2017-08-14

SOFTWARE-BASED TEACHING AID FOR DEMONSTRATING STOCHASTIC INTE)-~,POLATION METHOD BY ESTIMATING BRECIPITATION OVER LAKE TANA FOR 1996-2006

Yihenew, Getinet

A SOFTWARE-BASED TEACHING AID FOR DEMONSTRATING STOCHASTIC INTERPOLATION METHOD BY ESTIMATING PRECIPITATION OVER LAKE TANA FOR 1996-2006

Yihenew Getinet

BAHIR DAR UNIVERSITY
BAHIR DAR, ETHIOPIA

MAY 2009

A SOFTWARE-BASED TEACHING AID FOR DEMONSTRATING STOCHASTIC INTERPOLATION METHOD BY ESTIMATING PRECIPITATION OVER LAKE TANA FOR 1996-2006

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements for the Degree of
MASTER OF EDUCATION IN PHYSICS

By

Yihenew Getinet

AT
BAHIR DAR UNIVERSITY
BAHIR DAR, ETHIOPIA
MAY 2009

BAHIR DAR UNIVERSITY
DEPARTMENT OF
PHYSICS

The undersigned hereby certify that they have read and recommend to the Faculty of Education for acceptance a thesis entitled **"A software-based teaching aid for demonstrating stochastic interpolation method by estimating precipitation over Lake Tana for 1996-2006"** by **Yihenew Getinet** in partial fulfillment of the requirements for the degree of **Master of Education in Physics**.

Dated: May 2009

Advisor: ______________________
Baylie Damtie (PhD)

Committee Chair: ______________________
Tsegaye Kassa

External Examiner: ______________________
Habtu Zegeye (PhD)

Internal Examiner: ______________________
Geberegziabeher Kahsay (PhD)

BAHIR DAR UNIVERSITY

Date: **May 2009**

Author: **Yihenew Getinet**

Title: **A software-based teaching aid for demonstrating stochastic interpolation method by estimating precipitation over Lake Tana from 1996-2006**

Department: **Physics**

Degree: **M.Ed** Year: **2009**

Permission is herewith granted to Bahir Dar University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

Dedicated To

Kassahun Tilahun, Dr. Baylie Damtie, Mekedim Tibebu

Acknowledgements

I would like to thank all people working in the Washera Geospace and Radar Science Laboratory for their technical and moral support during the course of this work.

In particular, I would like to record a special thanks and sincere appreciation to my advisor Dr. Baylie Damtie for his invaluable suggestions, proper advice and dedication of his time in reading and correcting this thesis. His guidance has been practical, rigorous, matter-of-fact and realistic. I have been very fortunate and privileged to have had his counsel.

My sincerest appreciation and gratitude are extended to my father Ato Getinet Getu, my mother W/O Yeshialem Emekulu and my friend Kassahun Tilahun along with my younger brothers and sisters whose unconditional love, support and encouragement greatly motivated me to do this research. Kassishmar, thank you for your financial, material, academic and moral assistance that helped to sustain me during my academic pursuits. To all of you, I extend a thankful and a loving heart.

Thanks go to head of BDU physics department, Ato Kassa Demelash, Ato Zewedu Kebede, principal of Tana Haik Secondary School, Dr. Krishena, my friends Abiyu Zerfu and Gashaw Moges for their unconditional support and encouragement at times of difficulties and hardship.

I owe special thanks to Mekedim Tibebu, she gave me her computer through out my academic years in the graduate school. Mekedie you have a special place in my heart.

In addition to practical assistance, I have been encouraged and given moral support by many colleagues and friends, for which I am very grateful. Though difficult to list all of them, I specially mention Muluken, Dessu,Yirgalem, Muluye, Moges, Alex, Danie,Dejenie, Melesew, Ashu, Sisay, Kaleab, Biruh, Yibeltal, Bayeleyegne, Simeneh, Tesfaye, Gutema, Melese, and Melak.

Table of Contents

List of Tables

List of Figures

Abstract

Various kinds of estimation techniques are often used in science and engineering to extract optimal information from a measurement. Thus, many undergraduate and graduate curricula incorporate these methods and students are expected to understand and apply them in solving real scientific and engineering problems. This kind of pedagogical goal can be achieved by involving students in solving real problems using real measurements instead of presenting the topics in the traditional talk and chalk method. In the Ethiopian context, it is not easy to find real instrument to carry out measurements. In the present thesis, we have developed a software-based teaching aid that implements the stochastic interpolation method (traditionally called Kriging method) using real rainfall data from the Lake Tana watershed. The amount of rainfall around Lake Tana shows substantial spatial variability. Capturing such characteristics is challenging and this means that one needs to establish a very dense network of rain gauge over the lake and its watershed. However, we can estimate such spatial variations of rainfall by means of interpolation. In this study, the mean monthly precipitation on different locations on the Lake Tana is estimated using the Kriging interpolation method by using measurements collected at eight rain gauge stations from the lake catchment in the years 1996 to 2006. Based on this analysis, we found that Lake Tana gets on average 1337mm rainfall annually. A 0.97 correlation coefficient was obtained between estimated precipitation and runoff. The present thesis contains two major results. First it contains a virtual teaching aid with an interactive user interface that provides students and teachers a unique opportunity to learn stochastic interpolation technique by investigating real measurements collected from their own surroundings. Secondly, the thesis shows how one can improve rainfall estimates over Lake Tana and thereby increasing our understanding of its hydrology.

Chapter 1

INTRODUCTION

1.1 Background Information

Teaching of science and engineering present various pedagogical challenges around the world. Questions like how to attract students to these subjects, how to motivate them in learning and also how to teach these subjects efficiently and the like may be dealt in different manners. Most educators recommend practically assisted lecture, where teachers' are expected to demonstrate the theoretical aspects of their lecture via experimentation that allow students to actively participate. Such instruction model is hard to implement in practice in the developing countries including Ethiopia. This is due to the lack of appropriate scientific instruments and skilled personnel among other things. Thus, the instructions are usually highly theoretical and such conditions are prevalent in Ethiopia. As a result, a new paradigm shift has been proposed by the Ethiopian Ministry of Education, where students' active participation in the classroom are highly encouraged. The lack of scientific instruments can still make the new program ineffective. The Washera Geospace and Radar Science Laboratory (WaGRL) at Bahir Dar University has been developing different virtual teaching aids that can be used in lieu of real instruments and some of these systems are available in [1]. One may be concerned with the effectiveness of these kinds of teaching aids and research conducted at WaGRL showed that, in some cases these systems can even be more effective than a real scientific instrument [2].

The purpose of the present thesis is to develop a software-based teaching aid that can be used to demonstrate stochastic interpolation method using real rainfall measurements from the Lake Tana watershed. The system provides students a unique opportunity to learn about the method using hand on experience and it also gives a very flexible environment for them to explore different scenarios. In the following sections, we present the background information about the Lake Tana and the problem associated to estimate the rainfall amount over the Lake Tana. This shall be followed by the technical presentation of the stochastic interpolation method and the associated analysis software.

1.2 Background Information on Lake Tana

Lake Tana is one of the world's most important inland water bodies. It is the largest fresh water body in Ethiopia with a maximum depth of 14m and mean depth of 9m. It is situated around $11^{0}36'$N and $37^{0}23'$ E with average altitude of 1786m above seal level. The lake is estimated to have a surface area of about 3156 sq.km. It is the largest lake in Ethiopia and the third largest in the Nile Basin. Lake Tana is the source of the Blue Nile River, which is the only out flowing river from the lake that carries more than $(\frac{2}{3})^{rd}$ of the total volume of the Nile River at Khartoum, Sudan. The drainage of Lake Tana basin measured at the outlet of the Lake at Bahir Dar has a size of 15, $320km^2$. There are more than 40 rivers feeding the lake and Gilgel-Abay, Reb, Gumera and Magech contributes more than 93 percent of the inflow [3]. The lake is bordered by low plains in the north by Dembia, east by Fogera and southwest by Kunzila, which are often flooded in the rainy season (forming extensive wet lands).

The seasonal distribution of rainfall over the lake is controlled by the northward and southward movement of the inter-tropical convergence zone (ITCZ henceforth). Moist air masses are driven from the Atlantic and Indian Oceans during summer (June-September). During the rest of the years the ITCZ shifts southwards and dry conditions persist in the region between October and May. Generally, the southern part of the Lake Tana basin is wetter than the western and the northern parts.

Lake Tana is believed to have been formed due to Damming by lava flow during the Pliocene [4]. But the formation of the depression itself started in the Miocene [5]. The lakes bottom consists of volcanic basalts, usually covered with a muddy substratum with only little organic matter. At some places, volcanic peaks in the lake bottom form reefs, or even islands [6].

1.3 Statement of the Problem

Stochastic interpolation method is very important technique to infer missing information from other relevant neighboring measurements. This technique is included in many science and engineering curricula and students' are expected to understand it and use it in solving real problems. However, the presentation of this topic is usually highly theoretical in the Ethiopian context and teachers' have little alternative to change their delivery method. As a result, most students hardly appreciate this method and use it in problem solving. This is evident especially when one examines the lack of sufficient publications by Ethiopian authors using this method.

The present thesis presents a software-based system that can be used to demonstrate stochastic interpolation using real rainfall measurements from the Lake Tana watershed. The Ethiopian Meteorological Service has several gauging stations around the lake and the rainfall data is available. These data are often used in hydrological modeling and especially to analyze the behavior of the lake in response to changes in climate and human induced activities.

Lake Tana as the source of the Blue Nile contributes more than 80 percent of the Nile's supply of fresh water to the northern countries Sudan and Egypt originating from Ethiopian highlands [7]. Its variability plays a regulatory role for the annual Nile flow, hence any changes in the lake's water balance have significant consequences for riparian countries dependent on Nile water. Recent assessments of the hydrological system of Lake Tana basin show uncertainties with respect to major lake water balance terms. A major source of uncertainties in hydrological modeling is the specification of precipitation. This is not only the largest term in the water balance, but it is also very difficult to assess. One source of error appears to be associated with the spatial variability of precipitation around the lake catchment as shown in Figure 1.1. The accurate estimation of precipitation over the lake requires a very dense network of rain gauge, which entails large installation and operational costs. Also the failure to take samples or read gauges may result in even lower sampling density. It is thus necessary to estimate point precipitation at unrecorded locations from values at the adjacent sites. Even though, a number of methods have been proposed for the interpolation of precipitation over a given area, selecting the optimal interpolation technique in some sense is vital to reduce the estimation error and uncertainty. This problem provides an ideal scenario to develop a software-based teaching aid for demonstrating how the stochastic interpolation technique works by solving real problems.

