text{.....(A2)}$$

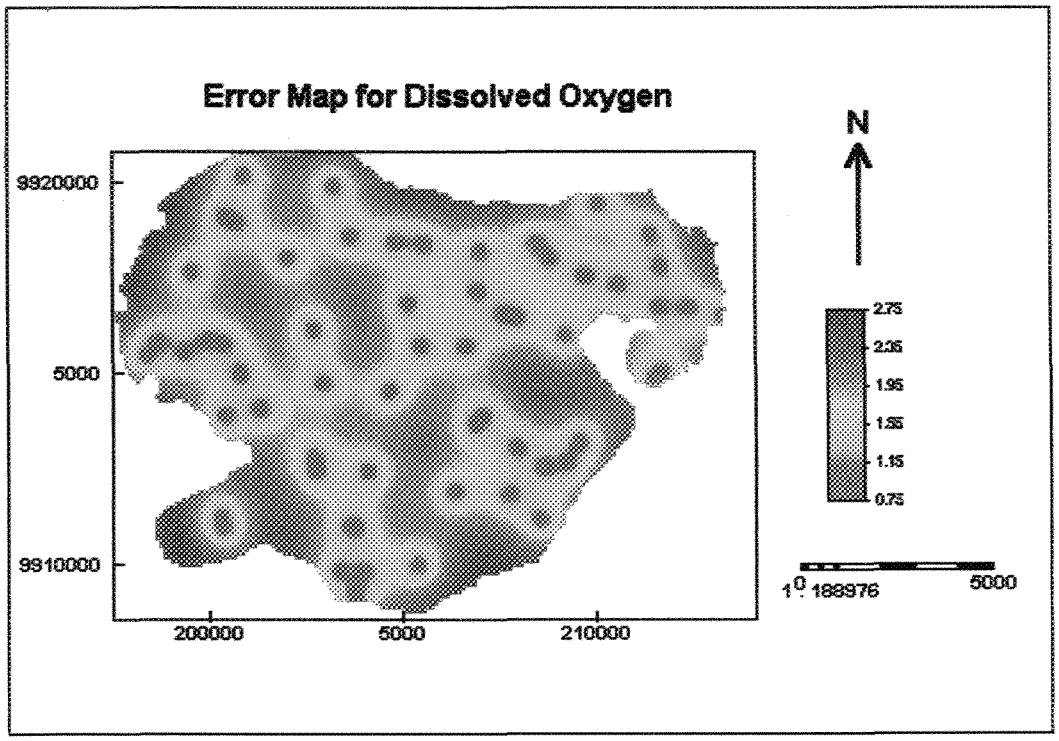
Where V_s ($\bar{x}$) represents the location of the nearest sampling point, $\bar{x}_e^j$ denote the nodes of a fine raster grid over the sampling area and n_e denote the number of raster nodes.

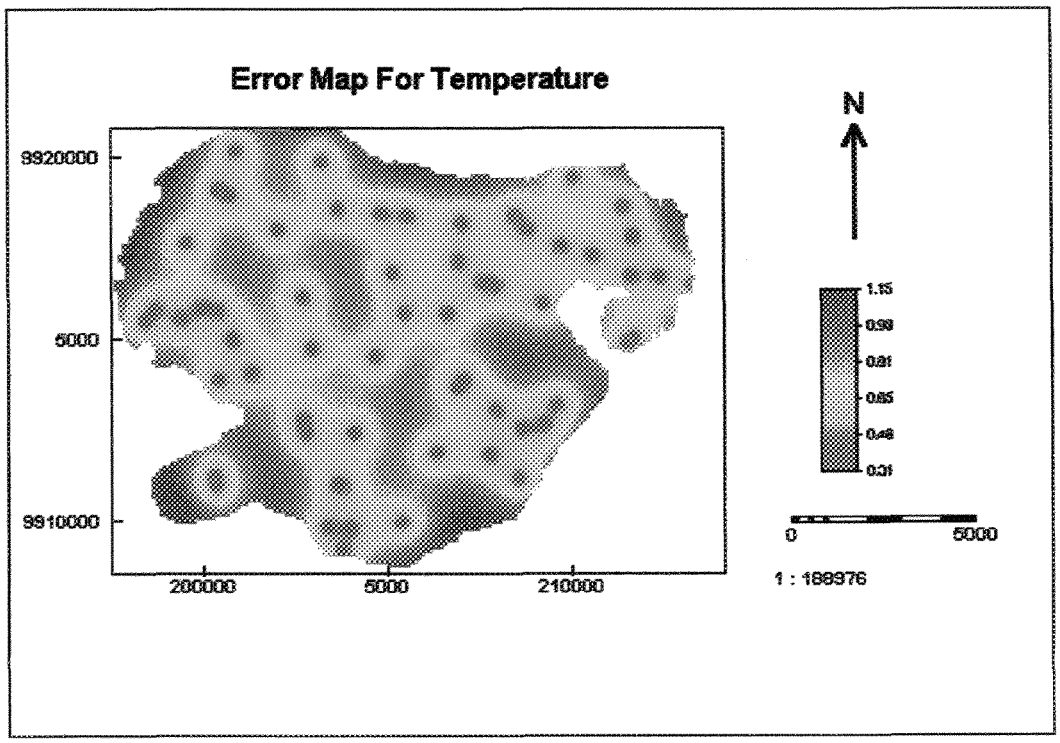
In the second instance, the minimizing kriging variance criterion (MK) was used; the aim of this criterion is the minimization of the of the ordinary kriging variance over the area of interest. The minimization function in this case is estimated using:

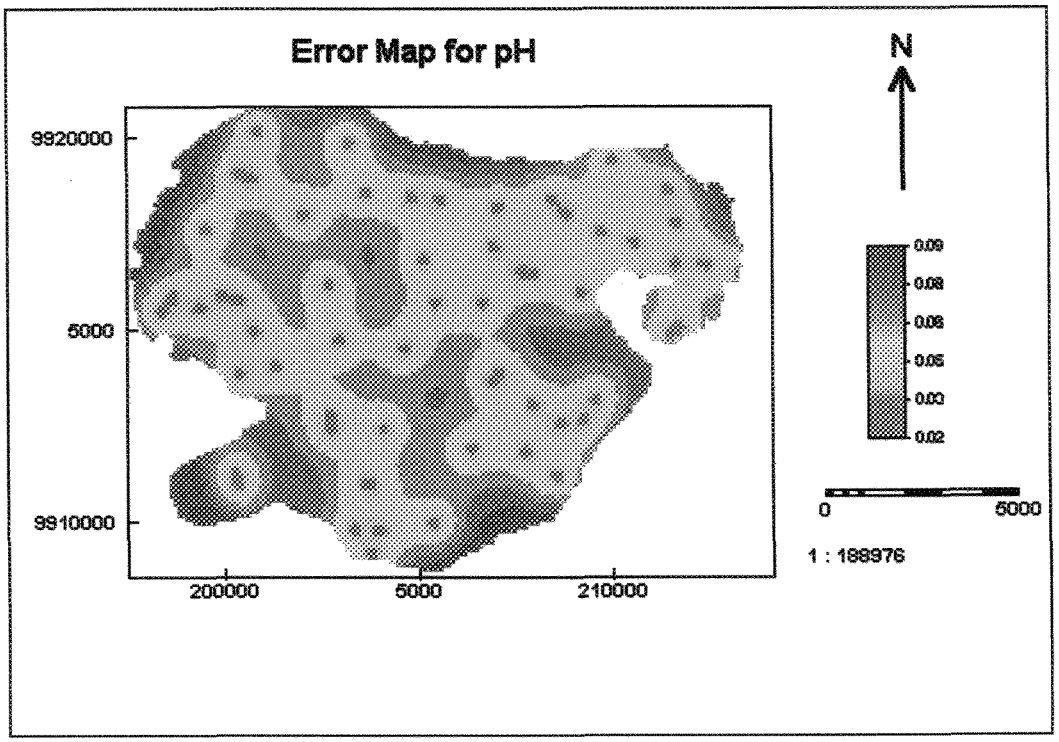
$$\phi_{MK}(S) = \frac{1}{n_e}\sum_{j=1}^{n_e}\sigma_{Mk}^2\left(\bar{X}_j|S\right) \qquad \text{.....(A3)}$$