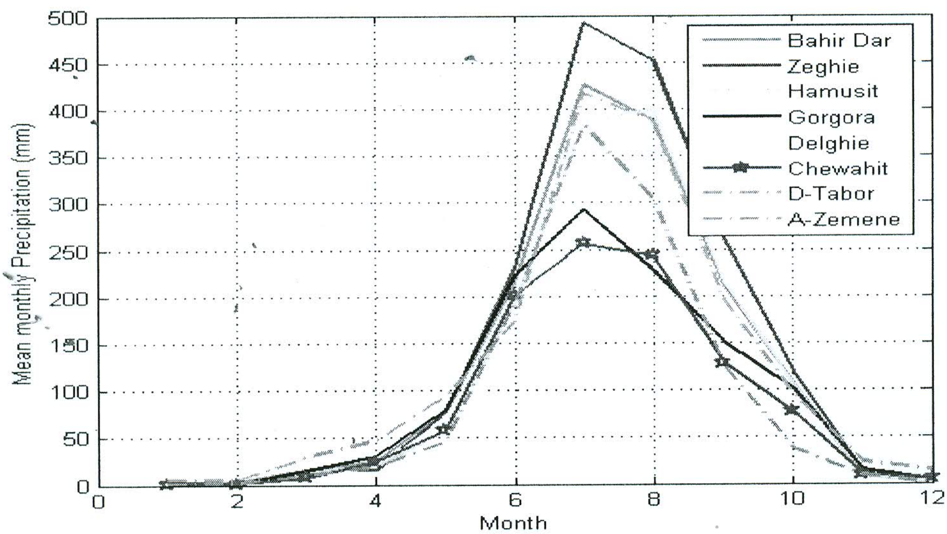

Figure 1.1: Mean monthly spatial variation of rainfall around Lake Tana catchment from 1996 to 2006.

1.4 Objectives of the Study

The objective of the present thesis is to develop a software-based teaching aid for demonstrating stochastic interpolation technique (Kriging method) using real rainfall measurements from the Lake Tana watershed. In addition to this pedagogical objective, its scientific objective is to show the spatial variation of rainfall over Lake Tana using sparsely located rainfall measurements around the lake and thereby leading to a better water balance calculation.

1.5 Significance of the Study

The significance of this study includes the following items.

- The thesis presents a very flexible and dynamic environment for learning or teaching stochastic interpolation technique using real rainfall measurements from the Lake Tana.

- The thesis has a potential to be a springboard for the development of similar systems for other topics in science and engineering and thereby facilitating active learning.

- The thesis investigates the spatial variability of the amount of rainfall over Lake Tana and thereby leading to a more optimal water balance calculation and other hydrological modeling.

- The information on spatial dependance of rainfall amount over Lake Tana can lead to other investigations like how rainfall amount varies with latitude, longitude and altitude.

- It can lead to investigations of the effectiveness of the present system pedagogically in teaching stochastic interpolation technique.

1.6 Research Questions

This work addresses the following research questions:

1. Can we develop a software-based and pedagogically sound teaching aids for demonstrating stochastic interpolation technique with interactive capability?

2. How can we estimate the amount of rainfall at any point over Lake Tana using sparsely located point measurements around the lake? What is the optimal estimator?

3. How do we validate our model?

Chapter 2

REVIEW OF RELATED LITERATURE

2.1 Introduction

Experimentally-assisted lectures are highly recommended by many educators around the world. Especially in science and engineering subjects hand on experience is vital. It is widely acknowledged that science and technology education in Ethiopia is highly theoretical and a new paradigm shift has been proposed to tackle the problem. One of the major pedagogical shift that has been put in practice is the students' center approach in teaching. This scenario assumes that, students are learning their subject matter by doing instead of the traditional talk and chalk method, which is highly teacher centered. The practical implementation of student centered instruction demands many inputs including various instruments and systems that engage students in many different activities. Such facilities are not easy to find for all the topics in the curriculum and thereby becoming major obstacles for the student centered instruction. However, modern computers provide unprecedented opportunities to develop various kinds of systems that can be used to engage students in practical work.

Many researchers around the world developed different virtual systems that allow students to be actively involved in their learning [8]. In Ethiopia the Washera Geospace and Radar Science Laboratory (WaGRL) has pioneered such endeavors and many virtual teaching aids, which can be used to teach various topics by actively involving students, have been developed. The effectiveness of such systems when compared with real instruments has also been investigated in this laboratory and it was found that in some cases such a system can be more effective [2]. This results encouraged WaGRL to continue developing such kinds of virtual systems including the three-dimensional virtual teaching aids that can be used to teach the physics of the Earth's magnetic field [9]. The present work is the continuation of this effort of WaGRL. In addition, we also have a scientific objective and thus the basic scientific works are reviewed in the following sections.

2.2 Precipitation Formation

Rain is formed by the conversion of atmospheric vapor into water. The water so formed then falls to the earth in the form of rainfall. In terms of hydrology rainfall constitutes the third phase of atmospheric division of the hydrologic cycle. Precipitation is a general term which includes all forms of falling moisture namely rainfall, snowfall, sleet, hail etc.

During a dry season, evaporation loss is highly accelerated from all types of free water surfaces. The water lost in evaporation finds room in an air mass. It adds to the atmosphere vapor storage. Although evaporation loss is excessive in a hot season the capacity of the air mass is also more. This large quantity of vapor makes the air mass moist. The change of state from atmospheric vapor to water takes place when the capacity of the air mass to hold the vapor particles exceeds. There are two main reasons for the change of state [10]. The first one is that hot air mass has large capacity to hold the vapor particles in suspension. When by some means this moist and warm air mass cools down its capacity to hold vapor particles is reduced. Finally vapor precipitates in the form of a rainfall. The second reason is the variation in pressure also brings about the change of state from vapor to rainfall. The actual mechanism of precipitation is nucleation. Ice or water crystals are formed up on the floating particles in the air mass (e.g. dust particles, salt particles, etc), the small crystals then grow in size by combining with other crystals. A stage comes when they fall down on the earth as snow or as rain water.

2.3 Types of Precipitation

Broadly speaking three types of precipitation occurs when warm and moist air mass gets lifted and subsequently cooled namely, cyclonic precipitation, convective precipitation, and orographic precipitation.

Cyclonic precipitation occurs from lifting of air which converges into a low pressure area or cyclone. When a moving air mass is obstructed by a stationary cold air mass, the warm air mass rises up as it is lighter than the cold air mass. The lifted air mass cools down at high altitudes and a continuous rain fall occurs till the warm front passes over the cold air mass.

Due to some local effects air gets heated up and rises up in the atmosphere as it is lighter than the cold air surrounding the area. At high altitudes, it gets cooled and convective precipitation occurs. On the other hand orographic type of precipitation occurs when a moving air moist air mass is obstructed by a barrier like mountains, the warm air mass rises to high altitudes and gets cooled and finally precipitates.

2.4 Spatial Distribution of Precipitation

Rainfall often shows high spatial and temporal variability. Factors such as nearness to sea, presence of mountains, direction of wind, and altitude are amongst others factors responsible for the inequitable distribution of rainfall over large area. Regarding to Lake Tana, its spatial rainfall distribution is significant as shown in Figure 1.1, which is controlled by ITCZ. It generally shows decreasing pattern from south to north. In year 2003 the rainfall variation within the lake area reached up to 40 percent [11].

Measured precipitation data is the essential driving input of all processes involved in the ground phase of the water cycle. For example, the ability of obtaining accurate lake precipitation is very important in the calculation of lake water balance. The lack of reliable measurement of areal rainfall is a major problem for many researchers in the field of hydrology, direct rainfall measurements can be obtained at a point scale [12]. In plain areas, if weather stations are uniformly distributed and the rainfall measurements do not vary considerably around the mean, the arithmetic average can be an acceptable method for estimating the average rainfall into the area [13].

However, the rainfall process is known to exhibit a high degree of variability both in space and time and capturing such aspects over a given area is great difficulty. Based on the assumption that rain gauge measurements can reliably account for the true point rainfall, aerial rainfall estimates have been obtained traditionally through some suitable interpolation method and aggregation technique. These are based on the hypothesis that rainfall estimates at ungauged sites can be obtained as linear or non-linear combinations of the values measured at a number of instrumented locations using the appropriate weights [12].

Many researchers have evaluated various methods for point interpolation of climate data, and the stochastic (Kriging) method has been rated superior over techniques such as Thiesen polygon, inverse distance weighting, least-squares polynomial regression and isohytal maps [14, 15].

2.5 Stochastic Interpolation Techniques

Geostatistics can be used to describe the autocorrelation of one or more variables in the 1-D, 2-D, and 3-D spaces or even in 4-D space-time, to make predictions at unobserved locations, to give information about the accuracy of prediction and to reproduce spatial variability and uncertainty. The shape, the range, and the direction of the spatial autocorrelation are described by the variogram, which is the main tool in linear geostatistics [16].

A basic assumption in geostatistics is that a spatiotemporal process is composed of deterministic and stochastic components. The deterministic components can be global and local trends. The stochastic component is formed by a purely random and an autocorrelated part. An autocorrelated component implies that on average, closer observations are more similar than more distant observations. This behavior is described by the variogram where squared differences between observations are plotted against their separation distances [16].

Stochastic interpolation attempts to obtain unbiased, minimum variance estimates of precipitation at points where measurements are not available, as a function of measurements available at a number of gauged sites. Stochastic interpolation has also become an important tool for the estimation of areal rainfall, due to the linearity of the process of aggregation from point rainfall rates to areal volumes [12].

The early work on stochastic interpolation was an unbiased, minimum variance estimator, based on the assumption that the covariance structure of the variable to be interpolated is known. The major advantage of the new approach is the possibility of quantifying the uncertainty in the resulting precipitation surface or precipitation volume as a function of the number and position of gauges and of the assumed spatial covariance structure [17], [12] and the reference within.

In the early seventies, developed the theory of regionalized variables which provide the theoretical foundations of Kriges's more practical methods [18]. These works led to the development of Kriging method, the optimal linear interpolation technique derived by Krige, and of the field of geostatistics, a new branch of statistical theory [12]. The fundamental idea of D. Krige was to use the variogram for interpolation as a means to determine the magnitude of influence of neighboring observations when predicting values at unobserved locations. Basic linear geostatistics includes two main procedures: variography for modeling the variogram and kriging for interpolation.

Chapter 3

METHODOLOGY

3.1 The data

There are numerous rainfall stations around Lake Tana catchment. However, the stations are not distributed uniformly around the lake catchment. For example, there is no rain gauge in the southwest of the lake catchment. Differences in network density will obviously have impact on the accuracy of a real rainfall estimation. At the Ethiopian Meteorological Agency, Bahir Dar branch, monthly records of rainfall data are available for the stations around the lake catchment. Our analysis utilizes mean monthly values for eight stations around the lake catchment from 1996 to 2006 in order to estimate precipitation over the lake (see Table 3.1). For stations with incomplete records, missing data are replaced using a simple arithmetic mean of the simultaneous rainfall records of three near by and evenly spaced stations around the station with missing records. The lack of consistent rainfall data in the sample stations is also the major problem on the accuracy of aerial rainfall estimation. Rainfall data in each station can be made consistent by adopting double mass curve techniques [10]. Results

Table 3.1: Monthly mean rainfall (mm) over Lake Tana catchment from 1996-2006.

Stations	Bahir Dar	Zeghie	Hamusit	Gorgora	Delghie
January	2.0	3.8	0.5	0.9	0.5
February	2.9	0.13	0.8	1.0	0.9
March	15.2	15.8	13.5	15.1	7.0
April	24.7	15.8	19.4	30.1	19.9
May	77.0	76.7	56.3	79.3	66.7
June	213.5	230.7	191.0	221.9	139.6
July	425.7	492.3	383.7	293.0	213.7
August	387.5	451.8	399.5	227.0	219.0
September	211.9	261.4	216.0	150.8	116.7
October	105.5	119.2	104.6	100.6	71.6
November	10.5	12.25	7.1	14.9	10.7
December	3.4	1.6	0.0	4.3	1.1
Annual mean	123.32	140.1	116.0	94.9	72.3

of the double mass analysis (for detail see Appendix 1) showed that inhomogeneity in the data series of some stations is also a major problem. We transformed the station locations held using latitude and longitude to provide an equal area projection of cartesian coordinates. The average monthly precipitation in these stations is ranging from 0mm in December to 492.3mm in July as shown in Table 3.1.

3.2 Kriging Methods

Kriging is named after its first practitioner, the South-African mining engineer D.Krige in 1951. Kriging is a stochastic technique, based on two tasks, namely semi-variogram and covariance functions (spatial autocorrelation) and prediction of unknown values. The following assumptions are needed to apply the Kriging method:

- all values in the study area are the result of a random process .i.e. Kriging is based on random processes with dependence.

- stationarity is another assumption that is often reasonable for spatial data to obtain the necessary replication. For semi-variograms, stationarity is the assumption that the variance of the difference is the same between any two points that are the same distance no matter which two points we choose. So, the semi-variogram depends only on the distance between the measurements and not on the location of the measurements.

- Isotropy for spatial correlation exists when the semi-variogram does't vary according to direction.

The linear combination of Kriging prediction may be given by

$$Z(\mathbf{S}_0) = \sum_{i=1}^{N} \lambda_i Z(\mathbf{S}_i), \tag{3.2.1}$$

where $Z(\mathbf{S}_0)$ is the estimate of Kriging at the prediction location $\mathbf{S}_0$, $Z(\mathbf{S}_i)$ is the measured value at the i^{th} location, λ_i is the Kriging unknown weights assigned to measurement and N is the number of measured values neighboring the prediction location. $Z(\mathbf{S}_o)$ is calculated by solving the Kriging equations which will be discussed in the next section.

3.2.1 The Ordinary Kriging Model

Let's assume that $Z(\mathbf{S})$ is a random variable defined at location $\mathbf{S}$ and $(Z(\mathbf{S}), \mathbf{Z}\epsilon\Omega)$, be a first-order stationary random field in a spatial domain Ω. The most commonly applied ordinary Kriging model is given by

$$Z(\mathbf{S}) = \mu + \varepsilon(\mathbf{S}), \tag{3.2.2}$$

where.

S=(X,Y) is a location and $Z(\mathbf{S})$ is the value at that location. The model is based on a constant mean μ for the data (no trend) and random errors $\varepsilon(\mathbf{S})$ with spatial dependence.

The Kriging predictor is formed as a weighted sum of the data, i.e.

$$\hat{Z}(\mathbf{S}_0) = \sum_{i=1}^{N} \lambda_i Z(\mathbf{S}_i), \tag{3.2.3}$$

where $\mathbf{S}_0$ is the prediction location. A natural optimality criterion for optimal estimator is the mean square error (MSE), which can be formulated as

$$MSE(\hat{Z}(\mathbf{S}_0)) = E(Z(\mathbf{S}_0) - \sum_{i=1}^{N} \lambda_i Z(\mathbf{S}_i))^2. \tag{3.2.4}$$

This measures the average mean square deviation of the estimator from the true value. To ensure the predictor is unbiased for the unknown measurement, the sum of the weight λ_i must be equal to one, i.e.

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{3.2.5}$$

The solution to eqn. (3.2.4), constrained by unbiasedness gives the matrix formula

$$\mathbf{\Gamma} * \lambda = \mathbf{g}, \tag{3.2.6}$$

which stands for

$$\begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{N1} & \cdots & \gamma_{NN} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} * \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \\ m \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ \vdots \\ \gamma_{N0} \\ 1 \end{pmatrix}. \tag{3.2.7}$$

Most of the elements in eqn. (3.2.7) are filled when we know the empirical semi-variogram model, which will be discussed in the next subsequent sections.

The gamma matrix $\mathbf{\Gamma}$ in eqn. (3.2.7) contains the chosen theoretically modeled semi-variogram values between all pairs of sample locations, where γ_{ij} denotes the chosen theoretically modeled semi-variogram values based on the distance between the two samples identified as the i^{th} and j^{th} locations. The 1s and 0s in the bottom row and the rightmost column arise due to the unbiasedness constraint.

The vector $\mathbf{g}$ in eqn. (3.2.7) contains the chosen theoretically modeled semi-variogram values between each measured location and the prediction location,where γ_{i0} denotes the chosen theoretical modeled semi-variogram values based on the distance between the i^{th} sample location and the prediction location. The 1s in the bottom row arise due to the unbiasedness constraint.

Once the Γ matrix and the $\mathbf{g}$ vector in eqn. (3.2.7) are filled, the goal is to solve the equations for all of the λ_i (the weights), so the predictor can be formed by using $\sum_{i=1}^{N} \lambda_i Z(\mathbf{S}_i)$. We can compute the Kriging weights vector by performing the simple matrix algebra to obtain

$$\lambda = \mathbf{\Gamma}^{-1} * \mathbf{g}, \tag{3.2.8}$$

where $\mathbf{\Gamma}^{-1}$ is the inverse matrix of $\mathbf{\Gamma}$.

3.2.2 Empirical Semi-variogram

The empirical semi-variogram is an essential step for determining the spatial variation in the sampled variable. It provides useful information for interpolation, sampling density, determining spatial patterns, and spatial simulation. The semi-variogram is formulated by

$$\gamma(d_{ij}) = \frac{1}{2} E[Z(\mathbf{S}_i) - Z(\mathbf{S}_j)]^2, \tag{3.2.9}$$

where $Z(\mathbf{S}_i)$ is the measured value at the i^{th} location, and $Z(\mathbf{S}_j)$ is the measured value at the j^{th} location. And $\gamma(d_{ij})$ is semi-variogram, which is dependent on the Euclidean distance between the i^{th} and j^{th} locations, and d_{ij} can be described by

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \tag{3.2.10}$$

From the plot of the average semi-variance (along Y-axis) versus average distance of the bins (along X-axis), various properties of the model are determined.

- The range: The range is the distance where the model first flattens out. Sample locations separated by distances closer than the range are spatially autocorrelated, whereas locations farther apart than the range are considered as spatially independent observations.

- The sill: The value that the empirical semi-variogram plot attains at the range (the value on the Y-axis) is called the sill. The partial sill is the sill minus the nugget.

- The nugget: Theoretically, at zero separation distance (i.e. lag=0), the semi-variogram value should be zero. However, at an infinitesimally small separation distance, the difference between measurements often does not tend to zero. This is called the nugget effect. The nugget effect can be interpreted as the variation among repeated measurements at the same point.

The nugget effect can be attributed to measurement errors or spatial variation sources at a distance smaller than the sampling interval (or both) [16]. Measurement error occurs because of the error inherent in measuring devices.

Though, the empirical semi-variogram is an essential step in the Kriging estimation, their values cannot be used directly in the Kriging equation, because a negative

Kriging standard errors may be introduced in the predictions. Instead, a model is fitted through the empirical semi-variogram values for various distances.

3.2.3 Theoretical Semi-Variogram Model

Normally, a theoretical semi-variogram model is fitted through the empirical semi-variogram values for the distant classes or lag classes. The semi-variogram properties; the sill, range and nugget can provide insights on which model will fit the best [16]. The most common models include the following:

- Spherical model: When there is a clear range and sill, a spherical model often fits the empirical semi-variogram well. The formulation of this model is based on

$$\gamma(\mathbf{h}) = c_0 + c_1[\frac{3}{2}(\frac{h}{h_{max}}) - \frac{1}{2}(\frac{h}{h_{max}})^3], \tag{3.2.11}$$

for $h\varepsilon(0, h_{max})$. Here $\mathbf{h}$ is the lag vector, and h is the length of $\mathbf{h}$ (distance between two locations), c_0 is the nugget value, c_1 is the partial sill (sill-nugget) of the model, and h_{max} is the range value.

- Exponential model: If there is a clear nugget and sill, but only gradual approach to the range, the exponential model is often preferred, which is based on

$$\gamma(\mathbf{h}) = c_0 + c_1[1 - exp(\frac{h}{h_{max}})], \tag{3.2.12}$$

for $h\varepsilon(0, h_{max})$.

- Gaussian model: The Gaussian model is fitted if the variation is very smooth and the nugget variance is very small compared to the spatially random variation and the model equation is

$$\gamma(\mathbf{h}) = c_0 + c_1[1 - exp(-(\frac{h}{h_{max}}))^2], \tag{3.2.13}$$

for $h \varepsilon (0, h_{max})$

- Linear model with no sill: semi-Variograms with no sill within the sampled area are represented by the linear model

$$\gamma(\mathbf{h}) = c_0 + bh. \tag{3.2.14}$$

- Linear model with sill: Linear models with sill within the sampled area are represented by

$$\gamma(\mathbf{h}) = c_0 + c_1(\frac{h}{h_{max}}). \tag{3.2.15}$$

Chapter 4

RESULTS AND DISCUSSIONS

4.1 Analysis Procedure in a Nutshell

The analysis procedure we have followed may be summarized in the following manner:

- prepare the data in a manner suitable for analysis using the MATLAB routines we have developed (for detail see Appendix 2);
- analysis of empirical semi-variograms for selection of model parameters including, nugget, range, sill, and semi-variogram type;
- Kriging calculations;
- estimating the precipitation time series for Lake Tana.

Table 4.1 shows the chosen theoretical variogram model, the nugget effect and the maximum range for each month. Figure 4.1, for example, shows how we choose a theoretical variogram model from a variogram estimator (an experimental semivariogram plot) for April. The precipitation at site of interest for April can be estimated

Table 4.1: Type, Nugget effect, and Range of variograms for each month.