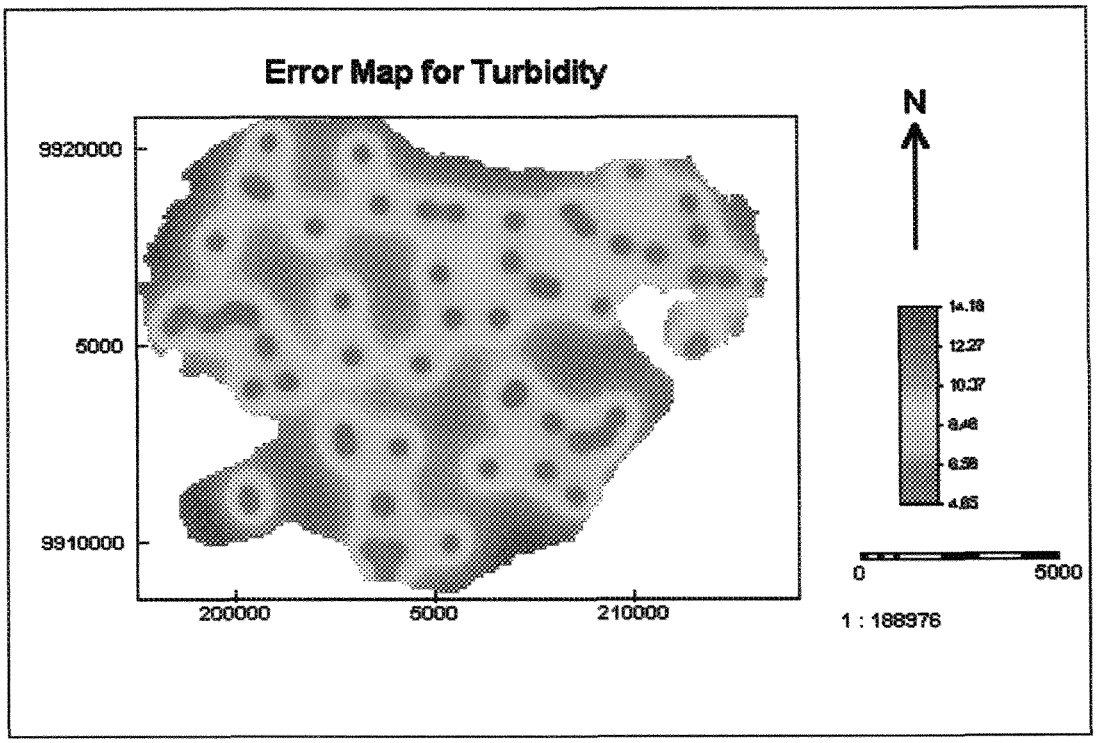
Where X denotes the location vector, σ_k^2 represents the kriging variance and the other terms were defined previously.