Month	Nugget	Sill	hmax (km)	Variogram type
January	0	3.0	52.46	Spherical
February	0	2.4	52.46	Spherical
March	0	37.83	52.46	Spherical
April	0	122.8	52.46	Spherical
May	0	391.4	52.46	Spherical
June	0	2.4	52.46	Spherical
July	0	2.4	52.46	Spherical
August	0	2.4	52.46	Spherical
September	0	1717	52.46	Spherical
October	0	873.9	52.46	Spherical
November	0	2.4	52.46	Spherical
December	0	33.55	52.46	Spherical

using the stochastic (Kriging) method in the following manner:

- First we calculate the semivariogram values and the corresponding distance between each pair of locations.
- We then plot the pairs formed by distance and semivariogram values on the x-axis and the y-axis, respectively.
- Since, the cloud of points (the scatter plot) are unmanageable, we group the pairs of locations within a specified range of distances (which is referred to as bin). The resulting average paired values (which is referred to as variogram estimator and the plot is empirical semivariogram plot) which gives us an idea of which theoretical model will fit best and the nugget effect, the sill and the range parameters are obtained from it.
- Spherical and exponential models are fitted into the variogram estimator as shown in Figure 4.1. One can see that the spherical model fits reasonably well

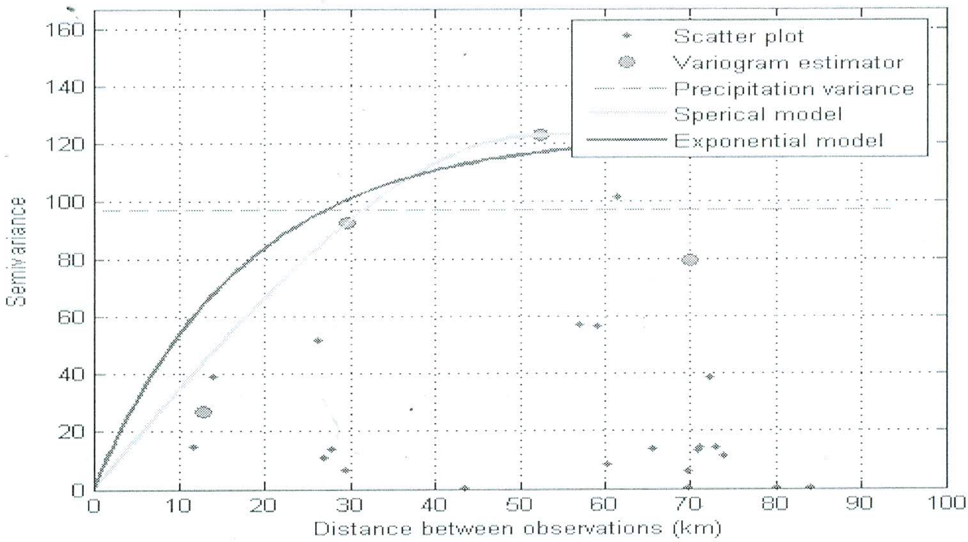

Figure 4.1: Fitting different theoretical model to a variogram estimator (empirical semivariogram) for April.

over the first three lags. Hence, we choose the spherical model equation in the Kriging calculations, because near by points are the most important for the Kriging estimate.

Figure 4.2 shows Kriging estimates of rainfall at different locations for April. This result was obtained by using the software that we have developed. The values on the X and Y axis are the transformed location values of longitude and latitude respectively for the gauge stations. The result shows the spatial variability of the amount of rainfall over Lake Tana.

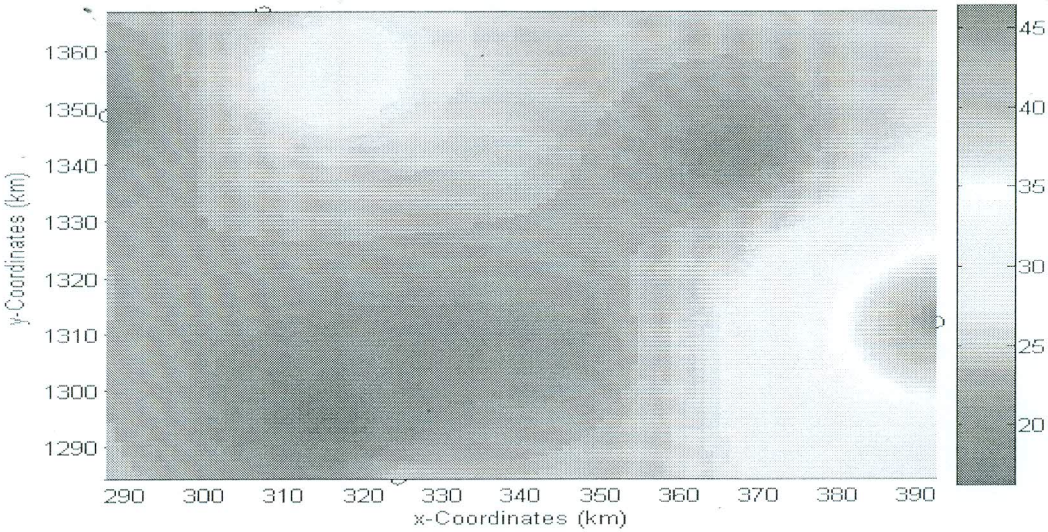

Figure 4.2: Kriging estimates of aerial precipitation over Lake Tana for April (1996-2006).

Following similar procedure for other months, precipitation estimates over Lake Tana can by obtained using the data in Table 3.1. The results are given in Table 4.2 and in Figure 4.3. Monthly values range from 0.7322mm in February to 369.0327mm in July. The estimate for annual precipitation over the lake was found to be 1337mm, which is nearly 7mm greater than the corresponding value in [11].

Table 4.2: Average monthly precipitation estimates (mm) using ordinary Kriging method over Lake Tana (1996-2006).

Month	Estimated precipitation(mm)
January	1.7461
February	0.7322
March	13.7069
April	21.3977
May	71.9398
Jun	205.0057
July	369.0327
August	339.9464
September	196.8778
October	102.9434
November	11.6330
December	2.0362
sum	1337

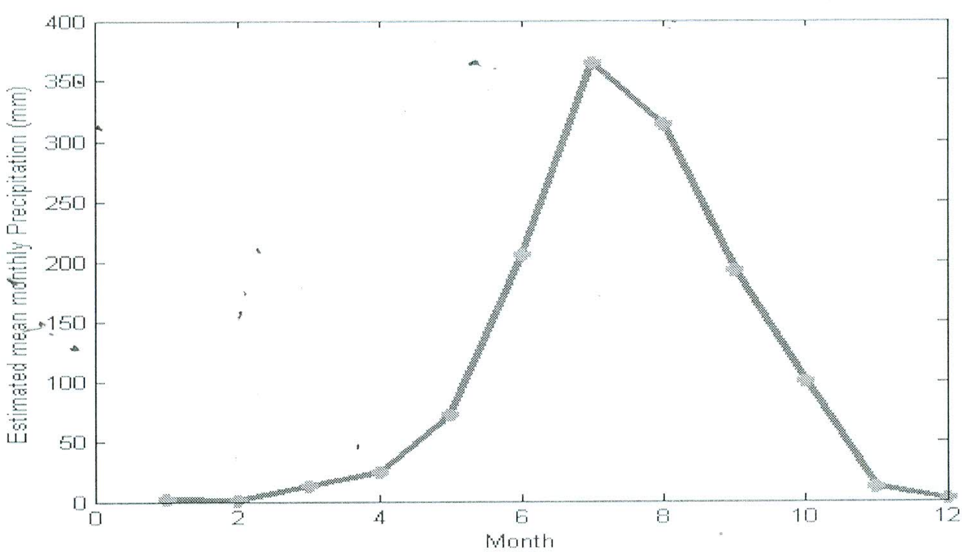

Figure 4.3: Mean monthly precipitation estimated over Lake Tana(1996-2006).

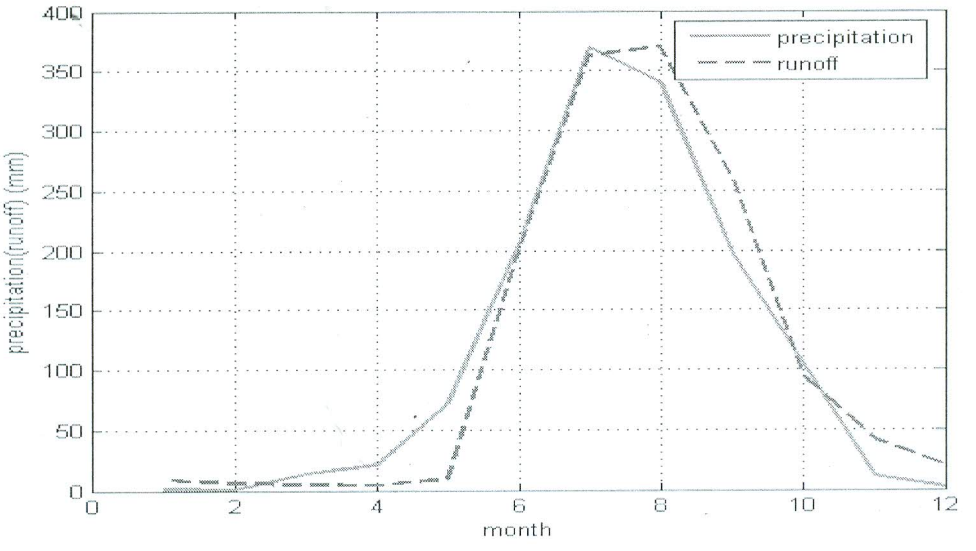

Figure 4.4: Mean monthly rainfall and runoff relation.

4.2 Model Validation

Model validation is one of the most important topics in science and engineering curriculum. Students should have the opportunity to explore the pros and cons of a model that has been presented to them theoretically. Such learning environment can be created in different ways. Model validation capability has been incorporated in the system we have developed in the present thesis. Due to the lack of reliable measurements of aerial rainfall, direct validation of the stochastic interpolation technique is hard to accomplish [12]. Hydrological rainfall-runoff models are widely used to transform the variability of rainfall in space and time to simulate the flow of river discharges into the basin [12]. Figure 4.4 shows the comparison between estimated mean monthly precipitation over Lake Tana and measured mean monthly runoff into the lake. One can see that the two curves have similar pattern.

This relation can be quantified by carrying out correlation analysis between estimated mean monthly precipitation and the corresponding mean monthly runoff. The result shows that there is a correlation coefficient of 0.97 between runoff data with the Kriging estimates. This means that the interpolation results can be considered as a more reliable estimates of the average aerial precipitation. These results were obtained by using the software-based teaching aid presented in this thesis and students can explore themselves by taking different scenarios.

Chapter 5

CONCLUSIONS

Many topics in science and engineering demand different teaching methodologies. The most widely accepted mode of delivery includes active learning, where students are involved actively. These kinds of dynamic learning environments can be realized in different ways depending on the nature of the topics and also the major pedagogical goals. In the present thesis, we have developed a software-based teaching aid for demonstrating the stochastic interpolation technique (Kriging method). First we have presented what one can learn from the rainfall data obtained around Lake Tana. Especially, the problem associated to determining the rainfall over Lake Tana due to the substantial spatial variability of the rainfall is emphasized. We then go on how this problem may be solved and stochastic interpolation has been proved superior in such scenarios. The basic idea and the theoretical formulation of this method is presented in a manner suitable for instruction using the well-known discovery method. This formulation is then implemented in a software and thereby creating software-based teaching aid that can be used to demonstrate the stochastic interpolation technique. This system provides an interactive learning environment for students and also to explore the method at different scenarios including model validations.

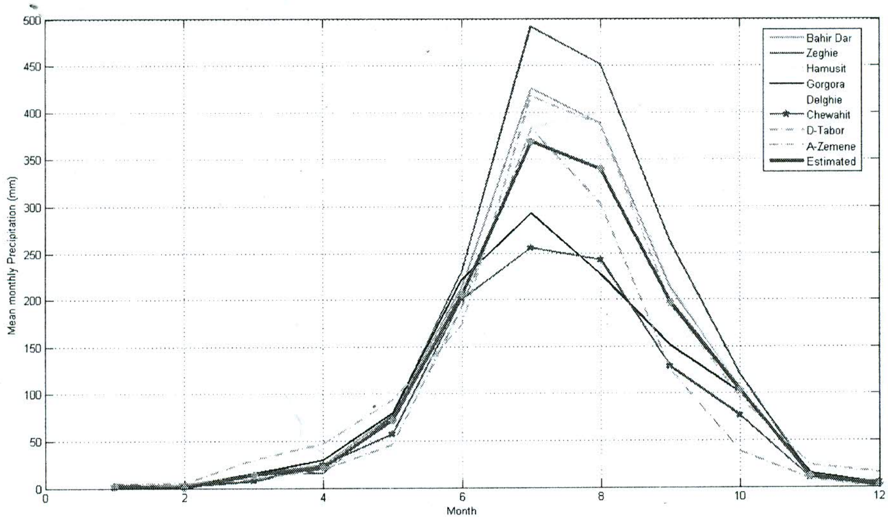

Figure 5.1: Comparison between measured mean monthly precipitation around Lake Tana catchment and estimated mean monthly precipitation on the lake from 1996 to 2006.

In addition to the pedagogical contributions, the present thesis has also scientific objective. The scientific goal is to estimate the spatial dependence of the amount of rainfall over Lake Tana using stochastic interpolation method. This is done by first assessing the rainfall characteristics inferred from rain gauge stations distributed around the lake catchment. The stochastic interpolation technique (Kriging) is then used to estimate the spatial distribution of precipitation over the lake. Model validity is also checked using rainfall-runoff correlation analysis. Figure 5.1 shows a comparison between estimated mean monthly precipitation over the lake and mean monthly precipitation around the lake catchment. The mean annual rainfall over Lake Tana is estimated to be 1248mm, which is about 7 percent lower than the rainfall on the surrounding catchments [11]. Our estimation shows that mean annual rainfall over Lake Tana for the period 1996 to 2006 was 1337mm. This result is in a good agreement with the estimate obtained [11] using IDW interpolation technique, which was 1335mm. On the other hand, there is a significant difference between previous published works that incorporates an annual mean of 1451mm lake precipitation to Lake Tana water balance calculations by taking simply the Bahir Dar station and the balance shows an error of +22mm per year [3]. Strictly speaking rainfall is the major aspects of water balance calculation and the balance error in the previous published work was due to over estimation of precipitation on the lake. Hence, we can conclude that rainfall at ungauged sites can be estimated as a linear combination of measured values from the adjacent stations using the appropriate weights. Most importantly of all linear interpolation techniques, the stochastic (Kriging) method is the best to estimate aerial precipitation.

5.1 Recommendations

As a continuation of the present work we recommend that:

- other researcher can conduct research that can measure the effectiveness of the present system pedagogically in teaching stochastic interpolation technique.
- other researcher can increase the optimality of the stochastic interpolation estimation techniques by taking couple of station gauges and by incorporating altitude in addition to longitude and latitude, because precipitation tends to increase as elevation increases.
- hydrologist may use the present result which leads to a more optimal water balance calculation and other hydrological modeling.

Bibliography

[1] www.sgo.fi/ baylie

[2] Kassa, T. Impacts of computer simulation on students' achievment, group interaction and attitude in direct current circuits: the case of 2007 first year physics majoring students at Bahir Dar university, Master thesis, Bahir Dar university, Bahir Dar, (2007).

[3] Kebede, S., Y.Travi, The Water Balance of Lake Tana and its Sensitivity to fluctuations in rainfall, Blue Nile basin, Ethiopia. Journal of Hydrology, **316**, 233-247 (2006).

[4] Mohor, P.A. The geology of Ethiopia. University-College Press, Addis Ababa, (1962).

[5] Chorwiz, J., The Tana basin, Ethiopia. Intra-plateau uplift, rifting and subsidence. Tectonophysics **295**, 351-367 (1998).

[6] Negelkerke, L. In Lake Tana, A unique fish fauna needs protection, In Bioscience, (1997).

[7] Conway, D. The Climate and hydrology of the upper Blue Nile river. The Geographical Journal, **166**, 49-62 (2000).

[8] Triona, L.M., Klahr, D., Computer simulation and video for hands- on equipment in an elementary school class, Cogn.Instruct, (2003).

[9] Nigussie, M., Three-Dimensional virtual teaching aid for demonstrating the physics of the earth's magnetic field. Master thesis, Bahir Dar university, Bahir Dar [Unpublished], (2008).

[10] Sahasrabudh, S.R., Irrigation engineering and hydrolic structure, 6^{th}ed. G.N. engineering college: Ludhia, (1996).

[11] SMEC, I.P., Hydrological Study of the Tana-Beles Sub-Basins. Part 1, (2007).

[12] Lanza, L.G., Ramire, J.A, Todini, E., Stochastic rainfall interpolation and downscaling. Hydrology and Earth system sciences,**5(2)**, 139-143 (2001).

[13] Te Chow, V., D.R. Maidment, and L.W. Mays. Hidrologa aplicada. Editorial Nomos S.A. Bogot. Colombia. (2000).

[14] Tabios and Salas, J.D., A comparative analysis of techniques for spatial interpolation of precipitation, Water Resources Bull. **21**, 365-380 (1995).

[15] Phillips, D, J. Dolph and D. Marks., A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain, Agri. and Forest Meteorol,**58**, 119-141 (1992).

[16] Martin, H.T., MATLAB Recipes for Earth Sciences, 2^{nd}ed. Springer-Verlag Berlin Heidelberg, (2007).

[17] Gandin, L.S., The Planning of Meteorological Station Networks. (Technical Note No. 111). WMO No. **265**, Geneva, 1970.

[18] Matheron, G., The theory of regionalised variables and its applications, Cahir du Center de Morphologie Mathematique, No.**5**, Fontainebleau, France, (1971).

Appendix 1.Double mass analysis of rainfall stations

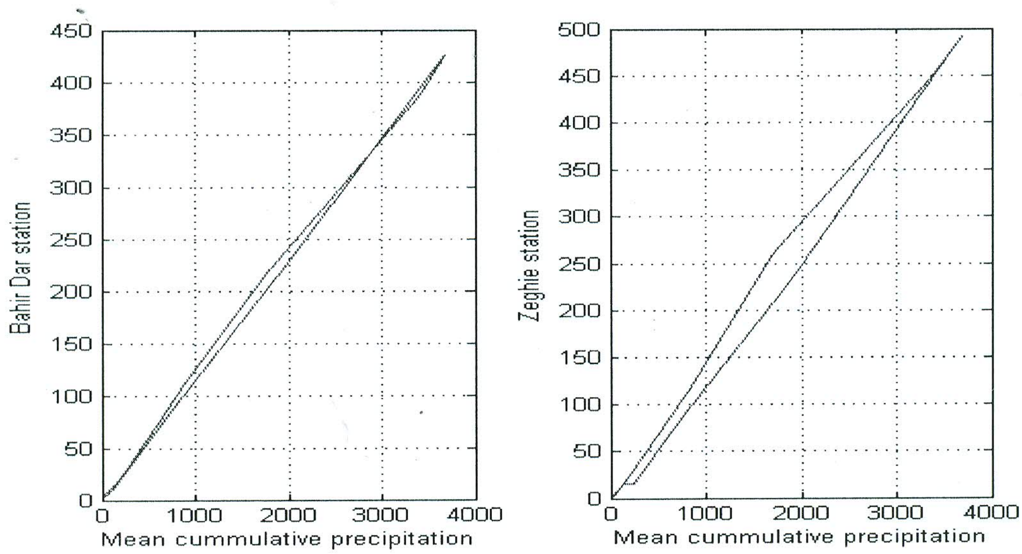

Figure 5.2: Double mass curve between mean monthly precipitation on a station and mean monthly cumulative precipitation of all stations

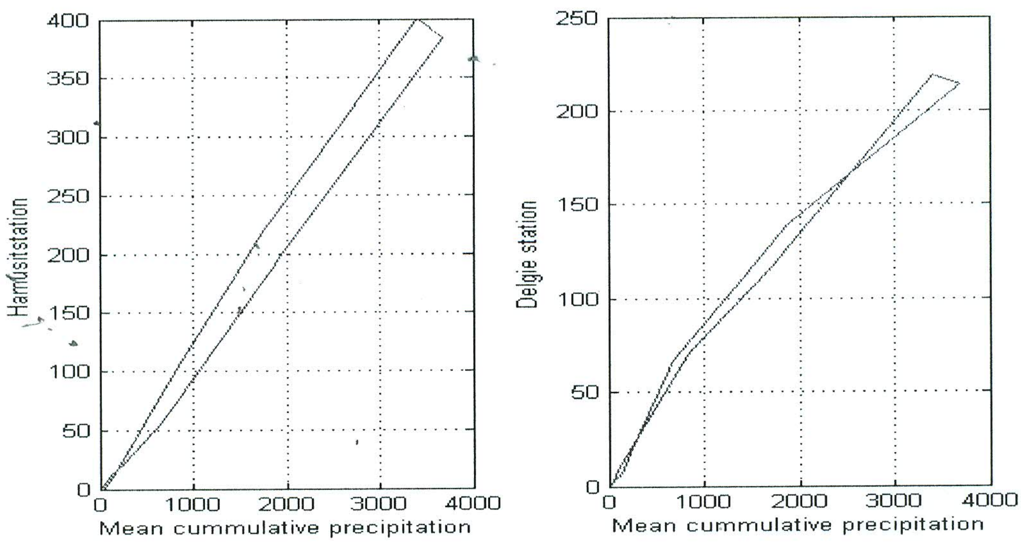

Figure 5.3: Double mass curve between mean monthly precipitation on a station and mean monthly cumulative precipitation of all stations

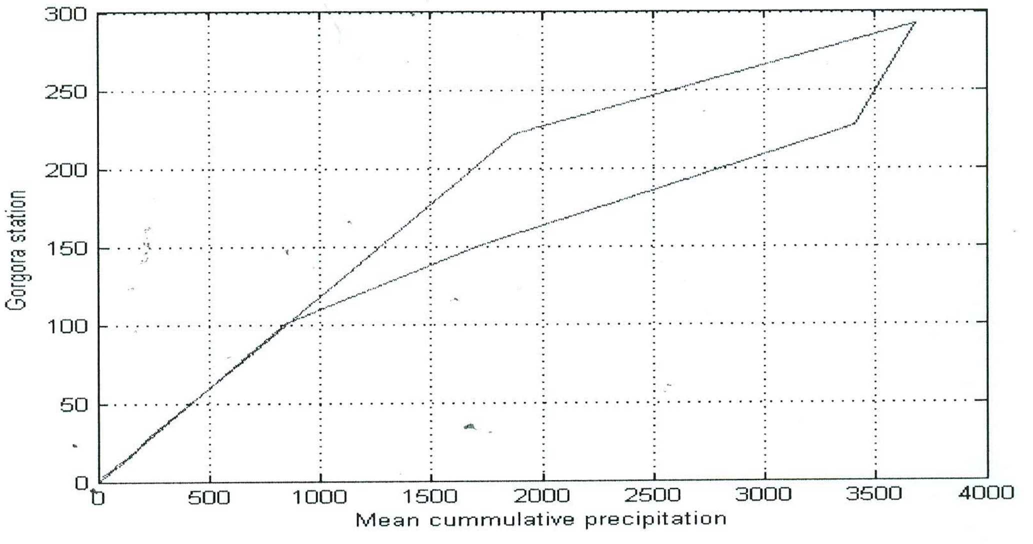

Figure 5.4: Double mass curve between mean monthly precipitation on a station and mean monthly cumulative precipitation of all stations