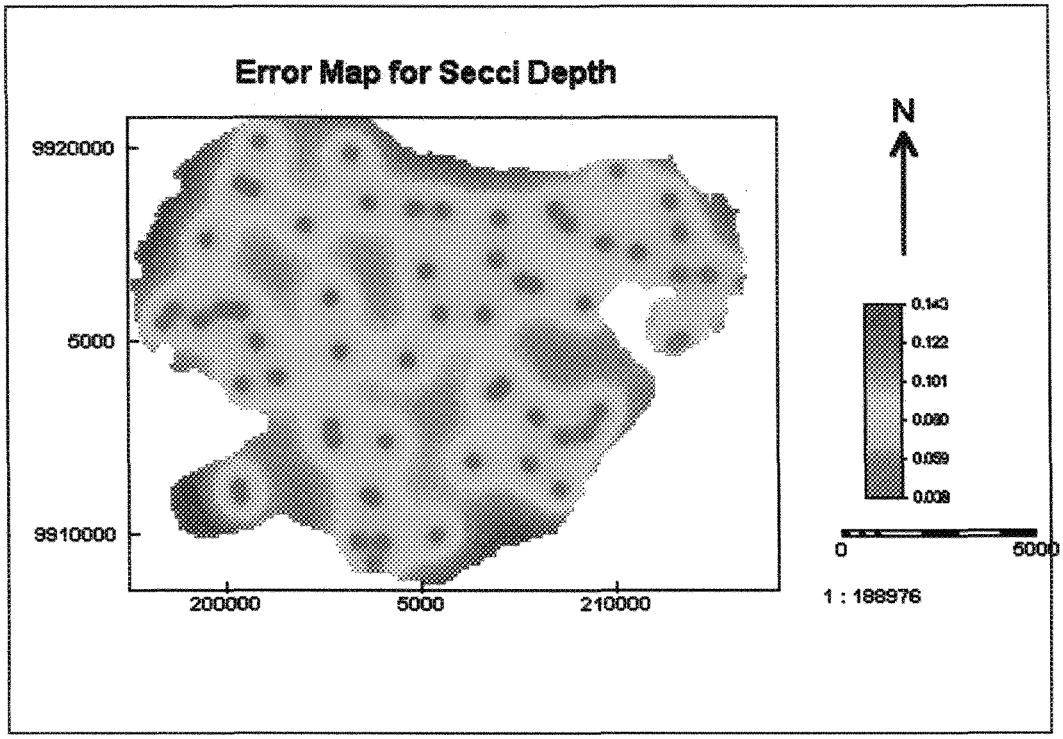
APPENDIX 2: ERROR MAPS OF PREDICTION



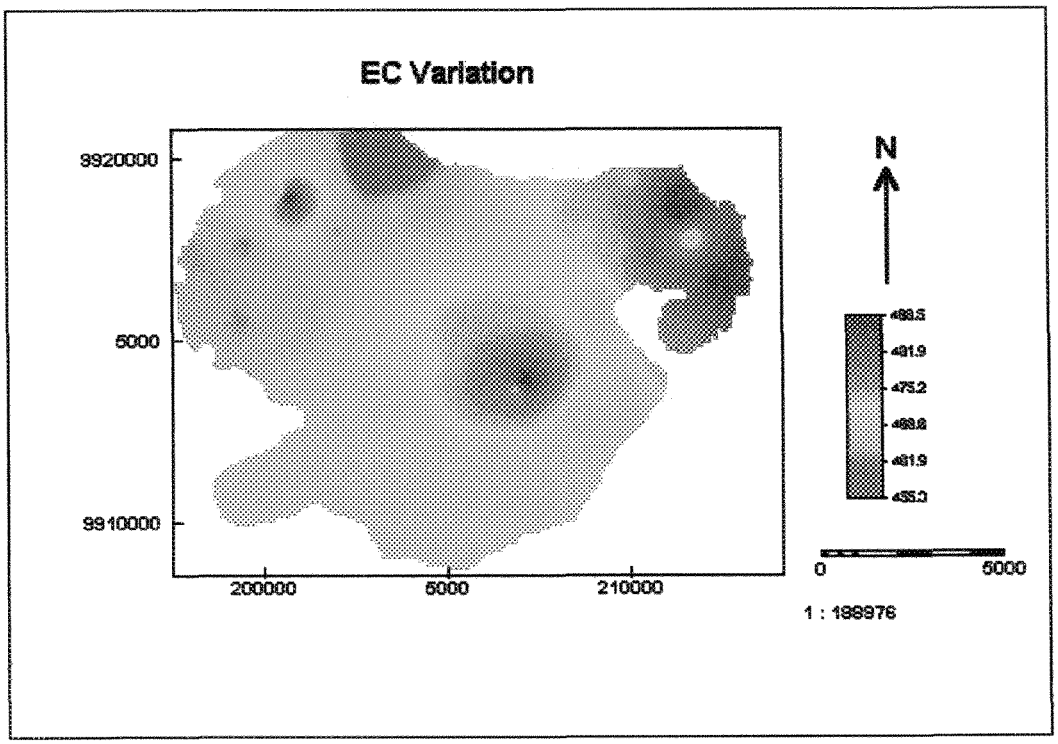



Error Map for pH
N
9920000
5000
9910000
200000
5000
210000
0.09
0.08
0.06
0.05
0.03
0.02
0
5000
1 : 188976


Error Map for Turbidity
N
9920000
5000
9910000
200000
5000
210000
14.18
12.27
10.37
8.46
6.56
4.65
0
5000
1 : 188976


Error Map for Secci Depth
9920000
5000
9910000
200000
5000
210000
N
0.143
0.122
0.101
0.080
0.059
0.038
0
5000
1 : 188976

APPENDIX 3: SPATIAL VARIATION IN EC



APPENDIX 4: DECISION MATRIX CALCULATIONS

The decision matrix was developed using the random sampling formula (equation 2.14)

$$n = \frac{t^2 s^2}{d^2}$$

as previously defined to give an initial estimate of the number of samples.

The variance for phosphorus, s2, as obtained from this study = 31(μgP/l) $^{2.}$

Assume a desired precision level, d, to be, e.g., ±5μgP/l of a mean phosphorous concentration of 194μgP/l. Assume also a 90 percent probability of being within the desired precision level of the mean Since the student's t value for this probability level varies as a function of sample size, an n = 30 is used to initialize the sampling formula so

$$n = \frac{(1.697)^2_{30,0.10}(31)}{(5)^2} = \frac{(2.88)(31)}{25} = 3.57$$

Round n to the next larger integer, 4; enter the t table at α=0.10 for n = 4 and repeat:

$$n = \frac{(2.132)^2_{4,0.10}(31)}{(5)^2} = \frac{(4.545)(31)}{25} = 5.6$$

Round to 6, re-enter the t table and repeat:

$$n = \frac{(1.943)^2_{1,0.10}(31)}{(5)^2} = 3.17 \text{ (Convergence at n=5)}$$

To be within ±5μgP/l of the mean epilimnion P concentration 90 percent of the time, therefore, requires 5 randomly collected samples. Similar analyses performed for turbidity, and chlorophyll to be with a desired precision level 90 percent of the time, indicated 27 and 11 samples are randomly collected.

APPENDIX 5: SYMBOLS

Water	SSW	g/m3	Suspended solids concentration
Water	TIPW	g-P/m3	Inorganic P water column
Water	TOPW	g-P/m3	Organic P water column
Water	TONW	g-N/m3	Organic N water column
Water	NH4W	g-N/m3	Ammonia N water column
Water	A	g-C/m3	Algal biomass
Water	NO3W	g-N/m3	Nitrate N water column
Bottom	AB	g-c/m3	total algal biomass in sediments
Bottom	TIPB	g-P/m3	Inorganic P sediment
Bottom	TOPB	g-P/m3	Organic P sediment
Bottom	TONB	g-N/m3	Organic N sediment
Bottom	H4B	g-N/m3	Ammonia N sediment
bottom	NO3B	g-N/m3	Nitrate N sediment
parm	Is	W/m2	Optimal light intensity
parm	achlc	ugChl/mgC	Chlorophyll to Carbon ratio
parm	umax	1/day	Maximum growth rate
parm	kres	1/day	Respitation rate
parm	kdie	1/day	Die-off rate
parm	tra	-	Temperature coeff die-off
parm	Tga	-	temperature coeff of growth
parm	kn	g-N/m3	Nitrogen monod constant
parm	kp	g-P/m3	Phosphorus monod constant
parm	Vsa	m/day	Settling velocity
parm	Vss	m/day	Fall velocity suspended solids
parm	POR	-	Sediment porosity
parm	RHO	kg/m3	Density suspended solids
parm	HB	m	Depth of sediment top layer
parm	Edif	m2/day	Diffusive exchange
parm	KpipW	m3/g SS	Partition constant P water column
parm	KpipB	m3/g SS	Partition constant P sediment
parm	fdpoW	-	Fraction DOP water coloumn
parm	fdpoB	-	Fraction DOP sediment
parm	TIPLB	g/m3	Inorganic P lower sediment layer
parm	TOPLB	g/m3	Organic P lower sediment layer
parm	fporg	-	Fraction organic P released by respiration
parm	apc	mgP/mgC	Phosphorus to Carbon ratio
parm	fdnoW	-	Fraction dissolved organic N water column
parm	fdnoB	-	Fraction dissolved organic N sediment
parm	TONLB	g-N/m3	Organic N lower sediment layer
parm	fnorg	-	Fraction organic N released by respiration
parm	anc	mgN/mgC	Nitrogen to Carbon ratio
parm	NH4LB	g-N/m3	Ammonia N lower sediment layer
parm	Kmn	g-N/m3	Ammonia preference constant
parm	Knit	1/day	Nitrification rate constant
parm	tnit	-	Temperature coefficient nitrification
parm	Kno	mg-O2/m3	Oxygen half sat. constant nitr.
parm	NO3LB	g-N/m3	Nitrate lower sediment layer

parm	KdenB	1/day	Denitrification rate constant sediment
parm	tdenB	-	Temperature ceoefficient denitrification sediment
parm	KdaB	1/day	Anaerobic decay algae sediment
parm	tdaB	-	Temperature coefficient algal decay sediment
parm	KminB	1/day	Anearobic decomposition rate
parm	tminB		Temperature coefficient anaerobic decomposition
parm	Kmin	1/day	Decomposition rate organic matter water column
parm	tmin	-	Temperature coefficient decomposition
parm	ma	g alg/g C	Biomass to Carbon ratio algae
parm	e0	m-1	Background extinction
parm	ealg	ug Chl/lm	Specific extinction
parm	eSS	mg SS/l m	Specific extinction susp-solids
parm	Sd0	m	Background secchi depth
parm	Sdads	-	Contribution of gelbstoff to inverse secchi depth
parm	Sdalg	m-1mg-1m3	Contriution of algae to inverse secchi depth
xt	Fres	g/m2 day	Resuspension flux
xt	T	oC	Temperature
xt	Ia	W/m2	Average light intensity
xt	L	hour	Day length
xt	Ads	m-1	Adsorption at 380 nm
flow	Z	m	Depth