Appendix 2.A software-based teaching aid that demonstrates the stochastic interpolation techniques (Kriging method), which estimates precipitation at the prediction (unrecorded) locations over Lake Tana for April (1996-2006).

```
% This is a program that implements a Stochastic Kriging
%interpolation techniques,which estimates precipitation
%at the prediction (unrecorded) locations. The in put data
%used in this program are the mean monthly ground rain fall data
%taken from at the Ethiopian Meteorological Agency Bahir Dar
```

```
    %branch from 1996 to 2006 for each sample stations.
    % Copyright@ Yihenew Getinet Bahir Dar University
clear all
clc
%%Group 1
RR1=xlsread('bahidar.xls');%% Call the Bahir Dar rainfall data from
%%1996 to 2006 for each month from the directory
RR2=xlsread('Zeghie.xls');
RR3=xlsread('B_Airport.xls');
RR4=xlsread('Meshentie.xls');
%%-----------------------------
e=-999;         %% error
[rows cols]=size(RR2);
for j=1:cols
    for i=1:rows
        if RR2(i,j)==e
            RR2(i,j)=(RR1(i,j)+RR3(i,j)+RR4(i,j))/3;
        end
    end
end
%%-----------------------------
e=-999;
[rows cols]=size(RR3);
for j=1:cols
```

```
    for i=1:rows
        if RR3(i,j)==e
            RR3(i,j)=(RR1(i,j)+RR2(i,j)+RR4(i,j))/3;
        end
    end
end
%%-----------------------------
e=-999;
[rows cols]=size(RR4);
for j=1:cols
    for i=1:rows
        if RR4(i,j)==e
            RR4(i,j)=(RR1(i,j)+RR3(i,j)+RR2(i,j))/3;
        end
    end
end
% %%---------------------------
%%%-------------------------------------
%% Group2
RR5=xlsread('Chandiba.xls');
RR6=xlsread('Chewahit.xls');
RR7=xlsread('Delgie.xls');
RR8=xlsread('Gorgora.xls');
RR9=xlsread('Maksegnit.xls');
```

```
%%------------------------------
e=-999;
[rows cols]=size(RR6);
for j=1:cols
    for i=1:rows
        if RR6(i,j)==e
            RR6(i,j)=(RR7(i,j)+RR9(i,j))/2;
        end
    end
end
%%------------------------------
e=-999;
[rows cols]=size(RR5);
for j=1:cols
    for i=1:rows
        if RR5(i,j)==e
            RR5(i,j)=(RR6(i,j)+RR7(i,j))/2;
        end
    end
end
%%------------------------------
e=-999;
[rows cols]=size(RR7);
for j=1:cols
```

```
    for i=1:rows
        if RR7(i,j)==e
            RR7(i,j)=(RR6(i,j)+RR5(i,j))/2;
        end
    end
end
%________________________________________
e=-999;
[rows cols]=size(RR8);
for j=1:cols
    for i=1:rows
        if RR8(i,j)==e
            RR8(i,j)=(RR5(i,j)+RR6(i,j)+RR7(i,j))/3;
        end
    end
end
%%------------------------------
e=-999;
[rows cols]=size(RR9);
for j=1:cols
    for i=1:rows
        if RR9(i,j)==e
            RR9(i,j)=(RR6(i,j)+RR7(i,j)+RR8(i,j))/3;
        end
```

```
    end
end
% %%----------------------------
%%%Group3
RR10=xlsread('Yifag.xls');
RR11=xlsread('Enfranz.xls');
%%------------------------------
e=-999;
[rows cols]=size(RR11);
for j=1:cols
    for i=1:rows
        if RR11(i,j)==e
            RR11(i,j)=(RR9(i,j)+RR10(i,j))/2;
        end
    end
end
%%______________________
%%Group 4
RR12=xlsread('D_tabor.xls');
RR13=xlsread('Hamusit.xls');
RR14=xlsread('Wanzaye.xls');
RR15=xlsread('A_Zemen.xls');
%%------------------------------
e=-999;
```

```
[rows cols]=size(RR13);
for j=1:cols
    for i=1:rows
        if RR13(i,j)==e
            RR13(i,j)=(RR1(i,j)+RR2(i,j)+RR14(i,j))/3;
        end
    end
end
%%------------------------------
e=-999;
[rows cols]=size(RR14);
for j=1:cols
    for i=1:rows
        if RR14(i,j)==e
            RR14(i,j)=(RR13(i,j)+RR12(i,j))/2;
        end
    end
end
%%------------------------------
e=-999;
[rows cols]=size(RR15);
for j=1:cols
    for i=1:rows
        if RR15(i,j)==e
```

```
            RR15(i,j)=(RR10(i,j)+RR11(i,j))/2;
        end
    end
end
%%------------------------------
%% Mean monthly RR
RR1=mean(RR1(:,1:cols));
RR2=mean(RR2(:,1:cols));
RR3=mean(RR3(:,1:cols));
RR4=mean(RR4(:,1:cols));
%%------------------------------
RR5=mean(RR5(:,1:cols));
RR6=mean(RR6(:,1:cols));
RR7=mean(RR7(:,1:cols));
RR8=mean(RR8(:,1:cols));
RR9=mean(RR9(:,1:cols));
RR10=mean(RR10(:,1:cols));
RR11=mean(RR11(:,1:cols));
RR12=mean(RR12(:,1:cols));
RR13=mean(RR13(:,1:cols));
RR14=mean(RR14(:,1:cols));
RR15=mean(RR15(:,1:cols));
%%______________________________
% %%%Calculating the empirical semivariogram
```

```
            RR15(i,j)=(RR10(i,j)+RR11(i,j))/2;
        end
    end
end
%%-----------------------------
%% Mean monthly RR
RR1=mean(RR1(:,1:cols));
RR2=mean(RR2(:,1:cols));
RR3=mean(RR3(:,1:cols));
RR4=mean(RR4(:,1:cols));
%%-----------------------------
RR5=mean(RR5(:,1:cols));
RR6=mean(RR6(:,1:cols));
RR7=mean(RR7(:,1:cols));
RR8=mean(RR8(:,1:cols));
RR9=mean(RR9(:,1:cols));
RR10=mean(RR10(:,1:cols));
RR11=mean(RR11(:,1:cols));
RR12=mean(RR12(:,1:cols));
RR13=mean(RR13(:,1:cols));
RR14=mean(RR14(:,1:cols));
RR15=mean(RR15(:,1:cols));
%%-----------------------------
% %%%Calculating the empirical semivariogram
```

```
%%instead of using for loops to run faster.

%%The plot of the experimental variogram is called the variogram

%%cloud. We get this after extracting the

%%lower triangular portions of the D and G arrays.

indx = 1:length(z);

[C,R] = meshgrid(indx);

I = R > C;

plot(D(I),G(I),'.' )

hold on

grid

xlabel('lag distance')

ylabel('variogram')


%To calculate the mean minimum distance of pairs

%% we have to replace the diagonal of

%the lag matrix D zeros with NaNs,

%otherwise the minimum distance will be zero:

 D2 = D.*(diag(x*NaN)+1);

% %lag = mean(min(D2));

lag=12.7922;

%As a rule of thumb, the half maximum distance is

%suitable range for variogram analysis. We obtain the half maximum distance

%and the maximum number of lags by:

hmd = max(D(:));
```

```
max_lags = floor(hmd/lag);
LAGS = ceil(D/lag);
for i = 1 : max_lags
SEL = (LAGS == i);
DE(i) = mean(mean(D(SEL)));
PN(i) = sum(sum(SEL == 1))/2;
GE(i) = mean(mean(G(SEL)));
end
%where SEL is the selection matrix defined by the lag classes in LAG, DE is
%the mean lag, PN is the number of pairs and GE is the variogram estimator.
%Now we can plot the classical variogram estimator (variogram versus mean
%separation distance):
DE1=[D2(1,2) D2(4,6)];
DE2=[D2(1,3) D2(2,3) D2(4,5) D2(5,6) D2(7,8)];
DE3=[D2(3,6) D2(3,7) D2(3,8) D2(6,7)];
DE4=[D2(1,5:8) D2(2,4:8) D2(3,4:5) D2(4,7)];
DE5=[D2(1,4) D2(4,8) D2(5,7:8) D2(6,8)];
DE=[mean(DE1) mean(DE2) mean(DE3) mean(DE4) mean(DE5)];
GE1=[G(1,2) G(4,6)];
GE2=[G(1,3) G(2,3) G(4,5) G(5,6) G(7,8)];
GE3=[G(3,6) G(3,7) G(3,8) G(6,7)];
GE4=[G(1,5:8) G(2,4:8) G(3,4:5) G(4,7)];
GE5=[G(1,4) G(4,8) G(5,7:8) G(6,8)];
GE=[mean(GE1) mean(GE2) mean(GE3) mean(GE4) mean(GE5)];
```

```
m=4;                          % m stands for the selected month
z=[RR1(1,m) RR2(1,m) RR13(1,m) RR6(1,m) RR7(1,m) RR8(1,m) RR15(1,m) RR12(1,m)];
% Z contains the selected mean monthly rain fall for the sample points
x=[324.344 315.194 342.642 307.352 287.747 312.580 367.560 392.309];
% x-coordinate of the sample points (stations) in meters
y=[1284.306 1294.762 1305.218 1366.649 1348.351 1356.193 1340.856 1311.754];
% y-coordinate of the sample points (stations) in meters
deltax=max(x)-min(x);
deltay=max(y)-min(y);
x0=min(x):0.01*deltax:max(x);
%x0 contains x-coordinate of the prediction locations in meters
y0=min(y):0.01*deltay:max(y);
%Y0 contains y -coordinate of the prediction location in meters
%%To calculate the experimental variogram
%%we first have to build pairs of observations
[X1,X2] = meshgrid(x);
[Y1,Y2] = meshgrid(y);
[Z1,Z2] = meshgrid(z);
%%The matrix of separation distances D between the observation points is
D = sqrt((X1 - X2).^2 + (Y1 - Y2).^2);
%%The experimental variogram G as half the squared differences between the
%%observed values is
G = 0.5*(Z1 - Z2).^2;
%%We used the MATLAB capability to vectorize commands
```

```
plot(DE,GE,'o','MarkerFaceColor',[.6 .6 .6] )
var_z = var(z);
b = [0 max(DE)];
c = [var_z var_z];
%hold on
plot(b,c,'--r')
yl = 1.1 * max(GE);
ylim([0 yl])
xlabel('Averaged distance between observations')
ylabel('Averaged semivariance')
%Spherical model with nugget
nugget =0;
sill =122.8;
range =52;
lags = 0:max(DE);
Gsph = nugget + (sill*(1.5*lags/range - 0.5*(lags/...
range).^3).*(lags<=range) + sill*(lags>range));
plot(lags,Gsph,'g','Linewidth',2)
% Exponential model with nugget
nugget =0;
sill =122.8;
range =52;
Gexp = nugget + sill*(1 - exp(-3*lags/range));
plot(lags,Gexp,'b','Linewidth',2)
```

```
% % % Linear model with nugget
% % nugget = 10;
% % slope = 100;
% % Glin = nugget + slope*lags;
% % plot(lags,Glin,'-m')
xlabel('Distance between observations')
ylabel('Semivariance')
legend('Scatter plot','Variogram estimator','Precipitation variance',...
'Sperical model','Exponential model')
hold off
%%The variogram model is a parametric curve fitted to the variogram estimator.
%%Variogram model,the exponential variogram model with a nugget, sill and range
G_mod = (nugget + sill*(1 - exp(-3*D/range))).*(D>0);
%%Then we get the number of observations and add a column and row vector of
%%all ones to the G_mod matrix and a zero at the lower left corner:
n = length(x);
G_mod(:,n+1) = 1;
G_mod(n+1,:) = 1;
G_mod(n+1,n+1) = 0;
%%Now the G_mod matrix has to be inverted:
G_inv = inv(G_mod);
deltay=max(y)-min(y);
Ry=min(y):0.01*deltay:max(y);
deltax=max(x)-min(x);
```

```
Rx=min(x):0.01*deltax:max(x);

[Xg1,Xg2] = meshgrid(Rx,Rx);

[Yg1,Yg2] = meshgrid(Ry,Ry);

%%and converted to vectors by:

Xg = reshape(Xg1,[],1);

Yg = reshape(Yg2,[],1);

%%Then we allocate memory for the kriging estimates Zg and the kriging variance

%%s2_k by:

Zg = Xg * NaN;

s2_k = Xg * NaN;

%DOR = Xg * NaN;

%%Now we are kriging the unknown at each grid point:

 for k = 1 : length(Xg)

 DOR = ((x - Xg(k)).^2 + (y - Yg(k)).^2).^0.5;

 G_R = (nugget + sill*(1 - exp(-3*DOR/range))).*(DOR>0);

 %G_R = (nugget + (sill*(1.5*DOR/range - 0.5*(DOR/range).^3))).*(DOR>0);

 G_R(n+1) = 1;

 G_R=G_R';

 E = G_inv * G_R;

Zg(k) = sum(z*E(1:n,1));

s2_k(k) = sum((E(1:n,1))'*G_R(1:n,1))+E(n+1,1);

 end

r = length(Rx);

Z = reshape(Zg,r,r);
```

```
SK = reshape(s2_k,r,r);
%%A subplot on the left presents the kriged values:
%subplot(1,2,1)
h = pcolor(Xg1,Yg2,Z);
set(h,'LineStyle','none')
axis equal
xlim([min(x) max(x)])
ylim([min(y) max(y)])
title('Kriging estimate Of areal precipitation for April')
xlabel('x-Coordinates (km)')
ylabel('y-Coordinates (km)')
colorbar
 hold on
 plot(x,y,'ok')
```

Appendix 3.Kriging Estimate of Aerial Precipitation for each month over Lake Tana (1996-2006)

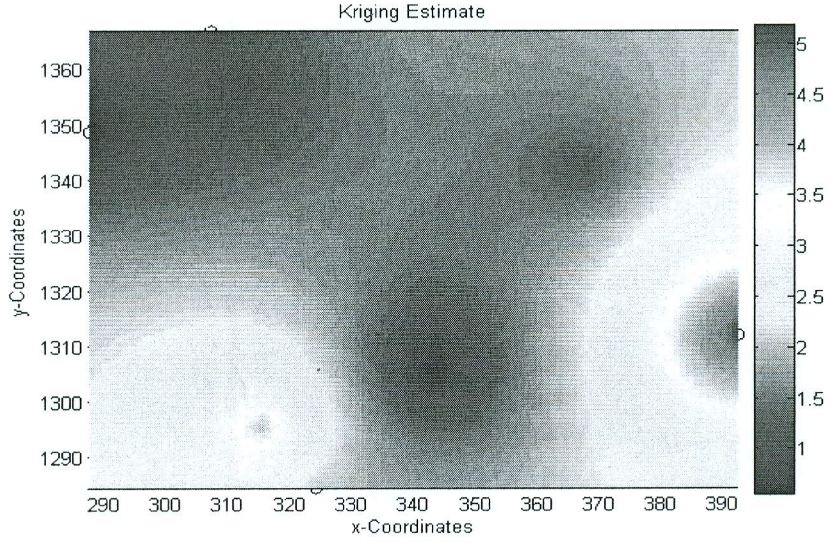

Figure 5.5: Kriging estimate of aerial precipitation over Lake Tana for January.

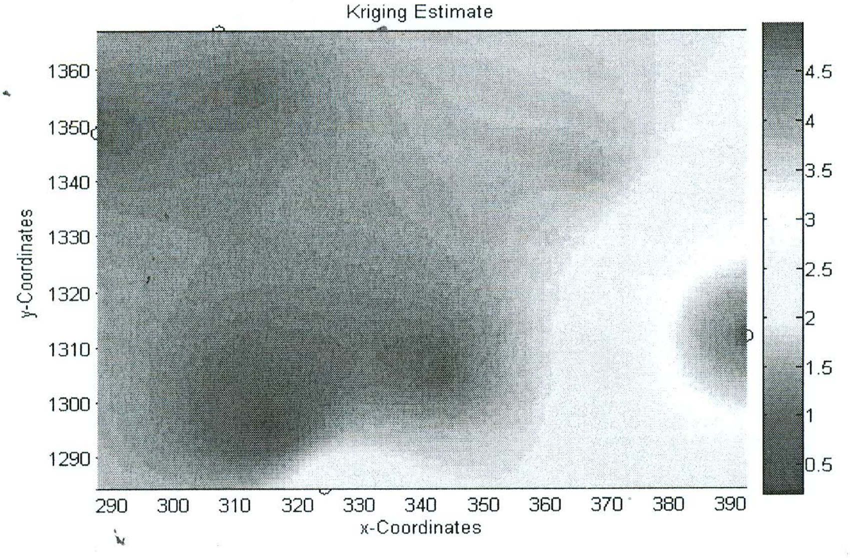

Figure 5.6: Kriging estimate of aerial precipitation over Lake Tana for February.

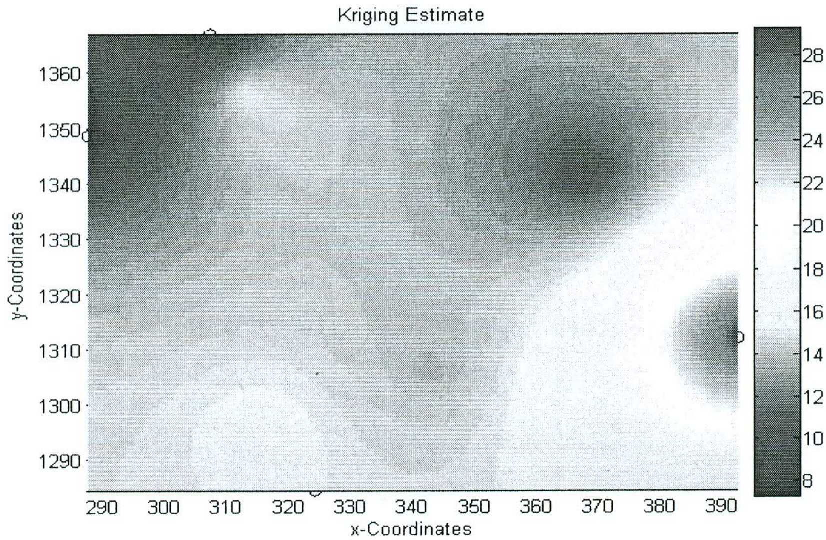

Figure 5.7: Kriging estimate of aerial precipitation over Lake Tana for March.

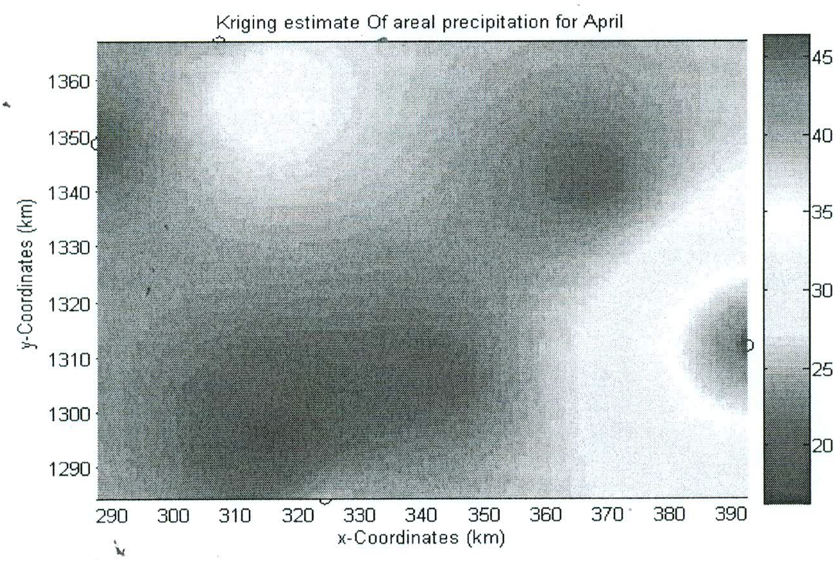

Figure 5.8: Kriging estimate of aerial precipitation over Lake Tana for April.

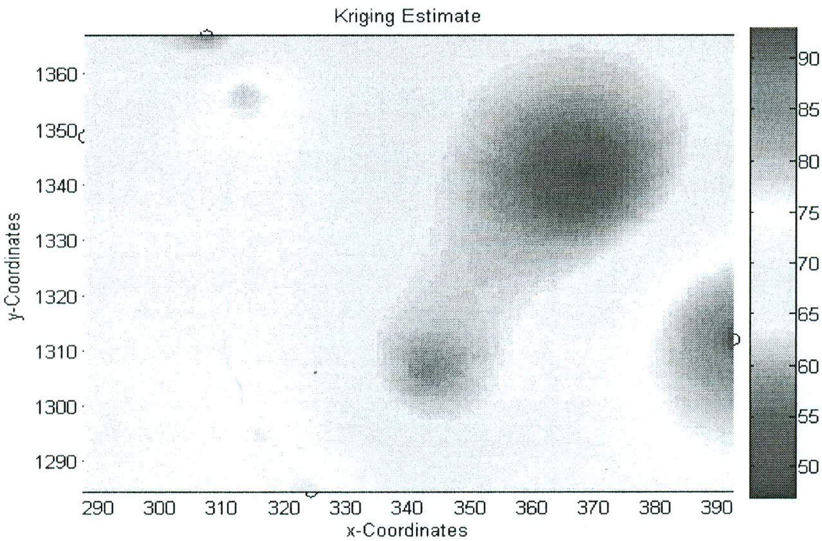

Figure 5.9: Kriging estimate of aerial precipitation over Lake Tana for May.

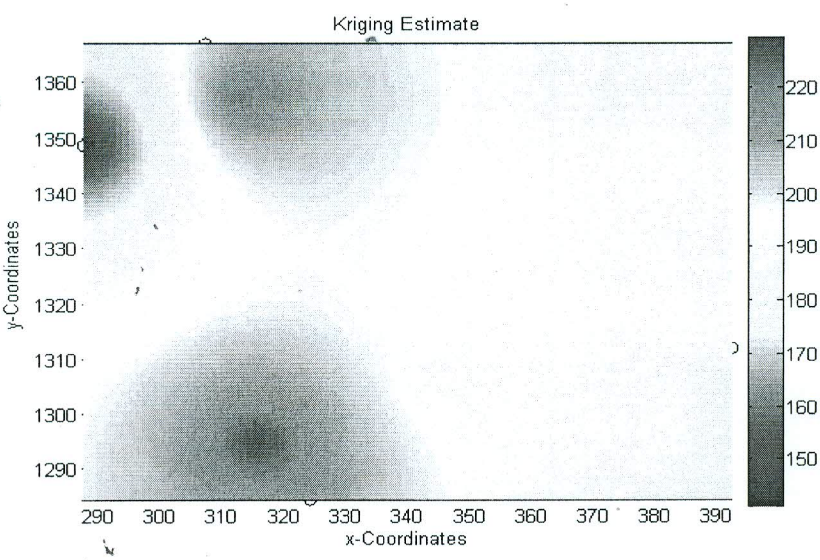

Figure 5.10: Kriging estimate of aerial precipitation over Lake Tana for Jun.

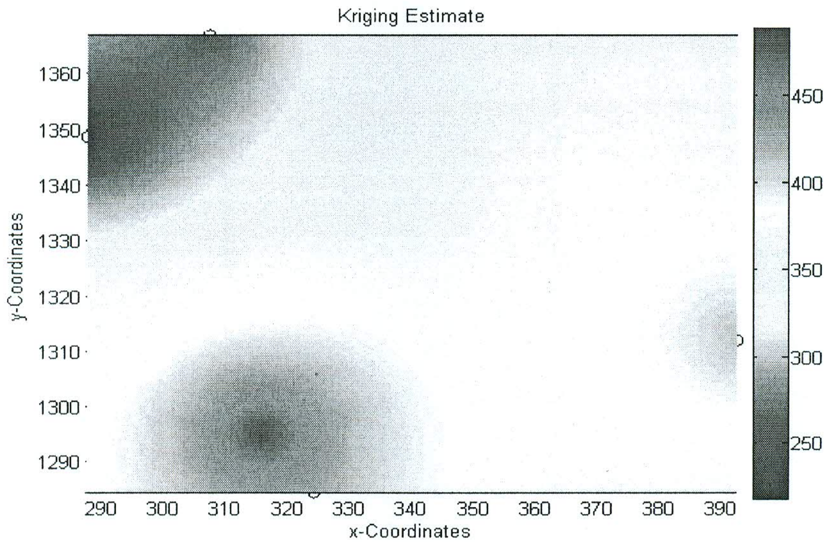

Figure 5.11: Kriging estimate of aerial precipitation over Lake Tana for July.

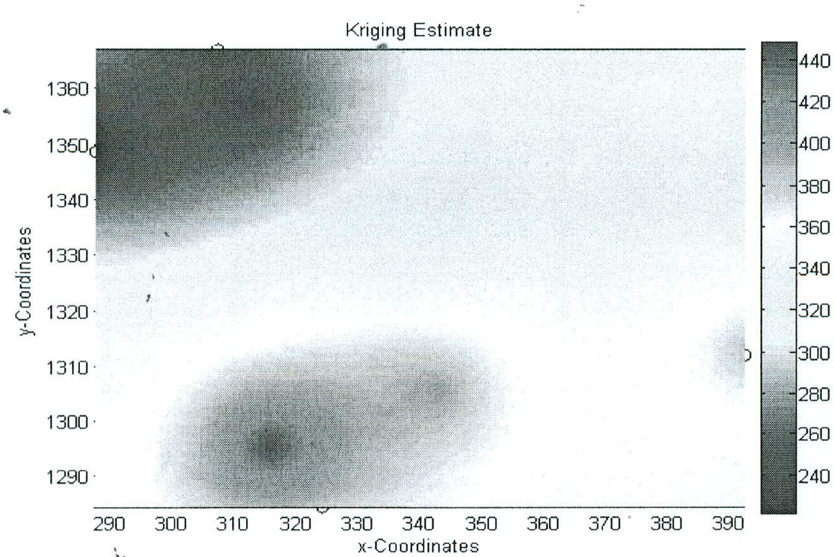

Figure 5.12: Kriging estimate of aerial precipitation over Lake Tana for August.

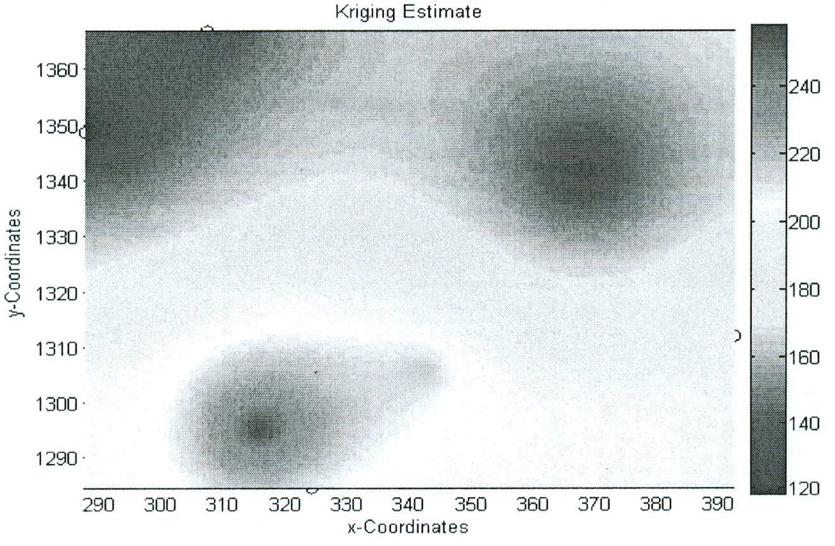

Figure 5.13: Kriging estimate of aerial precipitation over Lake Tana for September.

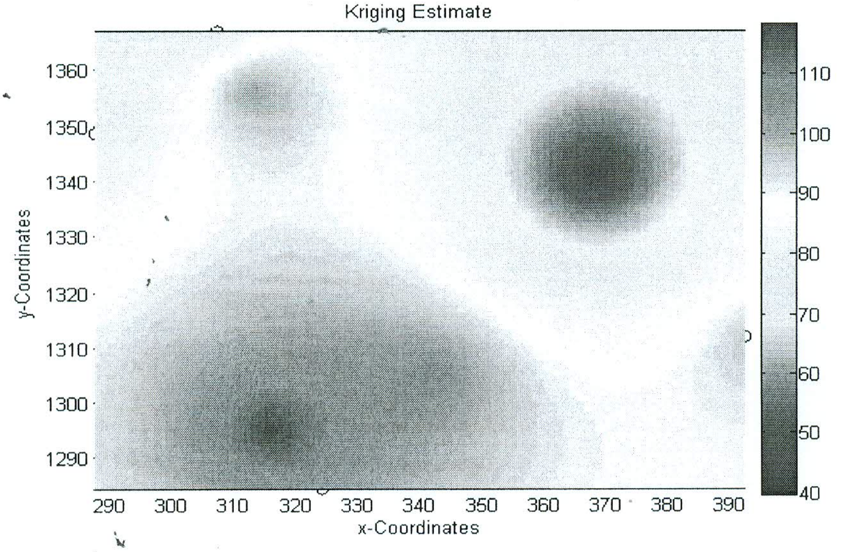

Figure 5.14: Kriging estimate of aerial precipitation over Lake Tana for October.

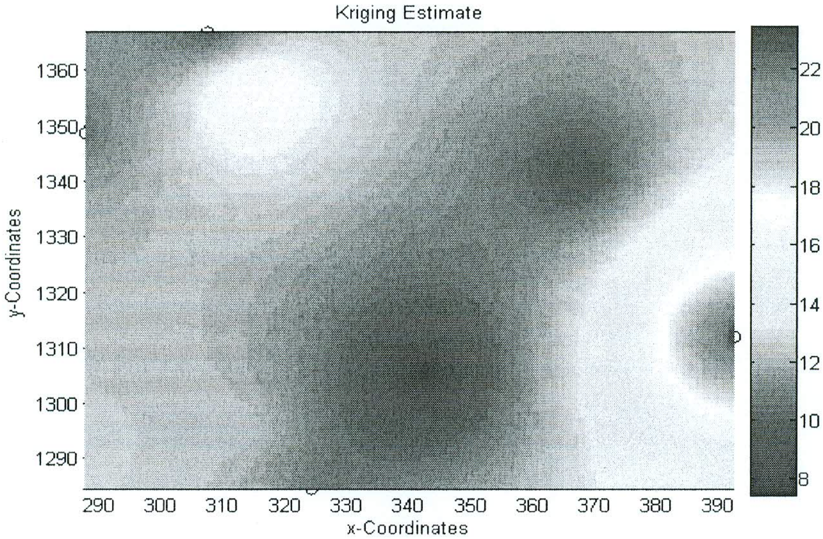

Figure 5.15: Kriging estimate of aerial precipitation over Lake Tana for November.

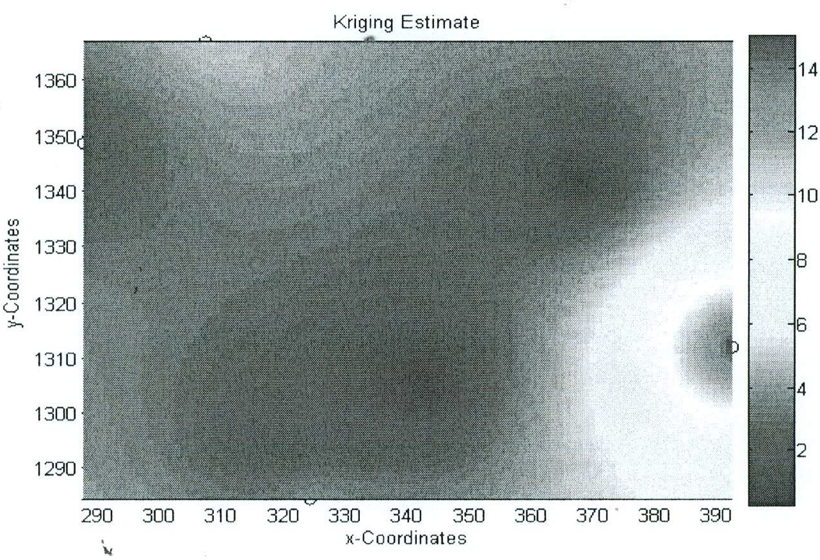

Figure 5.16: Kriging estimate of aerial precipitation over Lake Tana for December.

DECLARATION

I declare that the thesis is my original work and has not been presented for a degree in any other university, and that all sources of materials have been duly acknowledged.

Name: Yihenew Getinet

Signature:

This thesis has been submitted for the examination with my approval as university advisor.

Name: Dr. Baylie Damtie

Signature:

Date: