ESTIMATION OF LONG-TERM INFLOW INTO LAKE NAIVASHA FROM THE MALEWA CATCHMENT, KENYA

By

Md. Azizul Haque Podder

APRIL, 1998

ESTIMATION OF LONG-TERM INFLOW INTO LAKE NAIVASHA FROM THE MALEWA CATCHMENT, KENYA

By

Md. Azizul Haque Podder

This thesis is submitted in partial fulfillment for the degree of Master of Science in Water Resources Survey with emphasis on Watershed Management in the International Institute for Aerospace Survey and Earth Sciences (I.T.C.), Enschede, The Netherlands.

Supervisors:

Ir. H. de Brouwer
Drs. N.H.W. Donker

Submitted to the board of Examiners:

Prof. Dr. A.M.J.Meijerink (Chairman of board of examiners)
Ir. C. de Wijs (External Examiner)
Ir. H. de Brouwer (Supervisor)
Drs. R. Becht

Water Resources Survey Division
Department of Earth Resources Survey
International Institute for Aerospace Survey and Earth Sciences (ITC)
Enschede, The Netherlands
April 23, 1998

ABSTRACT

The study is about the estimation of long-term inflow into lake Naivasha from the Malewa catchment. The Malewa catchment is a sub catchment of the Naivasha basin, situated in the Rift Valley Province, Kenya. It has an area of about 1428 km^2. The Malewa river is the main source of fresh water supply into lake Naivasha which has the significant influence in agriculture, power generation, fishery, wildlife and tourism.

The study area suffers from inadequate data. The discharge series is not complete due to data gaps present at all the gauging stations. For the estimation inflow of the Malewa river, the data gaps were filled using linear interpolation and a multiple linear regression technique. To avoid the biasness, the gauging stations were selected from both sub catchments of the Malewa river.

The rainfall stations were far away from the study area and there was no correlation between rainfall and discharge. The rainfall data does not represent the areal precipitation of the catchment so that the rainfall data were not used for the prediction of discharge.

The rating curves were reconstructed and it was found that the reconstructed rating curves significantly differ from the supplied rating curves. The rating curve coefficients were optimized by assigning weights and without assigning weights. It was found that the developed optimization procedure with assigning weights produced reliable rating coefficients for low flow conditions and without assigning weights produced reliable rating coefficients for high flow conditions.

The random uncertainty of rating curves for main three gauging stations were calculated and it was found that the uncertainty of 2GB1 station was minimum (25.7% at the 95% confidence level) of the three gauging stations. The uncertainty of the lower part of the rating curves were higher compare to upper part of the rating curves.

The variation of yearly uncertainty of inflow of the Malewa river is not significant. The yearly uncertainty varies from 1.80% to 3.14% and the average uncertainty was 2.26%.

The variation of yearly inflow of Malewa river is significant. The maximum inflow (4428×86400 $m^3 \pm 3.14\%$) occurred during the year 1961 and the minimum inflow (613×86400 $m^3 \pm 1.89\%$) occurred during the year 1984. The average yearly inflow was 2486×86400 $m^3 \pm 2.26\%$ during the years 1960 - 1990.

ACKNOWLEDGMENTS

The completion of this thesis may not have been possible without the support, contributions and inspirations from the following persons:

Prof. A.M.J. Meijerink, Chairman of the Department;

Ir. J.A.M Hans de Brouwer, supervisor of this research, for his patience and indefatigable guidance;

Drs. N.H.W. Donker, supervisor of this research, for his guidance, support and encouragement;

Drs. R. Becht, for his close follow up, supplying necessary materials and guidance, especially during the field work; and

Ir. C. de Wijs, external examiner, who has spend his time on my examination

I would like to express my sincere gratitude to Ir. A.M. van Lieshout, who provided the necessary utilities during the research. My thanks to friends for their companionship and moral supports during the study. I appreciate all of the ITC staffs for their sincere cooperation during the study.

I would like to express my thanks to Allah for given me strength and wisdom to complete this research.

I am grateful to several people in Kenya around Lake Naivasha, i.e. the Lake Naivasha Owners Association and Mrs. S.Higgins for their cooperation, the owners of farms visisted for supplying vital information related to their farming system, and specially the working staff of Kenya Wildlife Service Centre of Naivasha for their excellent support and service provided during our field work. Also the project staff of the Kenyan Wetland project (Mrs. B. Van Helvoort, Mr. B. Zech et al) for their important support of this study.
Also a warm word of thank for the staff of the WRAP project, helping us wherever needed.

TABLE OF CONTENTS

Contents *Page*

Contents *Page*

Contents *Page*

LIST OF FIGURES

Figures **Page**

LIST OF PHOTOS:

LIST OF TABLES

APPENDICES:

CHAPTER ONE

INTRODUCTION

1.1 Introduction

Streamflow is generated by precipitation during storm events and by ground water entering surface channel. During dry periods, streamflows are sustained by ground water discharge. In areas where ground water reservoirs are below the stream channel streams cease to flow during protracted precipitation-free periods.

The nature of stream flow in a catchment is a function of the hydrologic input to that catchment (region) and the physical, vegetative and climatic characteristics. The natural basin features are very important element in runoff process, land-use features created by human (e.g., housing development, parking plot, road construction, deforestation, agriculture pattern) may in some cases, be the dominant one. Moreover, stream and rivers may exhibit not only even-based variation of flow but also seasonal variations which are largely a reflection of climate and in particular, of the balance between the precipitation and evaporation. In the Naivasha basin surrounding the Lake area evaporation is much higher than precipitation.

1.2 Importance of Study

Human cannot exist without water. It is true that water mismanaged, or during times of deficiency (drought), or times of surplus (flood) can be life threatening.

As water becomes more scarce and competition for its use expands, the need for improve water management. To provide water for the expanding population, new industrial development, food production, recreational demand, and for the preservation and protection of natural system, it becomes increasingly important for us to achieve a through understanding of the underlying hydrological process with which we must content.

This is the challenge to water resource engineers, hydrologists, planners, economists and other who must strive see that further allocations of water are sufficient to meet the needs for human and natural ecosystem.

The lake Naivasha is only the freshwater lake in the Rift Valley. This lake is highly significant for national freshwater resources. The irrigation crop farming is common near the lake Naivasha whereas other areas are rainfed agriculture. The socio-economic condition is significant with lake water such as agriculture, fisheries, tourism, and power generations. So, long term discharge series will help for the lake water management plan for sustainable development.

1.3 Objectives of the Study

The study is focused only on the water resources, mainly the estimation of long-term inflow into the lake Naivasha from the catchment, the main objectives of the study can be summarized as follows:-

i) to compare the main gauging stations and optimization of rating curve coefficients

iii) to infill the missing data gaps in discharge series.

iv) to asses the uncertainty of the inflow of the watershed.

1.4 Methodology

1.4.1 Rainfall analysis:

Correlation of daily and ten daily average rainfall between the pairs of stations. From the results it will be decided whether thiessen polygon is possible or not. Rainfall events will have to select for the prediction of runoff.

1.4.2 Rainfall-runoff analysis:

Analysis the relationship between the rainfall and discharge series. From the results it will be decided whether the rainfall is possible to use for infilling the missing data gaps.

1.4.3 Stream flow analysis:

Correlation of daily discharge between the pairs of stations. From the results it will be decided whether the discharge series of different stations are possible to use for infilling the missing data gaps.

1.4.4 Rating curve analysis

The stage-discharge relationship or rating curve at main gauging stations will be analyzed and compare with the existing rating curves to verify the existing rating curves.

1.4.5 Statistical analysis

1.4.5.1 Multiple linear regression

This technique was applied to test the combine effects of the different independent variables (gauging stations) on dependent variable. For "k" independent variables, X_1, X_2, ..., X_k the functional form of the multiple linear regression model is:

$$Y= \alpha+\beta_1X_1+\beta_2X_2 + \quad + \beta_kX_k \qquad (1.1)$$

where the βi's are the partial regression coefficients associated with each Xi and α is the interception of the line on the Y-axis (the value of Y when all the Xi's have zero values).

1.4.5.2 Standard error of estimate

The standard error of estimate is the standard deviation of the errors of the prediction and it provides an indication of their variability about the regression line in the population in which predictions are being made (Downier and Heath, 1984). Moreover, it is a good measure of the dispersion or scatter around the regression line of Y on X (Spiegel, 1961). It signifies the degree of variation expected to the encountered in making predictions through the use of regression equation developed with the use of regression analysis. In simple linear regression analysis, it can be computed through the following formulas:

$$Se = [\frac{\sum (Yi - Yest)^2}{N-2}]^{1/2}$$

Where Se = standard error of estimate
Yi = measured value
Yest = predicted or estimated value of Y
N= number of observation

1.5 Data and Materials used in the study

1.5.1 Rainfall, water level and discharge data

The rainfall, discharge and water level data series used in the study were received from the Water Development Ministry, Kenya.

1.5.2 MaterialS used in the study

The topographic maps at scale 1:50,000
The topographic maps at scale 1:200,000

1.6 Definition of Terms

Catchment is used synonymous with "watershed". The term "watershed" also connotes "drainage basin".

Inflow is the sum of the direct runoff and base flow. It also referred as "stream flow".

Rainfall is the form of precipitation observed in the rainfall station.

CHAPTER TWO

GENERAL ASPECTS OF THE STUDY AREA

2.1 Location of the Study Area

The Malewa catchment is situate in the Kenyan Rift Valley almost 70 km from the Nairobi (see figure 2.1). It is located between latitude 0° 09 to 0 ° 55 South and longitude 36 ° 09 to 36 ° 24 East. The maximum altitude is about 3990 meter and minimum altitude 1980 meter above the mean sea level. The area of the catchment is 1428 km^2.

The Naivasha basin is situated in the highest part of the Rift valley of about area of 3184 km^2. This basin has its internal drainage system and no outlet yet to be visible. The Malewa catchment is the sub-catchment within the upper part of the Naivasha catchment. The Malewa river drains into the Lake which is the main source of surface water. The main purposes of the Lake are:

* water supply for irrigation
* Water supply for generation of electricity
* Fish cultivation
* Drinking water & heritage of some wildlife
* Recreation
* Tourism

2.2 Climate:

Climatic conditions in the study area are quit diverse due to considerable differences in the altitude and land forms. Although the Lake is located within one degree of the equator and is thus "tropical", it generally experiences relatively cool conditions determined by altitude (Richardson and Richardson). The annual temperature range is approximately from 8 ° C to 30 ° C (Kenya government, 1976). The rainfall regime within the lake catchment is influenced by the rainshadow from the surrounding highlands of the Nyandarua rang (Aberdare) to the East, and the Mau Escarpment to the West. Two rainy seasons observed in this region. The "long rain" occurring in the March, April and May and the "short rain" in October and November. The rainfall pattern is controlled by relief, with much more rainfall in the higher altitudes than the lower altitude.

Naivasha experiences an average rainfall of 610 mm, and the wettest slopes of the Nyandarua mountains within the Lake's catchment receive as much as 1525 mm. The evaporation experienced by Naivasha is some 1,360 mm, so, the runoff from the nonimmediate catchment would seem to be broadly sufficient to maintain Lake level (East African Meteorological Department, 1963).

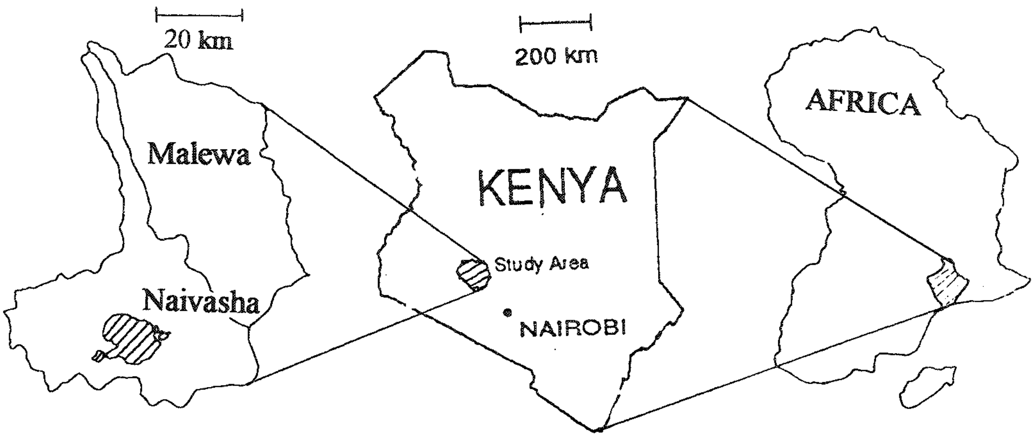

Figure 2.1: General location map of the study area

2.3 Topography

The Naivasha area of the Rift valley is confined by the Nyandarua Mountains (formerly the Aberdares mountains) to the East, (elevation to over 3990 meter) and the Mau Escarpment to the West (exceeding 3,000 meter). The Kinangop plateau forms a broad step between the Nyandarua range and the valley floor, east of Naivasha. The mount Longonot stands on the South of the Lake.

2.4 Rivers

In the Malewa catchment has a number of rivers and its tributaries. The main river is Malewa. The Turasha, Nandarasi, Engare Mugutyu, Wanjohi are the tributaries of this river. The Gigil, Karati are the other river of the Naivasha basin.

2.5 Land Use

The semi-arid climate and the topography are greatly influenced on the vegetation on this area. The catchment vegetation can broadly be grouped into:

* Forest
* Bushland
* Grassland
* Agricultural land

The natural forest within the study area comprises indigenous hardwood trees and grasses such as bamboo's. Menegai crater, the Eburru hills, Mau escarpment, Mount Longonot and the Nyandarua escarpment are all host of hardwood forests, whereas bamboo is confined to the Nyandarua and Mau escarpments. These form the main watersheds of the Lakes.

The greatest proportions of the low lying central part of the catchments are rangelands. Rangelands are land carrying natural or semi-natural vegetation that provide a habitat suitable for wild or domestic ungulates and varies from scrubland to grassland to bushland (Pratt and Gwynne, 1977). Naivasha shores are often encircled by ephemeral Papyrus colonies, and the surface covered by raft of Salvinia molesta (sometimes up to 25 % of the total surface area).

The main farming system in this area would commonly be referred to as mixed farming. Rainfed crop production is the most important activity within the catchments. Slopes of the Nyandarua, Mau mountains are the most common farming areas. Common crops include wheat, maize, potato, beans and sunflowers.

Irrigated crop farming is common near Lake Naivasha where large quantities of wheat, barley, French bean and fodder crops are grown. Another important horticulture products are flowers.

The livestock production by some dairy farms and by the Masai people is also common practice.

CHAPTER THREE

HYDROLOGICAL DATA ANALYSIS

3.1 Introduction

Engineering studies of water resources development and management depend heavily on hydrologic data. Hydrological modeling and any other farther analysis needs, basic examination to check whether the data to be used is consistent or not. There are some rainfall stations within the catchment but the data of these stations are not available. In this study the nearby Nakuru Met. station, Naivasha D.O, Milmet, Lake Nakuru National Park, Gilgil W\S and Elementaita N. Range Post rainfall stations' data were used for analysis(see figure-3.1). The discharge measurement station of Malewa catchment on river Malewa is 2GB1. The gauging stations in the tributary of Malewa river are 2GB3, 2GB4, 2GB5, 2GB7, 2GC4, 2GC5, 2GC7 and nearby gauging stations 2GA2, 2GA3, 2GA5 and 2GA6 of Gilgil river data were used for analysis.

3.2 Rainfall Data Analysis

3.2.1 Consistency and homogeneity of data

The rainfall data of the selected stations were checked by mass curve method.

3.2.1.1 Mass curve method

The accumulated rainfall data were plotted against time. Observing the mass curve, it seems to be there is a trend present in the Gilgil W\S station and Lake Nakuru National Park stations (figure -3.2d and 3.2e) and the data seems to be inconsistent. The rest of the station shows the rainfall data is consistent (see figure 3.2a, b, c, f). The period of data used in the analysis is too short. It may consider that there is a dry period and after that again wet period may come.

3.2.2 Relationship between average annual rainfall and altitude

In order to study the relationship of annual rainfall with changes of altitude, the average annual rainfall was plotted against their altitude. There is a good correlation coefficient (R^2=0.77) indicates a good relationship between average annual rainfall and elevation. It can be observed that the higher elevation shows much more average rainfall compare to lower elevation as shown in figure - 3.3

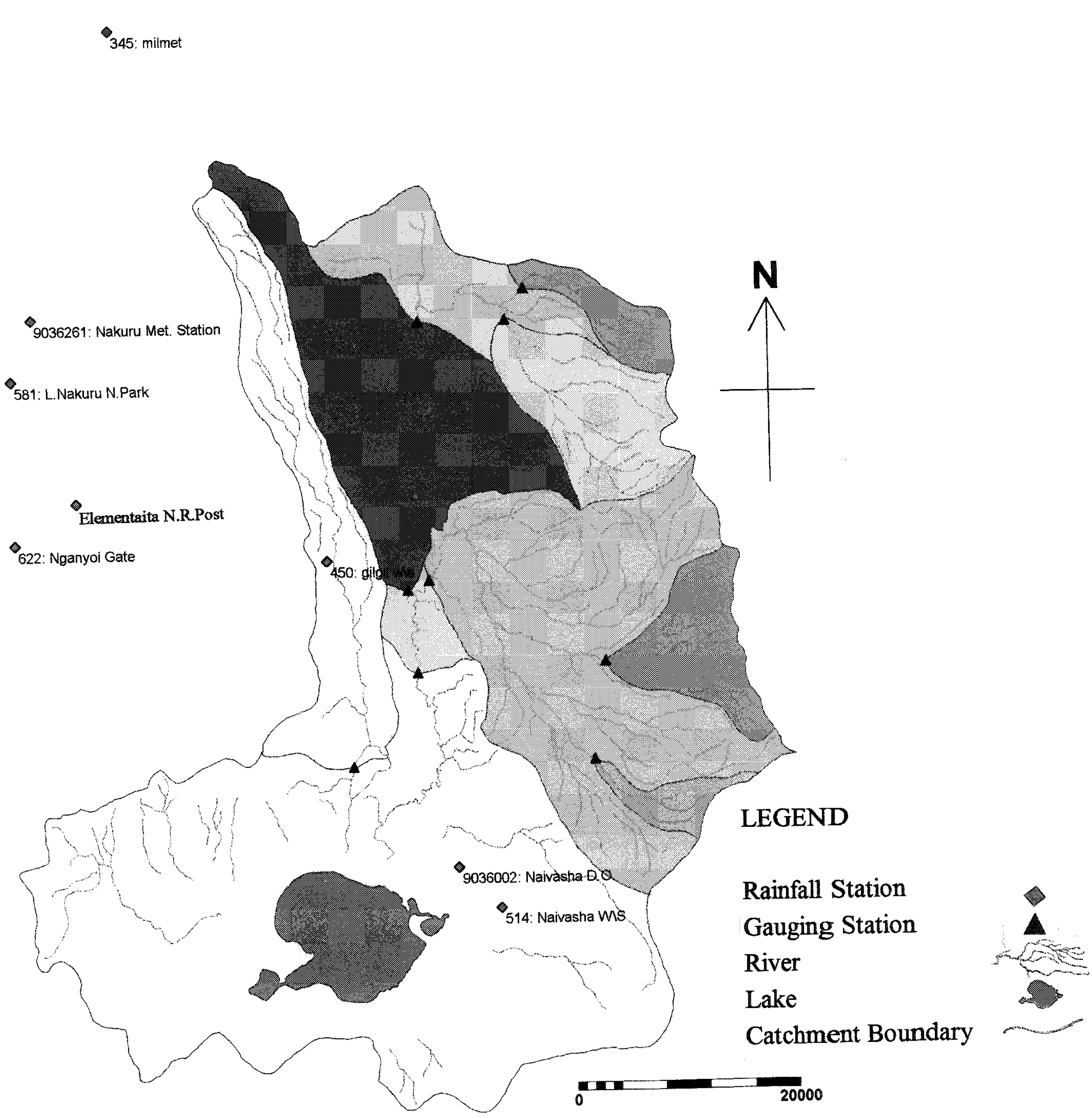

Figure - 3.1: Map showing rainfall stations and gauging stations

MASS CURVE OF RAINFALL

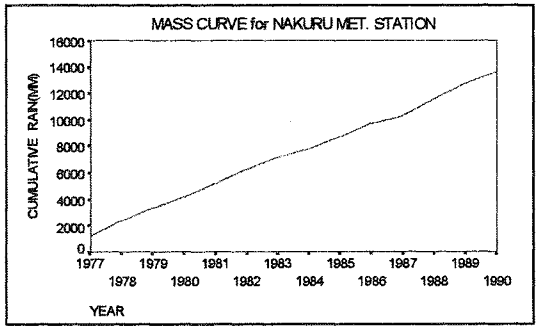

Figure 3.2 a

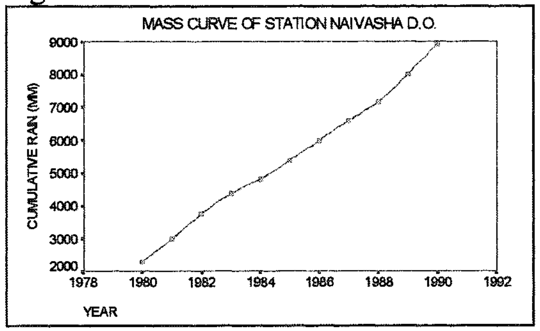

Figure 3.2.b

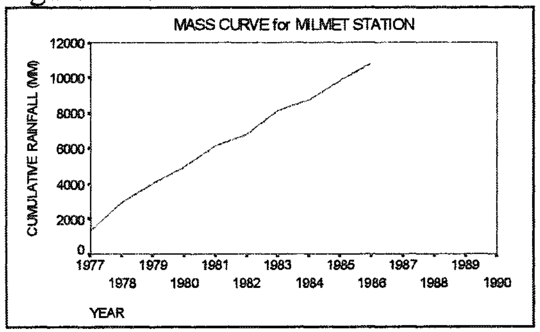

Figure 3.2 c

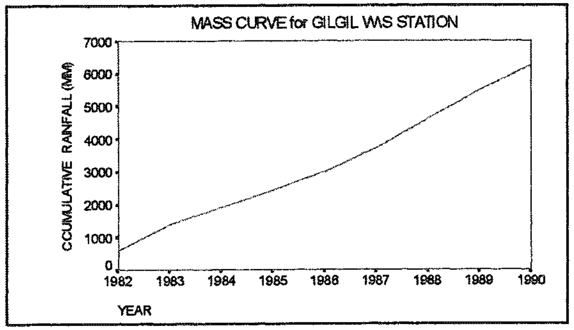

Figure 3.2 d

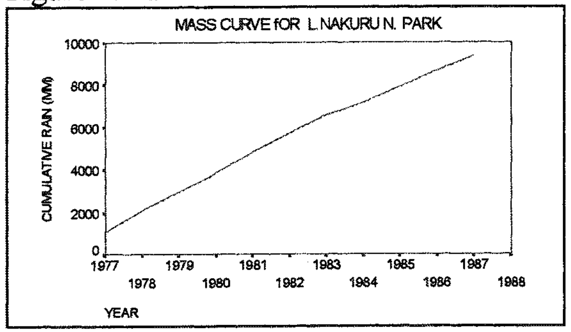

Figure 3.2 e

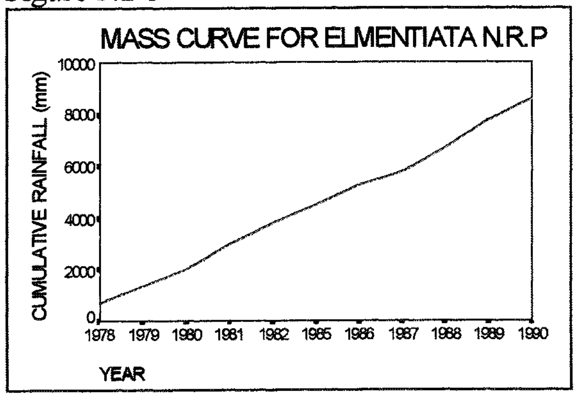

Figure 3.2 f

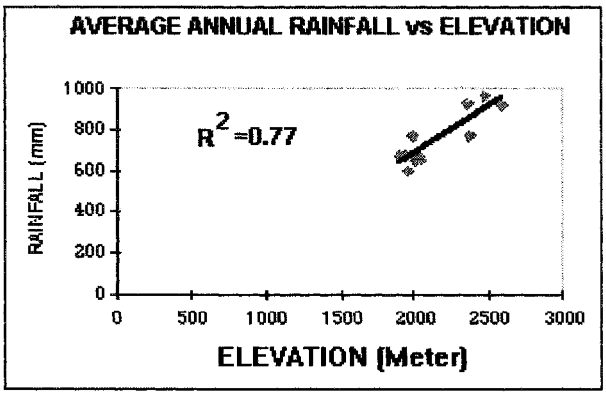

Figure - 3.3: Average annual rainfall vs. elevation

3.3 Spatial Distribution of Rainfall

The spatial distribution of rainfall over the catchment is not uniform. The rainfall frequency (number of rainy days) and the amount of rainfall varies significantly from each station. The rainfall data and the number of rainy days shown in Appendix-A (table - 3.1 and table - 3.2)

3.4 Correlogram

The correlation of rainfall have been examined as a function of distance between the stations. The rainfall data was selected as daily and ten daily average rainfall. The correlation coefficients and the distances between the stations are shown in tables 3.1A, 3.1B and 3.1C respectively.

The correlation coefficient was plotted against distance as shown in figure-3.4a and 3.4b. The figure shows the plotted points are scattered but there is a trend present. The increases of the distance decreases the correlation indicates that the spatial rainfall distribution not uniform. It also observed that the ten daily average rainfall shows the better correlation compare to daily rainfall and also shows the same trend.

Table 3.1 A

Distance between the stations (Km)

	Naivasha D.O	Naivasha W\S	Gilgil W\S	Ele.N.R.Post	Nganyoi Gate	L.Naku.N.Park	L.Naku.M.St.	Milmet
Naivasha D.O	0							
Naivasha W\S	5.09	0						
Gilgil W\S	29.98	34.5	0					
Ele.N.R.Post	47.095	51.804	23.266	0				
Nganyoi Gate	48.899	53.89	28.111	6.559	0			
L.Nak.N.Park	59.038	63.93	32.424	12.107	14.824	0		
L.Naku.M.St.	62.348	67.29	34.29	17.198	20.268	5.876	0	
Milmet	81.507	85.59	51.427	42.71	46.895	32.699	26.796	0

Table 3.1 B

Daily rainfall correlation coefficient between the stations

	Naivasha D.O	Naivasha W\S	Gilgil W\S	Ele.N.R.Post	Nganyoi Gate	L.Naku.N.Park	L.Naku.M.St.	Milmet
Naivasha D.O	1							
Naivasha W\S	0.49	1.00						
Gilgil W\S	0.21	0.25	1.00					
Ele.N.R.Post	0.20	0.21	0.21	1.00				
Nganyoi Gate	0.15	0.17	0.19	0.20	1.00			
L.Nak.N.Park	0.22	0.20	0.26	0.31	0.20	1.00		
L.Naku.M.St.	0.20	0.13	0.18	0.21	0.25	0.45	1.00	
Milmet	0.06	0.09	0.07	0.07	0.10	0.08	0.05	1.00

Table 3.1 C

Ten daily average rainfall correlation coefficient between the stations

	Naivasha D.O	Naivasha W\S	Gilgil W\S	Ele.N.R.Post	Nganyoi Gate	L.Naku.N.Park	L.Naku.M.St.	Milmet
Naivasha D.O	1							
Naivasha W\S	0.90	1.00						
Gilgil W\S	0.55	0.55	1.00					
Ele.N.R.Post	0.56	0.37	0.51	1.00				
Nganyoi Gate	0.56	0.56	0.75	0.71	1.00			
L.Nak.N.Park	0.54	0.52	0.63	0.62	0.74	1.00		
L.Naku.M.St.	0.54	0.45	0.60	0.62	0.84	0.81	1.00	
Milmet	0.35	0.55	0.48	0.44	0.54	0.48	0.37	1.00

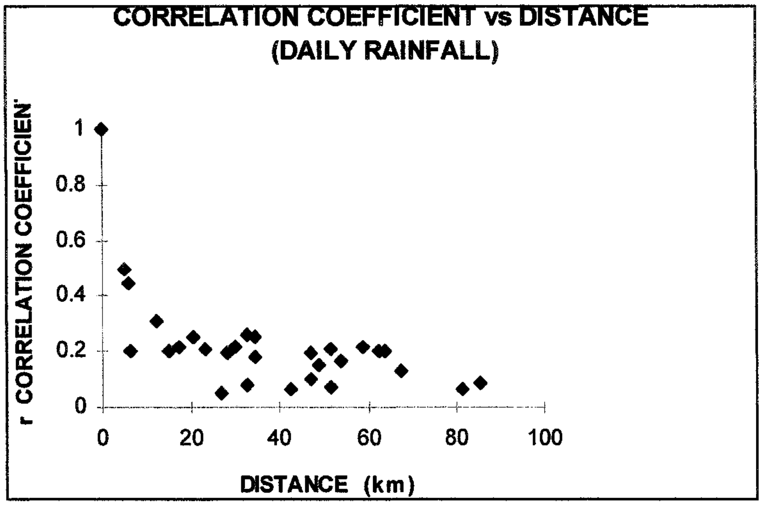

Figure - 3.4A: Correlation between daily rainfall vs. distance

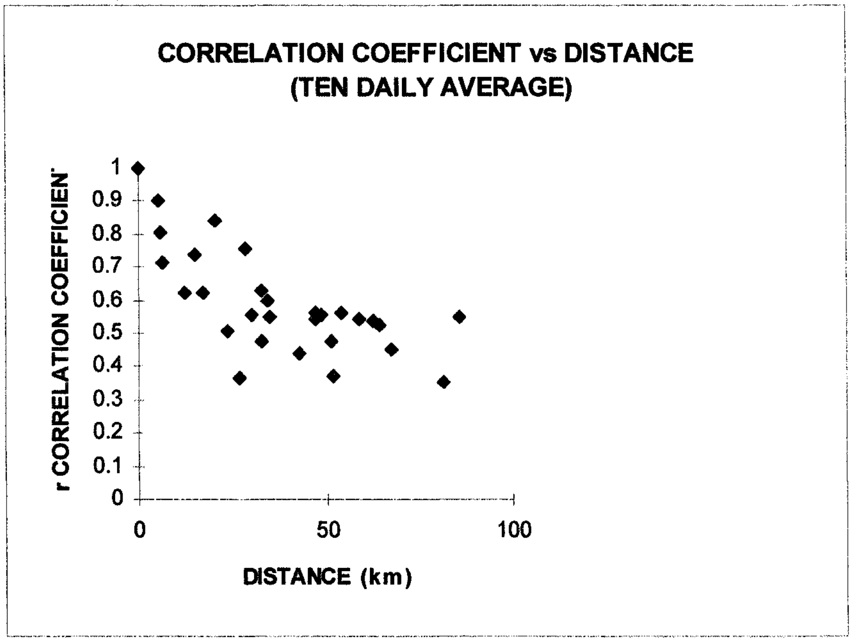

Figure - 3.4B: Correlation between ten daily rainfall vs. distance

3.5 Stream Flow Analysis

3.5.1 Consistency and homogeneity of data

The stream flow data in the selected stations were checked by double mass curve method.

3.5.2 Ten daily average

The stream flow records of the Malewa catchment are not continuous. Missing data gaps presents in the records varies from 1(one) day to years. Attempt was taken to make a ten daily average for analyzing the discharge to avoid the much discontinuity of data.

3.5.3 Double mass curve method

A plot of the accumulated ten daily average discharge against the accumulated ten daily average discharge of the surrounding stations were used to check the recorded data. The double mass curve was made as pairs of 2GB1 vs. 2GB5; 2GB1 vs. 2GC4; 2GC4 vs. 2GB5; 2GB1 vs. 2GC5; 2GB1 vs. 2GA3; 2GB1 vs. 2GB7; 2GB1 vs. 2GC7; 2GB5 vs. 2GA3; 2GB5 vs. 2GB7; 2GB5 vs. 2GA3; 2GB5 vs. 2GC5; and 2GB5 vs. 2GC7 as shown in figure - 3.5.

It can be observed that from the double mass curves station 2GB1 vs. 2GC5 (figure - 3.5D) and 2GB5 vs. 2GC5 (figure -3.5K) there seems to be a trend present in the discharge series of 2GC5 station and the data seems to be inconsistent. To verify the discharge series of 2GC5 station, the minimum discharge of each month was plotted. It was observed that there was no trend present in the minimum discharge. It was observed that after the year 1975, more records were presented during the low flow condition. It can be explain that during the ten daily average low flow records influenced the magnitude of the discharge series. It can be assumed that the discharge series of 2GC5 station is consistent.

3.6 Rainfall-runoff Relationship

A correlation between the ten daily average rainfall and ten daily average discharge was made as shown in table - 3.2. It can be observe that there is almost no correlation between rainfall and discharge. The maximum correlation (r=0.50) and minimum is no correlation (r=0). Negative correlation also observed. So, in the Malewa catchment these rainfall stations cannot be use for the prediction of runoff unless some rainfall station taking within the catchment.

3.7 Conclusions

- The rainfall and discharge records are consistent.
- The spatial distribution of rainfall is not uniform in the whole catchment.
- There is good relationship between annual rainfall and elevation.
- There is no correlation between rainfall and discharge due to rainfall stations are far away from the catchment.

DOUBLE MASS CURVE OF TEN DAILY AVERAGE DISCHARGE

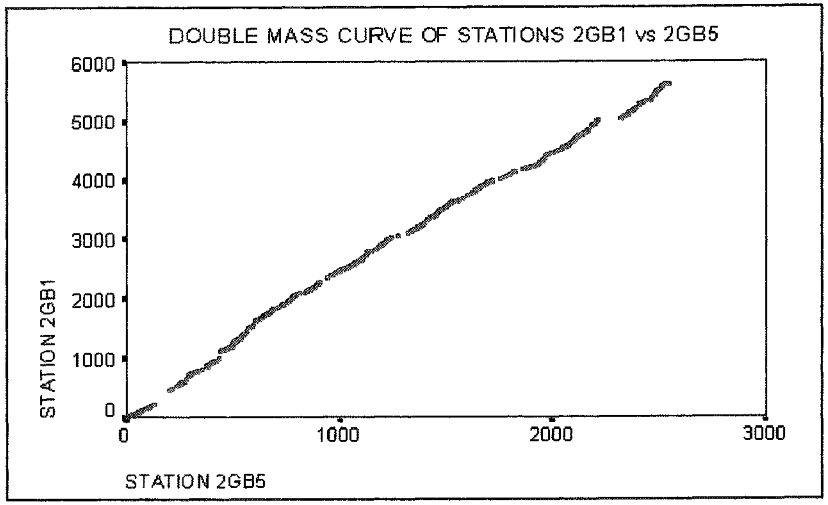

Figure A

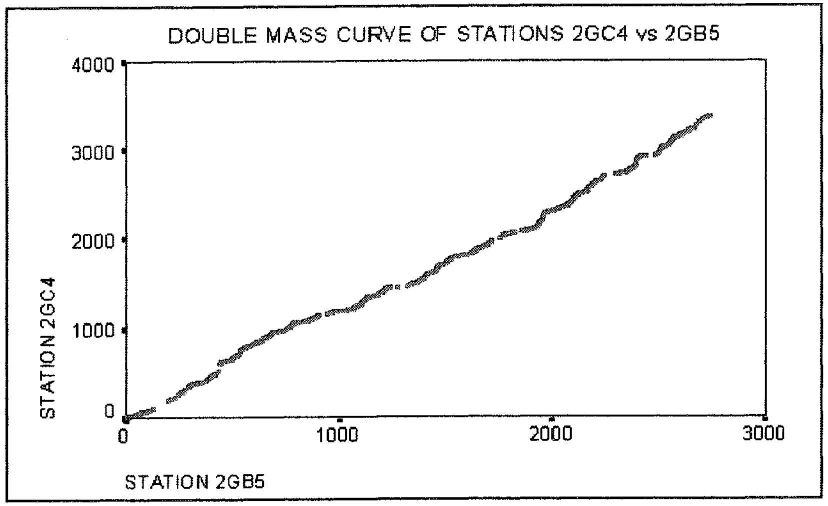

Figure - C

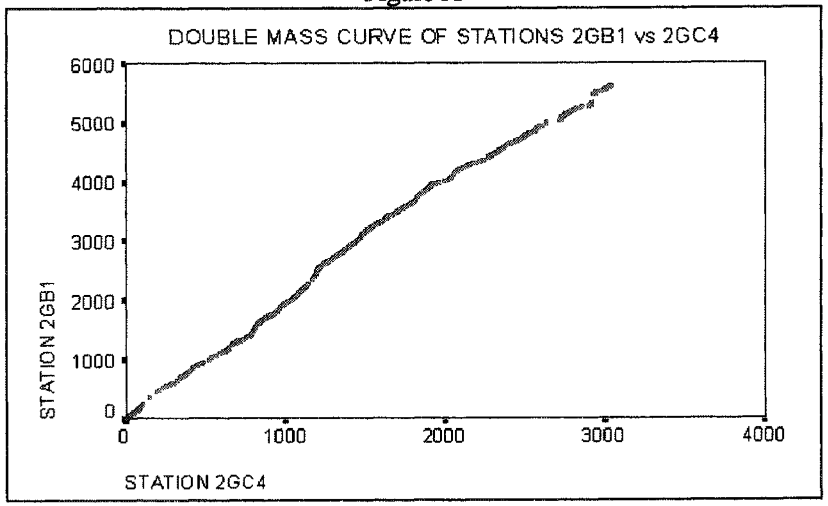

Figure B

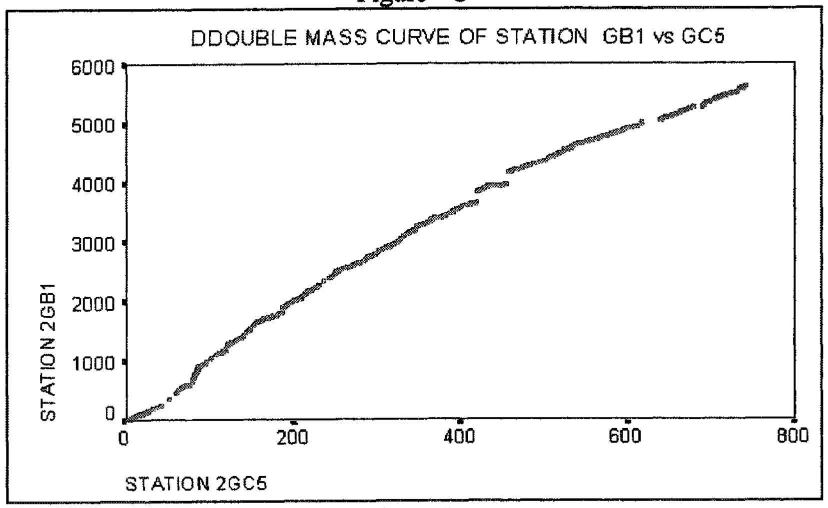

Figure D

Figure 3.5

DOUBLE MASS CURVE OF TEN DAILY AVERAGE DISCHARGE

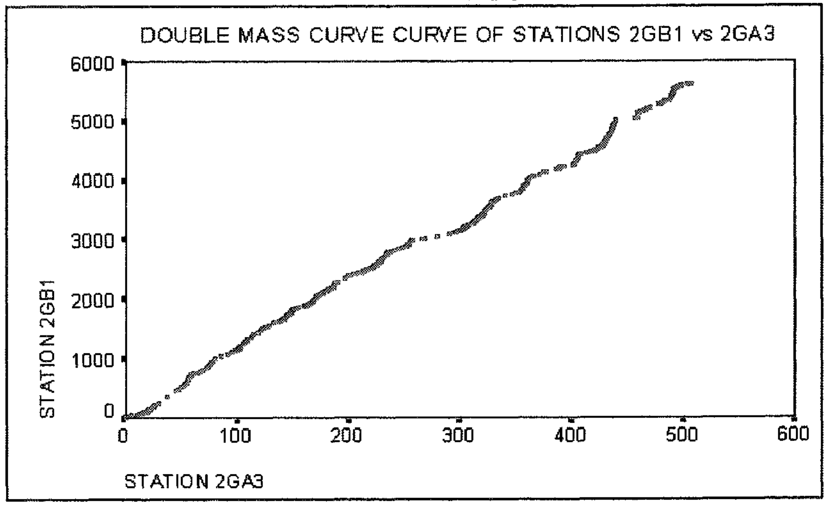

Figure E

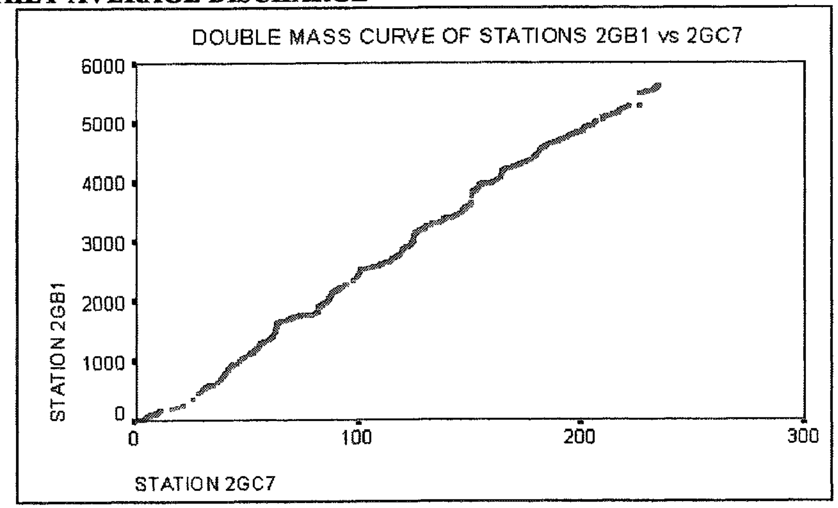

Figure G

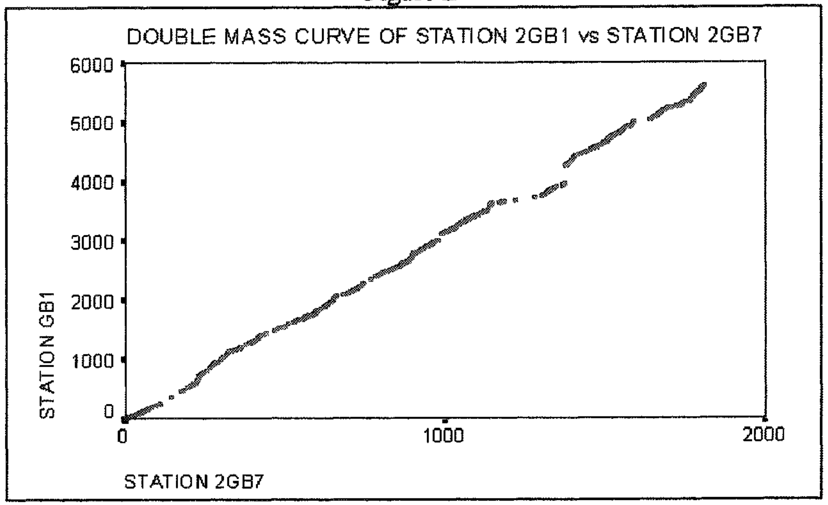

Figure F

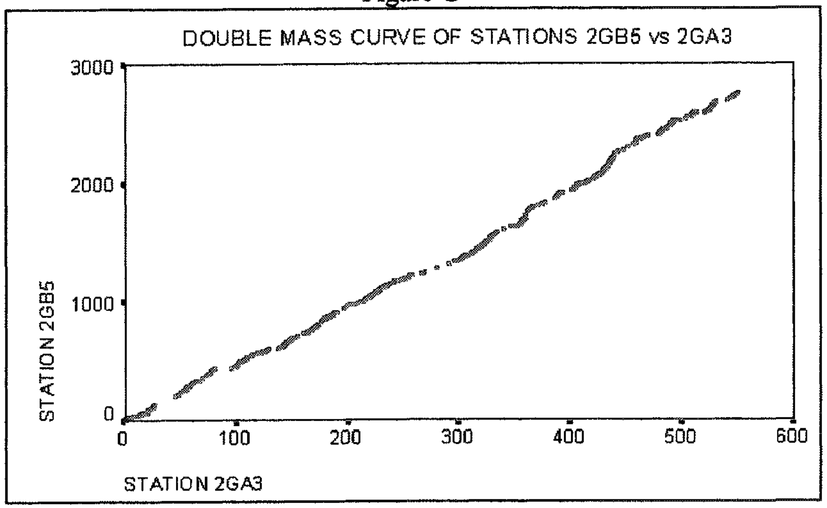

Figure H

Figure 3.5

DOUBLE MASS CURVE OF TEN DAILY AVERAGE DISCHARGE

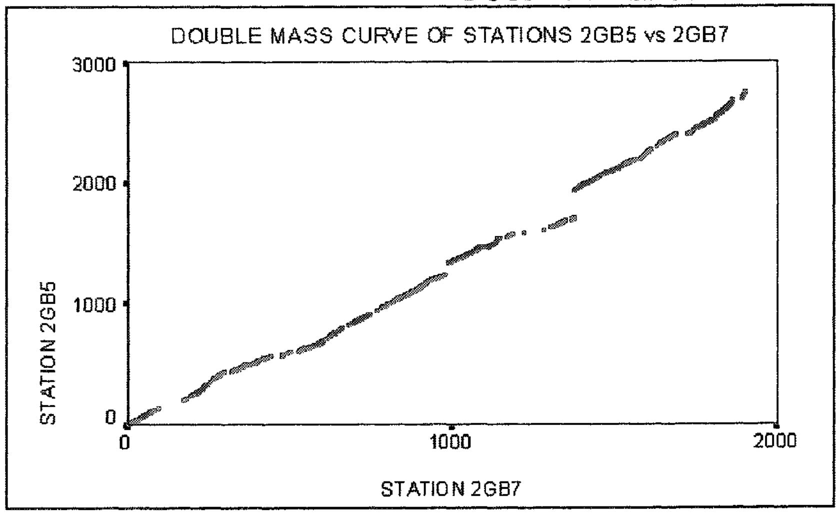

Figure I

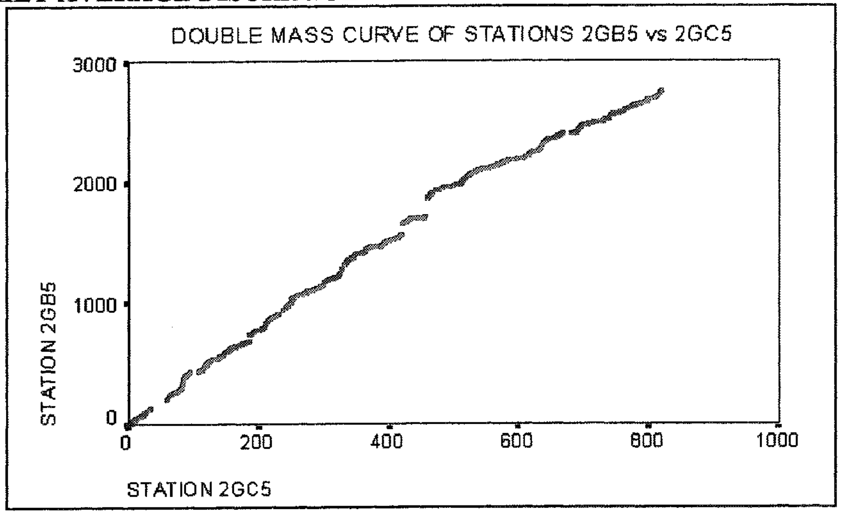

Figure - K

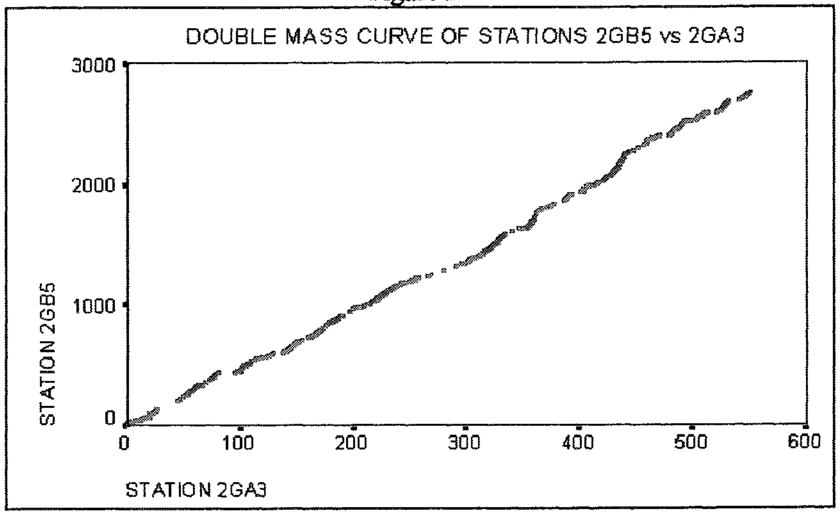

Figure J

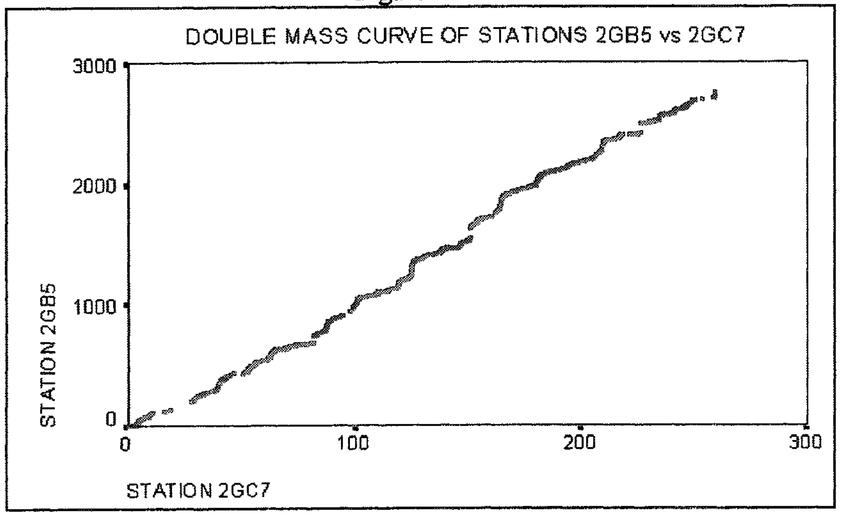

Figure L

Figure 3.5

Table 3.2

Correlation coefficient between the Rainfall and Gauging stations (Ten daily average)

RAINSTATION	2GA2	2GA3	2GA5	2GA6	2GB1	2GB3	2GB4	2GB5	2GB7	2GC4	2GC5	2GC7
Naivasha DO	0.49	0.18	0.12	0.03	0.41	0.12	0.40	0.24	0.27	0.41	0.36	0.50
Elemen.N.R.Post	0	0.23	0.22	0.18	0.39	0.48	0.40	0.34	0.37	0.42	0.30	0.34
Naku.Met.St.	0.31	0.23	0.26	0.02	0.44	0.17	0.49	0.36	0.32	0.40	0.34	0.41
Milmet	0.36	0.13	0.17	-0.02	0.27	0.12	0.36	0.18	0.19	0.25	0.25	0.28
Gilgil W\S	0	0.23	0.03	0.13	0.41	0.39	0.45	0.35	0.41	0.41	0.32	0.32
Naivasha W\S	0	0.08	-0.09	-0.01	0.02	0.35	0.32	0.21	0.27	0.37	0.28	0.45
L.Naku.Na.Park	0	0.17	0.16	0.06	0.33	0.19	0.43	0.29	0.27	0.38	0.29	0.36
Nganyoi Gate	0	0.09	0.21	0.25	0.33	0.24	0.47	0.05	0.43	0.26	0.09	0.17

CHAPTER FOUR

COMPARISON OF GAUGING STATIONS AND OPTIMIZATION OF RATING CURVE COEFFICIENTS

4.1 Introduction

The stage-discharge relationship or rating curve at a river cross-section is a fundamental technique in hydrology employed for determining discharge from catchments. Typically discharge ratings for the gauging stations are determined empirically by means of stream discharge and stage measurement in the field simultaneously. The problem in applying this empirical relationship (rating formula) is that the coefficients of the empirical formula vary with changing the channel cross-sections which are common in natural channel.

The rating curve needs to be optimized for time to time. This can be done by the field measurement of discharge and stage, and optimizing the rating coefficients. This optimization can also be done by comparing the discharge of the other stations of the same river if the field measurements were not taken in that time.

4.2 Selection of Gauging Stations

Three main gauging stations of the Malewa catchment were selected for comparison and optimization of the rating curve coefficients. These stations are: (i) station 2GC4 for the Turasha sub-catchment which contributes as major inflow to the Malewa river, (ii) station 2GB5 of the upper Malewa sub-catchment, and, (iii) stations 2GB1 for the main Malewa river (see figure 4.1).

4.3 Methodology

The techniques to verify the quality of discharge data vary from simple comparison to complex statistical analysis. The application of any particular method or technique depends on basin characteristics, drainage pattern or availability and data characteristics of the records. Three methods or techniques were used to verify the quality of available data and to optimize the rating curve coefficients as shown in flow diagram (see figure 4.2). These Techniques are: (a) Comparison of discharge between the gauging stations, (b) Reconstruction of rating curves and Comparison with existing rating curves, and (c) Develop a model and optimize the rating curve coefficients.

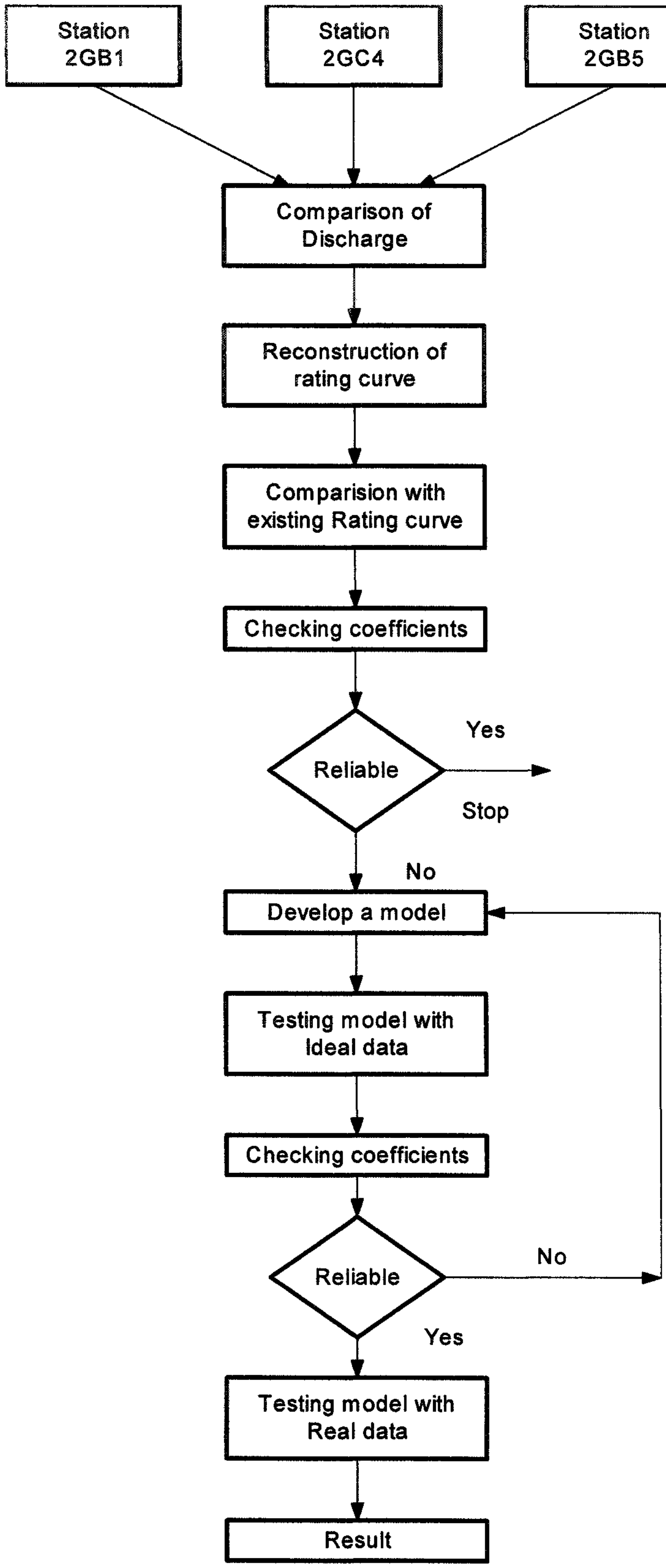

Figure 4.2 Flow diagram for comparison of the gauging stations and optimization of the rating curve coefficients

4.4 Caparison of Discharge of Three Gauging Stations

In order to compare the discharge of the above mentioned three gauging stations the following formula was used:

$$Q_{2GB1} = Q_{2GB5} + Q_{2GC4} + Q_M \quad (4.1)$$

Where Q_{2GB1} is the total inflow of the Malewa river (m^3/s), Q_{2GB5} is the inflow from the upper Malewa river, Q_{2GC4} is the inflow from the Turasha river (m^3/s) and Q_M is the inflow from the lower part of the Turasha and upper Malewa area.

The inflow between the station Q_{2GB1} and $(Q_{2GB5} + Q_{2GC4}) = Q_M$, Q_M is very low compared to the inflow of other stations. So, we can say:

$$Q_{2GB1} \approx Q_{2GB5} + Q_{2GC4} \quad (4.2)$$

The difference of discharge was calculated by adding the discharge of 2GB5 and 2GC4 stations (2GB4 + 2GC4) and subtracted them from 2GB1 station.

The average difference discharge between the measured discharge of 2GB1 station and calculated discharge (2GB5 + 2GC4) is "-.1038" m^3/s and standard deviation is 4.39 m^3/s. The negative sign (-) indicates that the calculated discharge higher than the measured discharge.

The results of the comparison between the observed discharge and predicted discharge can be seen in figure 4.3. It can be observed that during the years 1960 - 1970 the discharge data series seems to be of good quality. The predicted discharge is very close to observed discharge as shown in figure 4.3a, 4.3b and 4.3c. After the year 1970, it can be seen that the predicted discharge series sometimes significantly deviates from the observed discharge series as shown in figure 4.3d and 4.3e. So, the quality of discharge data series after the year 1970 seems to be not so good. Note that sometimes there is a gap in the displayed series (especially after 1970).

It can be describe that during the colonial time and just after the colonial time the quality of data is good due to proper management. After the colonial period due to budget limitation the data was not taken properly and even after September-1985 no data has been recorded for 2GB1 station.

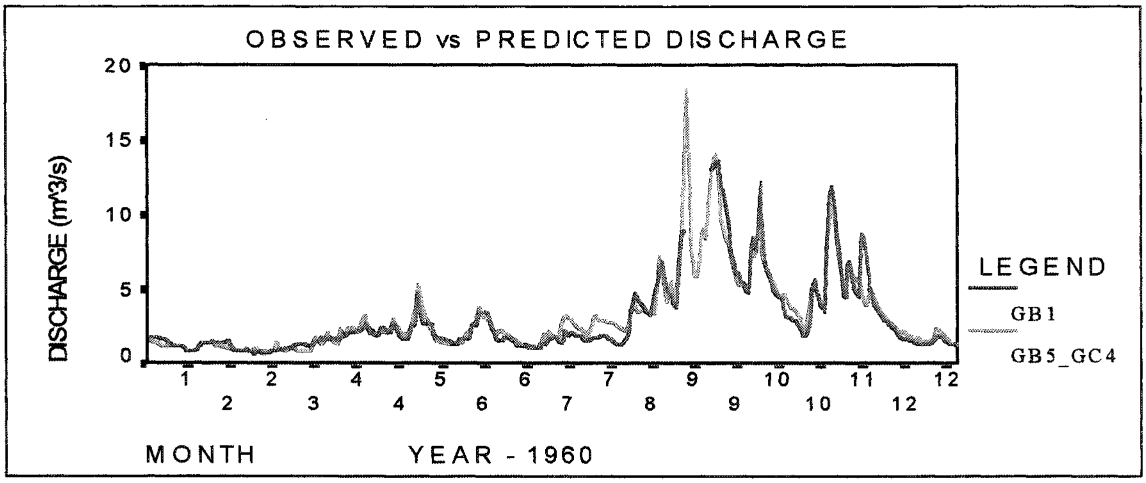

Figure 4.3a: Measured and predicted discharge (2GB5 and 2GC4) for the year-1960

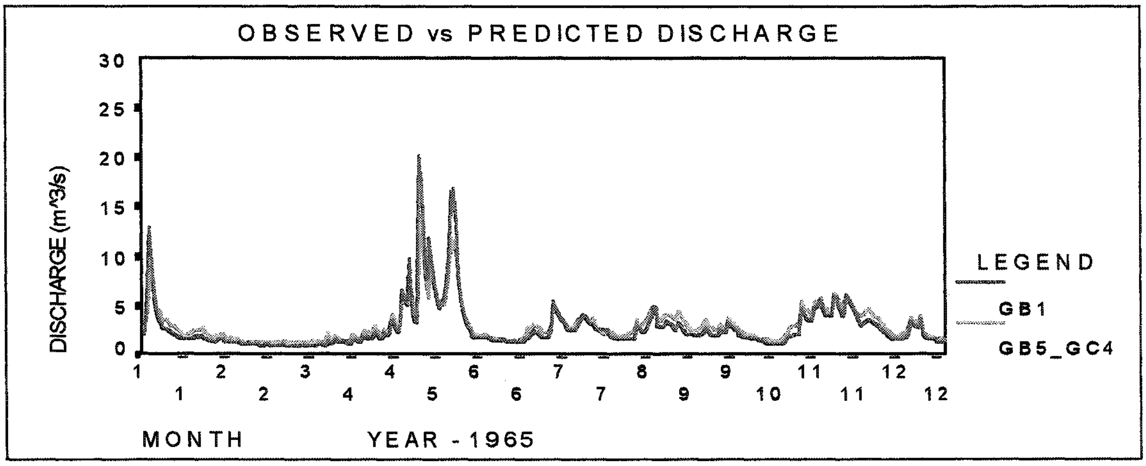

Figure 4.3b: Measured and predicted discharge (2GB5 and 2GC4) for the year -1965

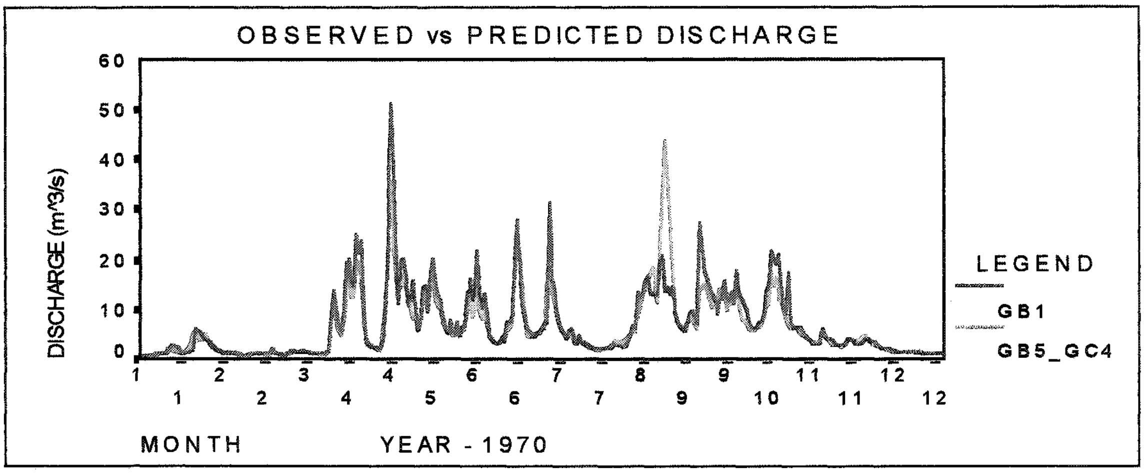

Figure 4.3c: Measured and predicted discharge (2GB5 and 2GC4) for the year - 1970

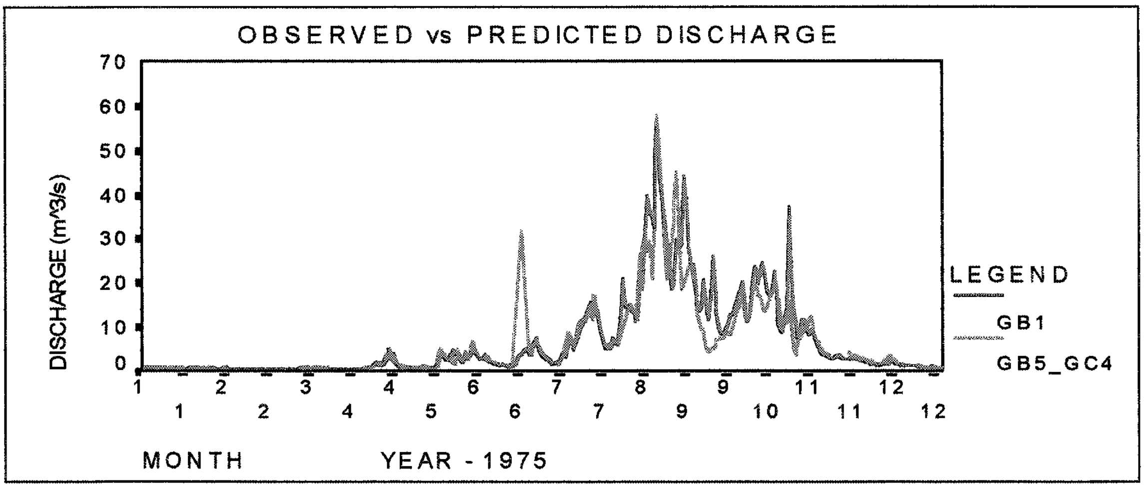

Figure 4.3d: Measured and predicted discharge (2GB5 and 2GC4) for the year - 1975

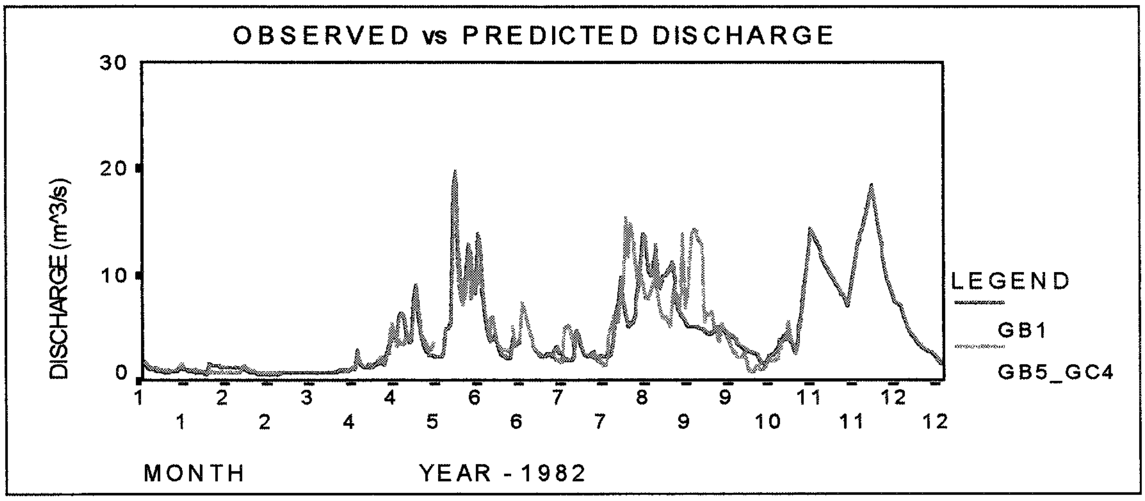

Figure 4.3e: Measured and predicted discharge (2GB5 and 2GC4) for the year - 1982

4.5 Reconstruction of the Rating Curve and Comparison with Existing Rating Curve

In order to compare the rating curves of the above mentioned three gauging stations, three rating curves were reconstructed. The rating curves were constructed with the field measurement of discharge and water level data. The constructed rating curves were compared with the rating curves used by the Ministry of Water Development, Kenya. The rating equations were established by plotting the logarithms of discharge against the logarithms of stage and at the same time the correlation coefficient and standard error also calculated as shown below.

4.5.1 Rating curve for the 2GB1 station

The field measured discharge was plotted against stage to obtain the rating curve in logarithmic scale as shown in figure 4.4. It can be observed that the rating curve shows a straight line without any break point. The stage-discharge (H-Q) relationship for this station that was calculated can be expressed as:

$$Q_{2GB1} = 27.29\,(H + 0.024)^{1.76} \qquad (4.3)$$

Where Q_{2GB1} is the discharge (m^3/s), H is the stage (meter), 27.29 and 1.76 are the constants and 0.024 is the stage correction for this station.

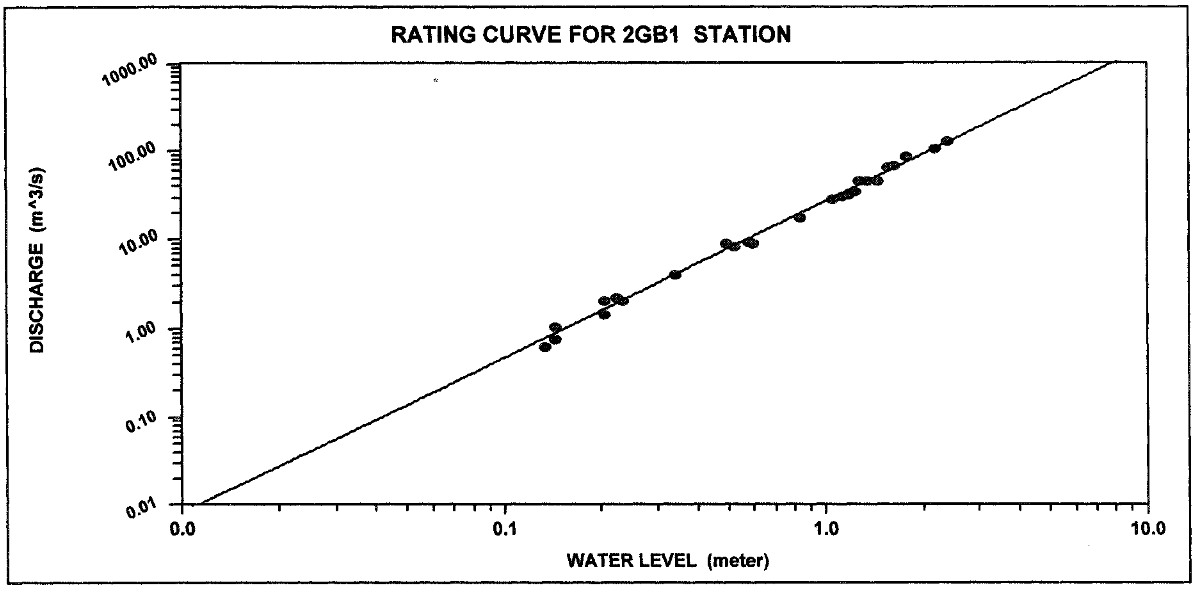

Figure 4.4: Rating curve for the 2GB1 station

The control section of 2GB1 station is complex. The lower discharge (base flow) of this station is controlled by artificial control section (complex rectangular weir) whereas the higher discharge is controlled as a natural channel as shown in photo 4.1. The overall power constant for this station is 1.76 which is higher than the theoretical value of a rectangular weir of 1.5, which can be expected for a complex rectangular weir (Operation Hydrology, Report no. 13, WMO- No. 519). For the natural control section the exponent (1.76) usually exceeds 2.0. Some parts of USA the exponent has been experienced a range from 1 to 4. The stage discharge relation shows a higher correlation (correlation coefficient R = 0.996) between the two and the standard deviation is 3.29. The natural control section for this station also as like as a rectangular channel due to topography. So, the rating equation for the 2GB1 station can be applied for both for artificial and natural control.

The rating curve used by the ministry for the same gauging station:

$$Q_{2GB1} = 27.33\ H^{1.78} \qquad (4.3a)$$

It can be observed that the rating curve used by the ministry calculates discharge very close to constructed rating curve but without shift. The discharge series of 2GB1 station has been recalculated for further analysis.

4.5.2 Rating curve for the 2GB5 station

Similar to 2GB1 station, the field measured discharge was plotted against stage to obtain the rating curve in logarithmic scale as shown in figure 4.5. It can be observed that the plotting points are scattered and show a break point in the rating curve. In order to fit a the stage-discharge (H-Q) relationship (rating curve) for this station the data series had been split at the break point. The lower discharge of the break point was termed as 'lower part' and the higher discharge above the break point was termed as 'upper part'. Two separate rating curves were establish for two parts of the discharge series.

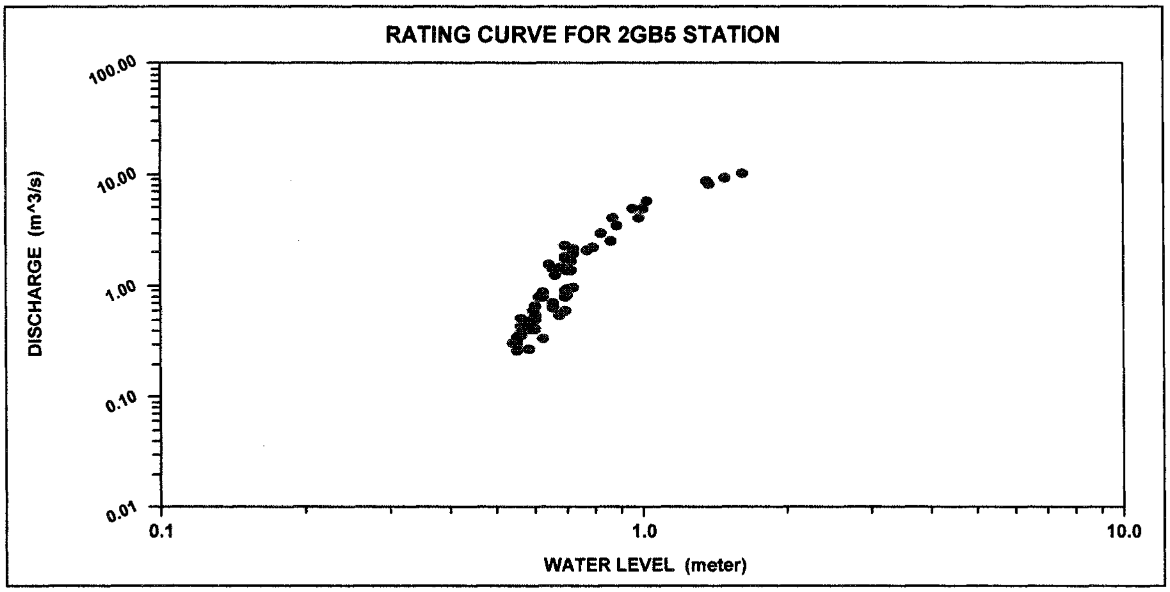

Figure 4.5 Rating curve for the 2GB5 station

(a) Rating curve for the 2GB5 station (Lower Part)

The stage-discharge(H-Q) relationship (rating curve) for this station (figure-4.6) that was calculated can be expressed as:

$$Q_{2GB5\ (Lower)} = 10.7\ H^{5.68} \tag{4.4}$$

Where $Q_{2GB5(Lower)}$ is the discharge (m^3/s), H is the stage (meter) and 10.7 and 5.68 are the constants for this station.

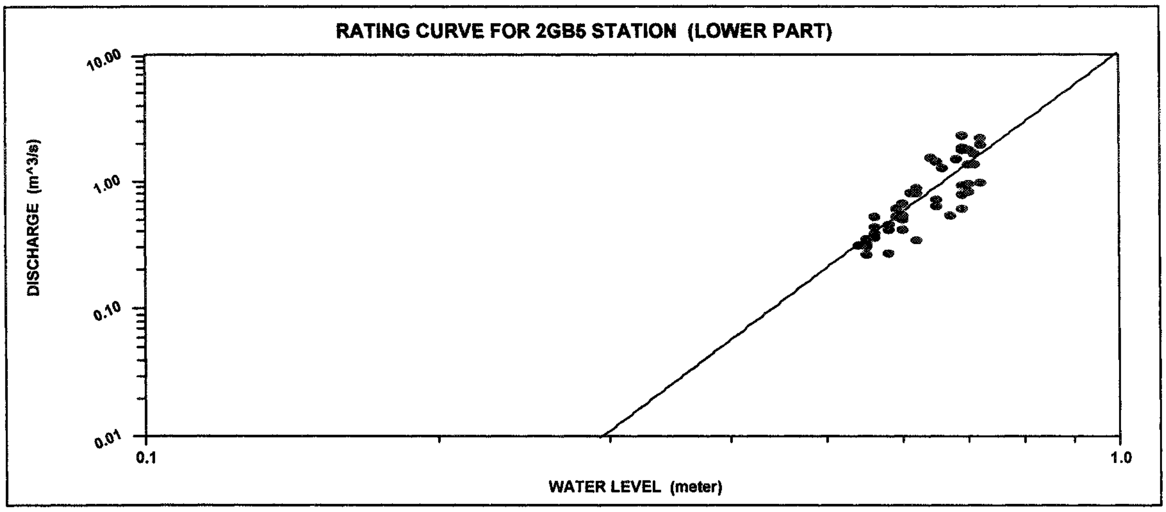

Figure 4.6: Rating curve for the 2GB5 station (Lower part)

(b) Rating curve for the 2GB5 station (upper Part)

The stage-discharge(H-Q) relationship (rating curve) for this station (figure-4.7) that was calculated can be expressed as:

$$Q_{2GB5\ (Upper)} = 4.56\ H^{1.8} \tag{4.5}$$

Where $Q_{2GB5(Upper)}$ is the discharge (m^3/s), H is the stage (meter) and 4.56 and 1.8 are the constants for this station.

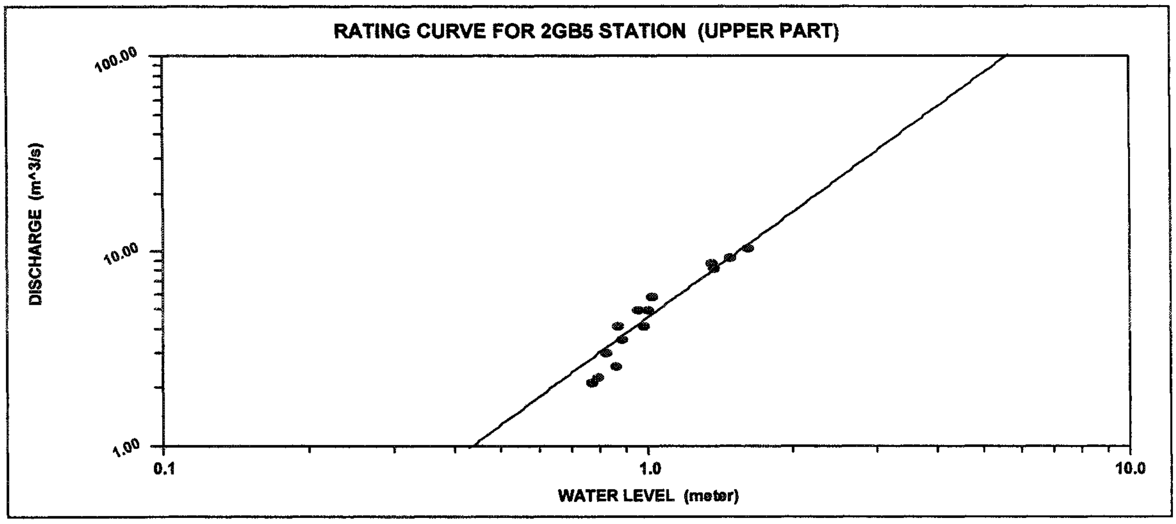

Figure 4.7 Rating curve for the 2GB5 station (upper part)

The control section of 2GB5 station is natural control section as shown in photo 4.2. The power constant for lower discharge (base flow) of this station is 4.68 which is higher than the expected value whereas the higher discharge (flood) shows the power constant for this station is 1.8 which is seems to be reliable. For the natural channel, depending on the control section the power constant ranges from 1 to 4 (Operation Hydrology, Report no. 13, WMO- No. 519). The stage discharge relation for the "lower part" shows a good correlation (correlation coefficient $R_{(lower)} = 0.79$) between the two and the lower standard error ($\sigma_{(lower)} = 0.36$). The "upper part" of the rating curve shows a very good correlation (correlation coefficient $R_{(upper)} = 0.97$) between the stage and discharge and the standard error ($\sigma_{(upper)} = 0.66$) which seems to be higher than the "lower part". It can be explain that the standard error is higher due to the magnitude of discharge during high flow.

The rating curve used by the ministry for the same gauging station:

$$Q_{2GB5\,(Lower)} = 31.411\ H^{8.47} \qquad (4.4a)$$

and

$$Q_{2GB5\,(Upper)} = 3.997\ H^{2.05} \qquad (4.5a)$$

It can be observed that the power constant for the "lower part" of the rating curve used by the Water Development Ministry, Kenya is too higher than the calculated one but both of them higher than expected and seems to be not reliable (power constant varies from 1-4; Operation Hydrology, Report no. 13, WMO- No. 519). The power constant for the "upper part" of the rating curve is seems to be reliable but it differs significantly which was calculated from the field measurement data. It can be observed that the rating curve which was established from field measurement significantly differs from supplied rating curves. So, it is necessary to recalculate the discharge series of this station to have the reliable data for further analysis.

4.5.3 Rating curve for the 2GC4 station

Similar to 2GB1 station, the measured discharge was plotted against stage to obtain the rating curve in logarithmic scale as shown in figure 4.8. It can be observed that, like station 2GB5, the rating curve shows the scattered points and a break point in the rating curve for this station too. In order to fit a the stage-discharge (H-Q) relationship (rating curve) for this station the data series was split at the break point. The lower discharge of the break point was termed as 'lower part' and the higher discharge above the break point was termed as 'upper part'. Two separate rating curves were establish for two parts of the discharge series. A number of measurements are available for lower discharge but only three measurement is available for higher discharge which is not enough for establish for reliable rating curve.

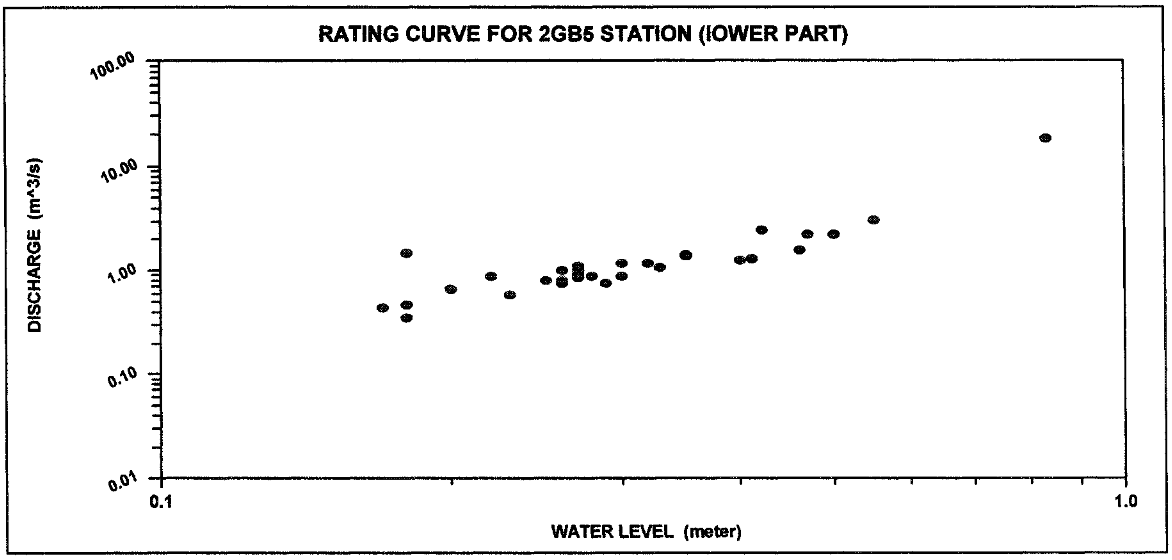

Figure 4.8 Rating curve for the 2GC4 station

(a) Rating curve for the 2GC4 station (Lower Part)

The stage-discharge(H-Q) relationship (rating curve) for this station (figure-4.9) that was calculated can be expressed as:

$$Q_{2GC4\ (Lower)} = 6.52\ H^{1.52} \qquad (4.6)$$

Where $Q_{2GC4(Lower)}$ is the discharge (m^3/s), H is the stage (meter) and 6.52 and 1.52 are the constants for this station.

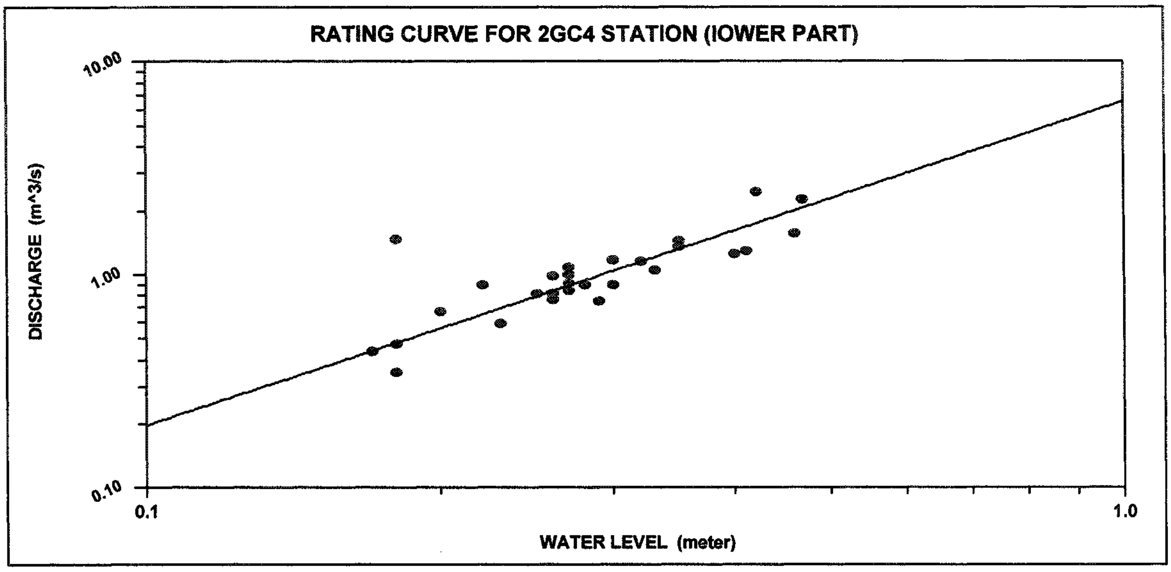

Figure 4.9 Rating curve for the 2GC4 station (Lower part)

(b) Rating curve for the 2GC4 station (upper Part)

The stage-discharge(H-Q) relationship (rating curve) for this station (figure-4.10) that was calculated can be expressed as:

$$Q_{2GC4\,(Upper)} = 37.62\ H^{3.99} \qquad (4.7)$$

Where $Q_{2GC4(Upper)}$ is the discharge (m^3/s), H is the stage (meter) and 37.62 and 3.99 are the constants for this station.

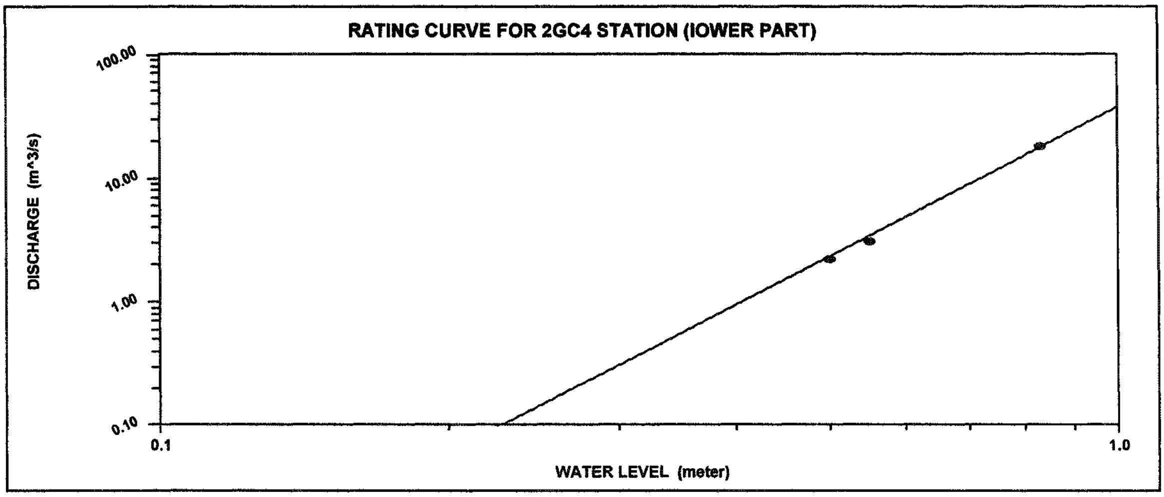

Figure 4.10 Rating curve for the 2GC4 station (Upper part)

The control section of 2GC4 station is complex like 2GB1. The lower discharge (base flow) of this station is controlled by artificial control section (complex rectangular weir) whereas the higher discharge is controlled as a natural channel as shown in photo 4.3. The power constant for lower discharge (base flow) of this station is 1.52 which is very close to the theoretical value. Besides this, the higher discharge (flood) shows the power constant for this station is 3.99 which is also reliable. The stage discharge relation of "lower part" of the rating shows a lower correlation (correlation coefficient R=0.79) but the "higher part" shows a higher correlation (correlation coefficient R=0.999). The standard error for the both parts of the of the rating curve is same (0.29).

The rating curve used by the ministry for the same gauging station:

$$Q_{2GC4\ (Lower)} = 4.95\ H^{1.55} \tag{4.6a}$$

$$Q_{2GC4\ (Upper)} = 14.95\ H^{2.79} \tag{4.7a}$$

It can be observed that the rating curve used by the Water Development Ministry, Kenya is seems to be reliable for both high and low flow conditions but it is significantly differs from the calculated one. So, the discharge series of this station was recalculated for further analysis.

STATION 2GB1

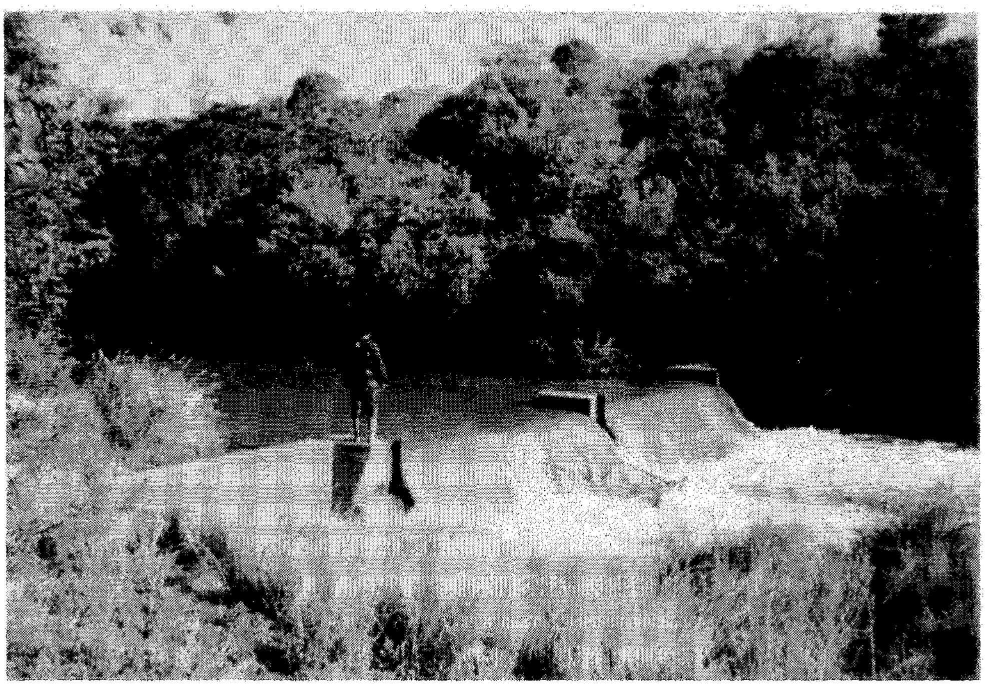

Photo 4.1 Lower discharge control by artificial structure

STATION 2GB5

Photo 4.2 The natural channel control

STATION 2GC4

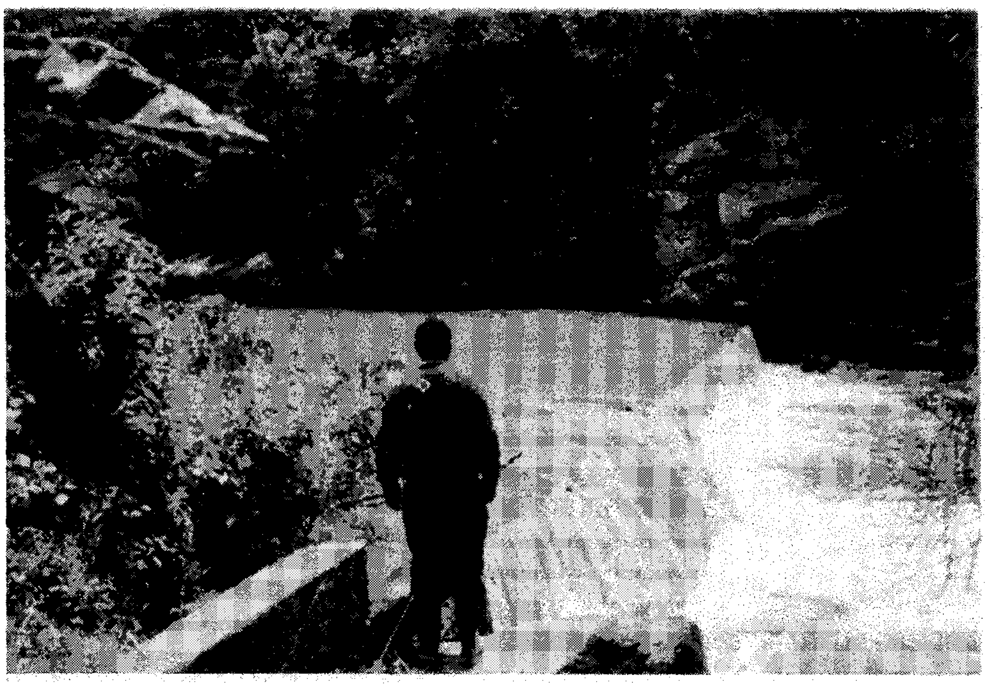

Photo 4.3 Lower discharge control by artificial structure

4.6 Develop a Model for Optimization of the Rating Curve Coefficients

In order to optimize the rating curve coefficients, a model was developed. For developing a model the ideal situation was considered. It was assumed that the combined inflow of the stations 2GB5 and 2GC4 was equal to the inflow of 2GB1 station. There was no inflow and outflow in between these stations. Three rating curves were taken for three gauging stations which are similar to the reconstructed rating curve for that station as shown in equation 4.8.

$$Q_{2GB1} \quad = \quad Q_{2GB5} \quad + \quad Q_{2GC4} \qquad 4.8$$

$$[C_1 (H_1+a)^{n_1}] \qquad [C_2 H_2^{n_2}] \qquad [C_3 H_3^{n_3}]$$

where $Q_{2GB1} = [C_1 (H_1+a)^{n_1}]$ is the rating curve for 2GB1 station, $Q_{2GB5} = [C_2 H_2^{n_2}]$ is the rating curve for 2GB5 station, $Q_{2GC4} = [C_3 H_3^{n_3}]$, is the rating curve for 2GC4 station , H is the water level (Stage), a is the datum correction and C, n are constants.

The discharge was calculated for each station using the respective rating curve. The discharge of station 2GB5 and 2GC4 was added and subtracted this from 2GB1. In order get the difference of discharge in absolute value the differences were squared and were summed them. The sum of squares was minimized by changing the rating coefficients (C, n). This procedure continued till the optimum (satisfactory) results come and at the same time the effect of the changing coefficient observed as shown in flow diagram figure 4.11.

4.6.1 Testing the model with ideal data

In order to test the model, three sets of artificial water level data were created and three sets of rating curve coefficients were taken which are similar to the rating curves of stations 2GB1, 2GB5 and 2GC4 respectively. The discharge data was synthesized by using the water level data series and the corresponding rating curves. The model was tested by changing the rating curve coefficients by minimizing the sum of difference squares and at the same time the changes in magnitude of the optimized discharge data series from the original discharge series was observed. It was also observed that the number of rating coefficients needs to constraint for this type of model for reliable optimization. The summary of optimization results have been presented in table 4.1

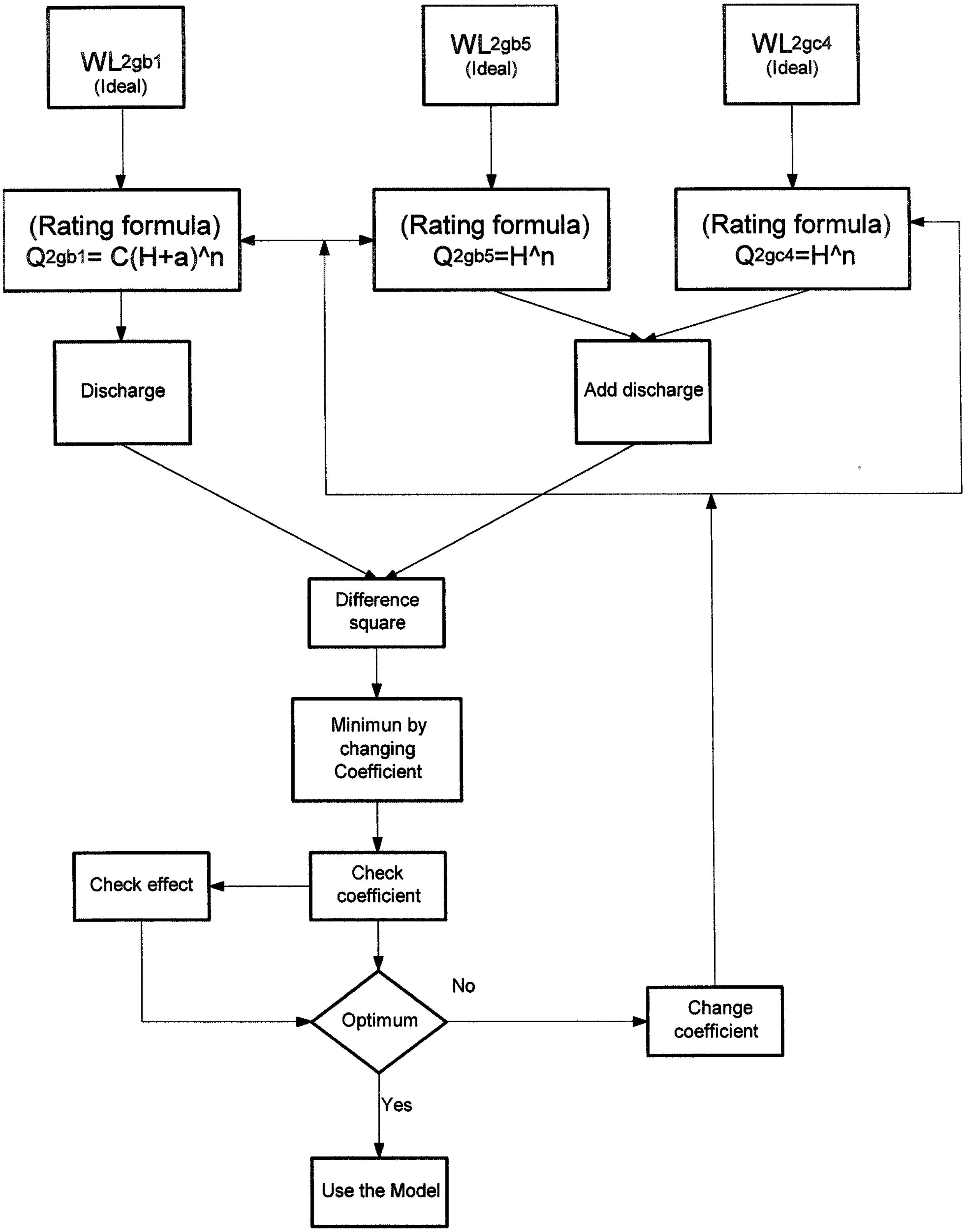

Figure 4.11: Flow diagram for optimization Model

Table 4.1 Optimization of the rating curve coefficients (Ideal condition)

Sl. No.	Iteration	Rating coefficient Station 2GB1		Rating coefficient Station 2GB5		Rating coefficient Station 2GC4		Min.sum of Diff.sqares	Effect on Discharge series
		C_1	N_1	C_2	N_2	C_3	N_3		
1	2	3	4	5	6	7	8	9	10
*	Original	25	1.8	10	2.05	14.5	1.6	2.18	
1	Change	* 25	* 1.8	* 10	* 2.05	* 14.5	6	1980.5	No effect on
	Calculated	* 25	* 1.8	* 10	* 2.05	* 14.5	1.64	1.96	discharge
2	Change	* 25	* 1.8	* 10	* 2.05	5	6	1304.9	No effect on
	Calculated	* 25	* 1.8	* 10	* 2.05	14.53	1.64	1.94	discharge
3	Change	* 25	* 1.8	* 10	6	5	6	1764.1	No effect on
	Calculated	* 25	* 1.8	* 10	1.98	14.53	1.66	1.82	discharge
4	Change	* 25	* 1.8	20	6	5	6	4784.8	No effect on
	Calculated	* 25	* 1.8	10.43	1.96	14.11	1.7	1.67	discharge
5	Change	* 25	6	20	6	5	6	1636.4	decrease of
	Calculated	* 25	-7E-07	10	-6.80E-07	14.03	-7.60E-07	3.00E-11	discharge
6	Change	10	6	20	6	5	6	7447.4	large decrease
	Calculated	0.02	0.93	0.01	1.025	0.013	0.91	3.00E-06	of discharge
7	Change	10	* 1.8	20	* 2.05	5	* 1.6	3110.86	large decrease
	Calculated	2.1	* 1.8	8.00E-08	* 2.05	1.30E-07	* 1.6	1.50E-16	of discharge
8	Change	* 25	6	* 10	6	* 14.5	6	173.4	small decrease of
	Calculated	* 25	1.3	* 10	1.44	* 14.5	1.17	1.39	discharge

Not: '*' indicates the cell value Constraints

The table 4.1 shows that two rating curve coefficients can be changed freely (sl.no. 1 - 4), without any changes in discharge data series. When only one coefficient "C_1" was taken constant (sl.no. 5) and remaining coefficients can be changed freely, the optimization procedure calculates discharge series which decreases in magnitude from the original discharge series for all the gauging stations. It can be observed that when all the rating coefficient can be changed freely (sl.no. 6), the optimization procedure calculates the large decreases in magnitude of discharges from the original discharge series for all the stations. When all the "C" coefficients can be changed freely (sl.no. 7), the optimization procedure also calculates the large decreases in magnitude of discharges. On the other hand, when all the "n" coefficients can be changed freely, the optimization procedure calculates the small decreases of discharges. The optimization results shows that the changes of "C" coefficient of a rating curve, significantly changes the magnitudes of discharges compare to "n" coefficient.

So, it is clear that, for the reliable optimization of the rating coefficients, maximum two rating curve coefficients can be changed freely by this model.

4.6.2 Applying model with real data

In order to optimize the rating coefficients of the above mentioned three gauging stations, the optimization model was applied for the years 1960 - 1970 due to quality of data. The discharge series was split into high flow data series and low flow data series. The water level of 2GC4 station was selected for the splitting the discharge series. The water level above the break point of the rating curve was treated as "high flow" discharge series and water level below the break point was treated as "low flow" discharge series. The problem is that some low flow portion the discharge series of 2GC4 station falls into both high flow and low flow series of 2GB5 station and 2GB1 station and vise versa. Similarly when the discharge series of 2GB5 station selected for splitting the discharge series repeats the same situation for remaining stations. Two types of optimization were performed for both high flow and low flow series depending upon weight. These are: (a) optimization without weight and (b) optimization with weight.

(a) Optimization without weight:

The optimization was performed in normal way without minimizing the effect of large differences. The higher differences predominate in this optimization procedure.

(b) Optimization with weight:

The optimization was performed in such a way that the effect of large differences was minimized and every difference has almost the same effect on the optimization procedure. This was done by dividing the difference of discharge with the sum of discharge of three gauging stations.

The rating coefficients of 2GB1 station were taken as constraints and the rating coefficients of gauging stations 2GB5 and 2GC4 were used for optimization for both cases. The results of optimization have been shown in table 4.2a and 4.2b.

Table 4.2a: Optimization of the rating curve coefficients (No Weight)

	Rating coefficient Station 2GB1		Rating coefficient Station 2GB5		Rating coefficient Station 2GC4		Min.sum of Diff.sqares	Flow Condition
	C_1	N_1	C_2	N_2	C_3	N_3		
1	2	3	4	5	6	7	8	9
Original	27.29	1.76	10.7	5.68	6.52	1.52	36853	Low flow
Optimized	27.29	1.76	5.21	2.86	24.00	3.53	1121	
Original	27.29	1.76	4.56	1.80	37.62	3.99	9118	High flow
Optimized	27.29	1.76	5.09	1.60	21.80	3.21	4383	

Table 4.2b: Optimization of the rating curve coefficients (With Weight)

	Rating coefficient Station 2GB1		Rating coefficient Station 2GB5		Rating coefficient Station 2GC4		Min.sum of Diff.sqares	Flow Condition
	C_1	N_1	C_2	N_2	C_3	N_3		
1	2	3	4	5	6	7	8	9
Original	27.29	1.76	10.7	5.68	6.52	1.52	18.14	Low flow
Optimized	27.29	1.76	5.86	3.89	16.69	2.16	9.15	
Original	27.29	1.76	4.56	1.80	37.62	3.99	9.37	High flow
Optimized	27.29	1.76	5.42	1.84	26.40	4.28	5.73	

It can be observed that the optimized rating coefficients of 2GB5 station reduced its magnitude during the "low flow" condition for both with assigning weight and without weight and the coefficients are within the reliable limit (see table 4.2a and 4.2b). The optimized rating coefficients for the same station during the "high flow" condition, the "C" constant increases but the power constant decreases for without assigning weight. In weight case, the both the coefficients increase their magnitude. For both the cases (weight and without weight), the coefficients are within the reliable limit (see table 4.2a and 4.2b).

For 2GC4 station, the optimized rating coefficients increase its magnitude for both with assigning weight and without weight for "low flow" conditions. The optimized rating coefficients for this station during "low flow" condition without weight is higher than the reliable (n_3 =3.53 expected upto 2.00) and with assigning weight it seems to be close to reliable limit (N_3=2.16, see table 4.2b) for complex rectangular weir. During the high flow condition without assigning weight optimized power coefficient seems to reliable limit (N_3= 3.21 <4, see table 4.2a) and with assigning weight, it seems to be higher than the reliable limit (N_3= 4.28 > 4, see table 4.2b) as it flow as a natural channel during high flow condition (Operation Hydrology, Report no. 13, WMO- No. 519).

4.7 Conclusions

- The quality of discharge data series of three main gauging station is better during the year 1960-1970 and after that the quality of data series decreases.
- The rating curve coefficients of 2GB1 and 2GC4 gauging stations, seem to be reliable for both high and low flow conditions. The rating curve coefficients of the gauging station 2GB5 seems to be reliable only high flow condition.
- The reconstructed rating curves significantly differ from the supplied rating curves.

- The optimization is possible for two rating curve coefficients at the same time using the three rating curves by the developed model (procedure).
- The optimization with assigning weight procedure produces reliable rating coefficients during low flow condition.
- The optimization without assigning weight procedure produces reliable rating coefficients during high flow condition.

CHAPTER FIVE

INFILLING OF MISSING DATA

5.1 Introduction

Design operations of water resources management systems in many countries, particularly in the developing world, country like Kenya often suffer from inadequate data.

For the planning, design and operation of complex water resources systems, discharge data are required at several sites simultaneously. Although a few series may be sufficient long, it is generally found that several are inadequate length. This may be due to gaps commencement of measurement has not been simultaneously. Therefore, it is necessary to infill missing data before practical application of the series.

5.2 Description of Drainage System

The Malewa is the main river of the Malewa catchment and the main source of surface water of the lake Naivasha. The Malewa river has two main tributaries named Malewa and Turasha and their branches. A number of small tributaries started from the north and north-east of the Nyandarua Ranges (Aberdares mountain) joining each other flowing to the down stream as Malewa by the name (see figure 5.1). The gauging stations of this river are 2GB1, 2GB3, 2GB4, 2GB5 and 2GB7. The 2GB1 and 2GB5 stations are the most important for this river. Many small tributaries started from the east and south-east from the same mountains joining together flowing to the down stream as a Turasha river. The gauging stations of this river are 2GC4, 2GC5 and 2GC7. The 2GC4 station is the most important for this river. This river meets with Malewa river from the left side at the upstream of the main gauging station 2GB1 and the combined flow is called the main Malewa river.

5.3 Previous Study

A relatively great number of investigations have been carriedout concerning the lake and its surroundings. Mr. A. E. Tetley study with discharge series of Naivasha basin (published 13/02/1948), after that no study was done with the discharge series on the study area.

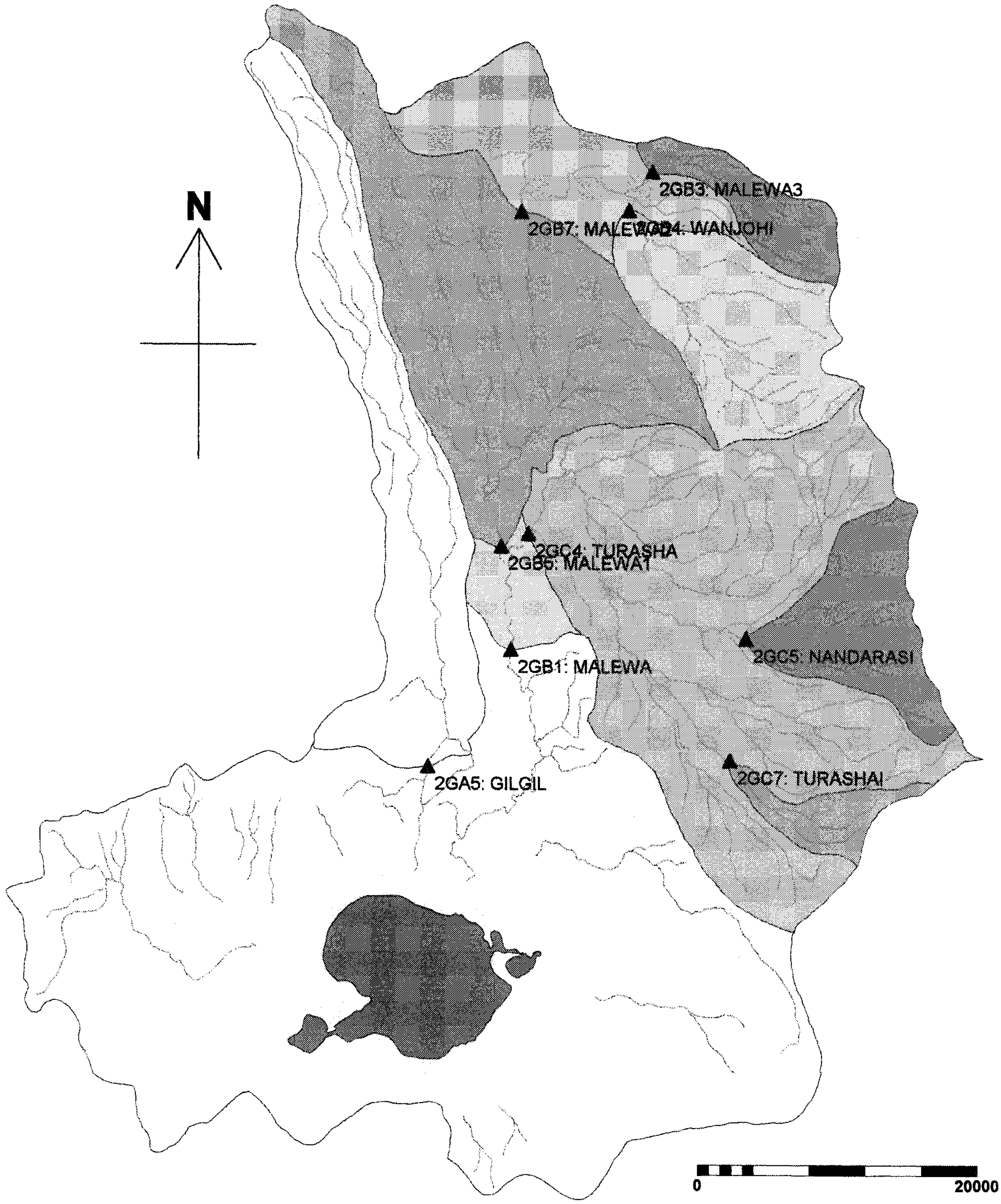

Figure - 5.1: Drainage system of the study area

5.4 Methodology

The techniques for the infilling of missing data vary from simple interpolation to models and complex statistical analysis. The application of any particular method or technique depends on the length of gaps, the season, the climatic region, basin characteristics, or availability and data characteristics of the records. Three methods were used for the infilling the missing data are shown in flow diagram (see figure 5.2). These methods are: (a) correlation, (b) Linear interpolation and (c) Multiple Linear Regression.

5.5 Input Data

In the study area, rainfall and base flow are the sources of inflow of the river. The flow data are available from the following gauging stations: 2GB1, 2GB3, 2GB4, 2GB5, 2GB7, 2GC4, 2GC5 and 2GC7. The nearby Nakuru Met. station, Naivasha D.O., Milmet, Lake Nakuru Nation Park, Gilgil W/S, and Elementaita N. Rang Post rainfall station's data were available for analysis. The missing gaps of flow data would be infilled for the station 2GB1 covering from January-1960 to December-1990.

5.6 Organization of Data

Three types of data were collected from the Ministry of Water Development, Kenya. These are discharge data series, water level data series and rainfall data series in the text format. The data were organized (in excel) as two-dimensional way tables with column shows the months and row shows the dates and years with the description of the data and the area of the catchment for each year

For the analysis of comparing the data, it needs to organize the data in two-dimensional table as a way that each column contains one variable (discharge, water level and rainfall of each station) and the rows are in cases (dates).

Reformatting the data

Steps:

- In order to reform the data for the year 1960 to 1990 were selected where available.
- For generating the continuous record, the texts between the years were removed.
- The missing value was assigned with known value (9999) by macro (see appendix-B)

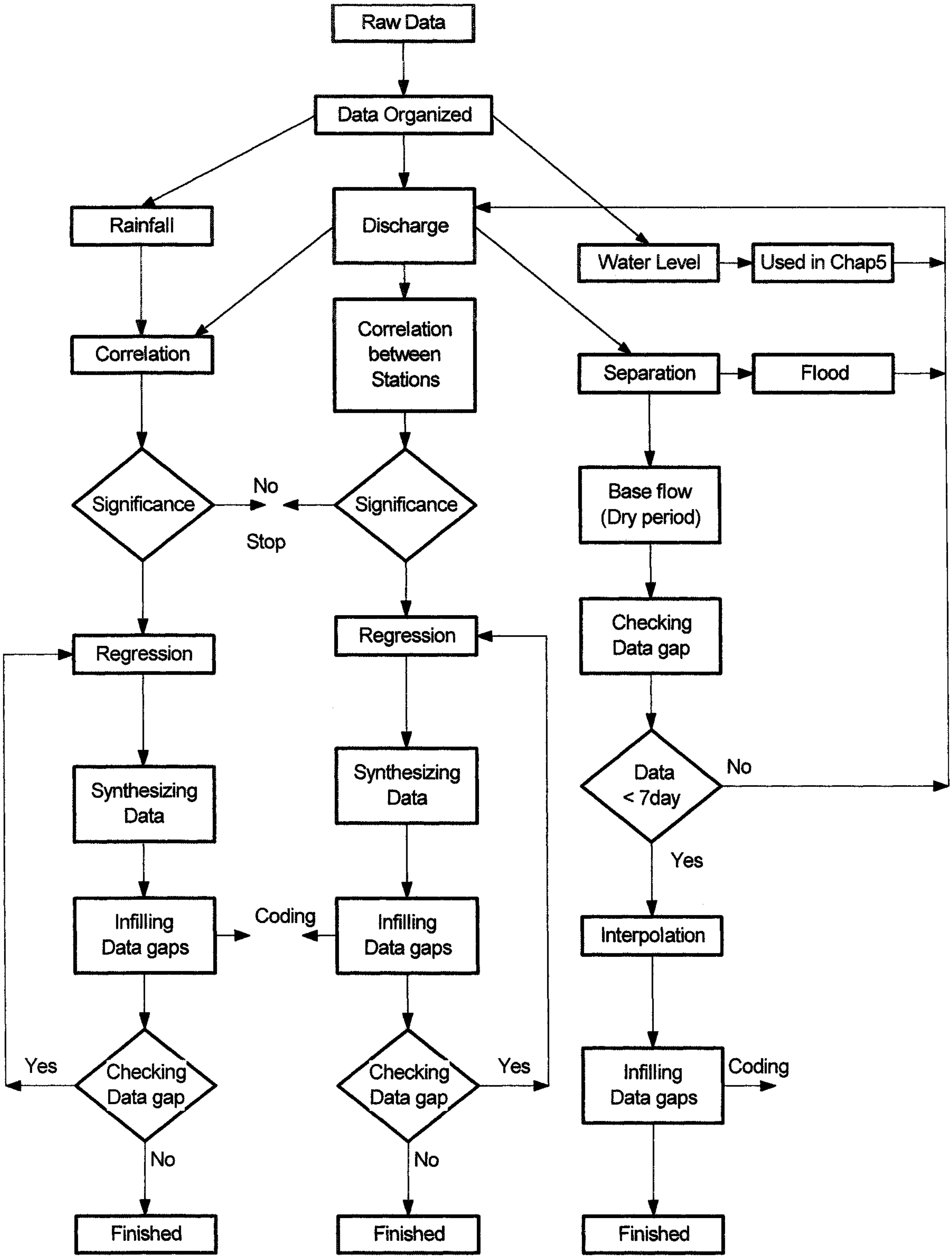

Figure 5.2: Flow diagram for infilling missing data

- The spreadsheet data was converted to dbase. The macros were (see appendix-B) developed and were run those for reformatting the data such a way that each column contains variable (station) and row contains the dates considering leap-year and months <31 days.

- The data was converted from dbase to spreadsheet and the data gaps were checked by macros (see appendix-B).

- All the steps were performed for each gauging and rainfall stations.

- In case of water level data, all the above steps and averaging were done where daily two observations (morning and evening) were recorded.

- Finally the data was organized in spreadsheet with each column contains one variable (discharge, water level, rainfall) and row with cases (dates).

5.7 Description of Missing Data

An exact description of missing data satisfying the varying the nature of its interpretations by various researchers is difficult. Because, the significance and complexity of a missing data varies according to time-scale usage of a time series. Further more, missing data of a variable may occur at on or more locations within a time series. At the station 2GC4 of Turasha river, for some days has two daily records (morning and afternoon) and the other stations of the catchment have daily one record.

Missing data was assigned in the time series when a single observation (data) in a day was not found in the discharge series. In the observation, it was found that all the stations of the catchment have missing data and it varies from couple days to years. The main station of the catchment 2GB1 of Malewa river has no records after September-1985 to 1990.

5.8 Correlation

Correlation analysis can be applied to determine the relationships or degree of association between the sets of variables. Correlation analysis is often a useful tool for the examination and selection of data that can be use in regression analysis.

In this study two types of correlation analysis were performed: (a) Correlation between rainfall and discharge; and (b) Correlation between different gauging stations

5.8.1 Correlation between rainfall and discharge

The heterogeneous topography in the mountainous parts of a watershed results in an equally heterogeneous rainfall pattern (gray, 1970). A correlation between the ten daily average rainfall and ten daily average discharge was made. It was observed that there was almost no correlation between rainfall and discharge due to the rainfall stations are being faraway from the catchment which was discussed in chapter-3.

5.8.2 Correlation between different gauging stations

The results of the correlation analysis of the discharge series among the eight gauging stations within the catchment are shown in table 5.1. It can be observed that there is a good correlation among the stations. The maximum correlation can be seen between the 2GB1 and 2GB5 station (r^2=0.85); 2GB1 and 2GC4 stations (r^2=0.81) which is expected as these two stations (2GB5 and 2GC4) are the main feeder of 2GB1 station and very close to 2GB1 station are shown in figure - 5.1. The most of the gauging stations show the good correlation with 2GB1 station except 2GB3 station (r^2=0.44). The station 2GB3 also shows the less correlation with all other stations. It can be explained that the sub-catchment of 2GB3 station is very small (53 km^2) compare to other sub-catchments, that is, situated in the upper part of the catchment which is far away from 2GB1 station. So, the rainfall is frequent compare to the whole catchment. So, the lowest correlation is found in 2GB3 station with other stations. It can be observed that the stations within a sub-catchment show higher correlation compare to inter sub-catchments.

Table - 5.1: Correlation coefficients between different gauging stations

	2GB1	2GB3	2GB4	2GB5	2GB7	2GC4	2GC5	2GC7
2GB1	1							
2GB3	0.44	1						
2GB4	0.80	0.42	1					
2GB5	0.81	0.39	0.74	1				
2GB7	0.73	0.32	0.68	0.76	1			
2GC4	0.85	0.20	0.67	0.58	0.50	1		
2GC5	0.76	0.13	0.57	0.53	0.59	0.64	1	
2GC7	0.63	0.20	0.55	0.41	0.51	0.65	0.64	1

5.9 Linear Interpolation

The linear interpolation technique was applied for the infilling the missing data. The discharge series was separated as "base flow" and "flood" depending upon water level. The "base flow" data gaps which are less than seven days was infilled by this technique.

5.10 Multiple Linear Regression

Multiple linear regression techniques are used to identify the mathematical dependence between the observed values of physically related variables and thus account for the additional information contained in correlated sequences of events (Viessman, 1996). In this case, sampling errors are reduced and the reliability of estimate is improved. This technique was applied to test the combined effects of the different independent variables on the dependent variable. In this procedure, any variables suspected to effect the dependent variable "Y" was included in the analysis. For "k" independent variables, X_1, X_2,........X_k the functional form of the multiple linear regression model is:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + \beta_k X_k \qquad 5.1$$

where the βi's are the partial regression coefficients associated with each Xi and α is the interception of the line on the Y-axis (the value of Y when all the Xi's have zero values).

The estimated regression equation is:

$$Y_{(est.)} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \qquad + \beta_k X_k \qquad (5.2)$$

The estimation procedures are outlined in a detailed way in appendix - C.

5.10.1 Used variables

In order to study the dependency of discharge series of different gauging stations, the multiple linear regression technique was used. The daily discharge series of gauging station 2GB1 as dependent variable was correlated with the rest of the gauging station as independent variables. In order to avoid the biasness during the analysis the independent gauging stations were taken from both sub-catchment each time.

5.11 Infilling of Missing Data and Coding

In order to infill the missing data for the station 2GB1, the following steps were done:

- The data which are available from the 2GB1 station either the original value calculated from the water level or infilled by interpolation technique has been identified.
- For each available value, the code was assigned as original "O".

- The regression analysis was done to obtain the regression parameters (α, β_1, β_2) using the discharge series which are highly correlated for the first step (2GB5 and 2GC4).

- The estimated regression equation was made with those coefficients.

- The synthesized discharge series was calculated when both the independent discharge series data were available.

- The data gaps of 2GB1 station was infilled with synthesized data and coding was made for each data such as "1" which was infilled.

- The remains' data gaps were checked for infilling.

- Again the regression analysis was done to obtain the regression parameters using the remaining discharge series which are higher correlated with 2GB1 station and so on.

- The remaining data gaps were infilled with synthesized data and separate coding was made for each data for each time.

- This procedure continues till the all data gaps were infilled as shown in figure - 5.2.

5.12 Results and Discussions

A number of multiple linear regression have been applied to infill the missing data of the station 2GB1 for the years 1960-1990. The depend variables (gauging stations) were selected depending on the degree of correlation with the 2GB1 station, standard deviation, position of the gauging station and availability of data where the missing data presents in the discharge series of the 2GB1 station. The results of infilling missing data are shown in table 5.2.

5.12.1 Data gaps infilled using the 2GB5 and 2GC4 stations

It can be observed that during the multiple linear regression analysis, the maximum correlation ($R^2 = 0.786$) can be obtained by using above mentioned gauging stations. The calculated standard error of these stations is minimum ($\sigma = 5.29$). The position of these two stations is very close to the 2GB1 station (see figure 5.1). These two gauging stations are more suitable to infill the missing data. Therefore, major portion (12.29%) of missing data (total data gaps in discharge series = 22.50%) was infilled by using these two gauging stations (see table 5.2).

5.12.2 Data gaps infilled using the 2GB7 and 2GC4 stations

It can be observed that the degree of correlation in the multiple regression analysis by using these two stations slightly decreases (R^2 = 0.741) and the standard error increases (σ = 5.36). It can be explained that due to location of the 2GB7 station which is little far away from the 2GB1 station (see figure 5.1). The second major portion of missing data (5.26%) was infilled using these stations.

5.12.3 Data gaps infilled using the 2GB5 and 2GC7 stations

It can be observed that the multiple regression analysis using these two stations, decreases the degree of correlation (R^2 = 0.622) and increases the standard error (σ=5.696). The 2GC7 station is located far away from the 2GB1 station (see figure 5.1). Using these two gauging stations only "1.02%" of missing data was infilled.

5.12.4 Data gaps infilled using the 2GB4 and 2GC5 stations

It can be observed that the above mentioned gauging stations are far away from gauging station the 2GB1 (see figure 5.1). The degree of correlation (R^2 = 0.628) decreases and the standard error of these stations is maximum (σ = 7.233). Using these two stations "3.26 %" of missing data was infilled due to availability of data.

5.12.5 Data gaps infilled using the "2GB3 + 2GC5" and "2GB5 + 2GB7" stations

Due to the location of the 2GB3 station and the 2GC5 station, the degree of correlation is very low (R^2 = 0.44). The degree of correlation using the 2GB5 and 2GB7 stations is higher (R^2 = 0.61) than the "2GB3 + 2GC5" stations as they are into the same sub-catchment (see figure 5.1). The standard deviation is almost same of the other stations (σ=5.23 and σ=5.69 respectively). Only 0.67% missing data was infilled using these stations.

Table 5.2: Summery results of infilling missing data

Iteration No	Independent Station	Multiple R	Square of Multiple R	Standard error	Code	Percent Infilled
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	Original	-	-	-	0	77.50
2	2GB5 and 2GC4	0.887	0.786	5.2880	1	12.29
3	2GB7 and 2GC4	0.861	0.741	5.3620	2	5.26
4	2GB5 and 2GC7	0.789	0.622	5.6960	3	1.02
5	2GB4 and 2GC5	0.792	0.628	7.2331	4	3.26
6	2GB3 and 2GC5	0.656	0.439	5.2323	5	0.63
7	2GB5 and 2GB7	0.779	0.607	5.6899	6	0.04

In order to compare the predicted discharge series with measured discharge series, the original discharge series of the 2GB1 station, predicted discharge series which was synthesis from other gauging station and the infilled data series were plotted together. The discharge series were plotted where the measured data discharge series of the 2GB1 station was available for comparison. The missing data in the discharge series during the years 1960 - 1980 was infilled using the 2GB5 and 2GC4 stations. After this period, the missing data were infilled by all the above mentioned gauging stations depending on availability of data series and above mention criteria (see section 5.12)

It can be observed that during the years 1960 - 1973, the predicted discharge series are very close to measured discharge series (see figure 5.3a, 5.3b, 5.3c and 5.3d). The number and length of missing data in the discharge series during this period are smaller. After that the predicted discharge series deviate from measured discharge (see figure 5.3e, 5.3f and 5.3g) and the length data gaps are larger. The discharge data series seems to be good quality during the years 1960-1973.

During the year 1981, the missing data in the discharge series was infilled partially using the stations "2GB5 + 2GC4", "2GB4 + 2GC5" and "2GB5 + 2GC7" are shown in figure 5.3g, 5.4 and 5.5 respectively. It can be observed that measured discharge series sometime deviates from predicted series. Figure 5.6 and figure 5.7 show the partially infilled discharge series by the station "2GB7 + 2GC4" and "2GB5 + 2GC4" for the year 1983 and 1985 respectively. During the year 1984 the measured discharge data is not enough for comparison. The measured discharge data of the 2GB1 station are not available for comparison after the year 1985.

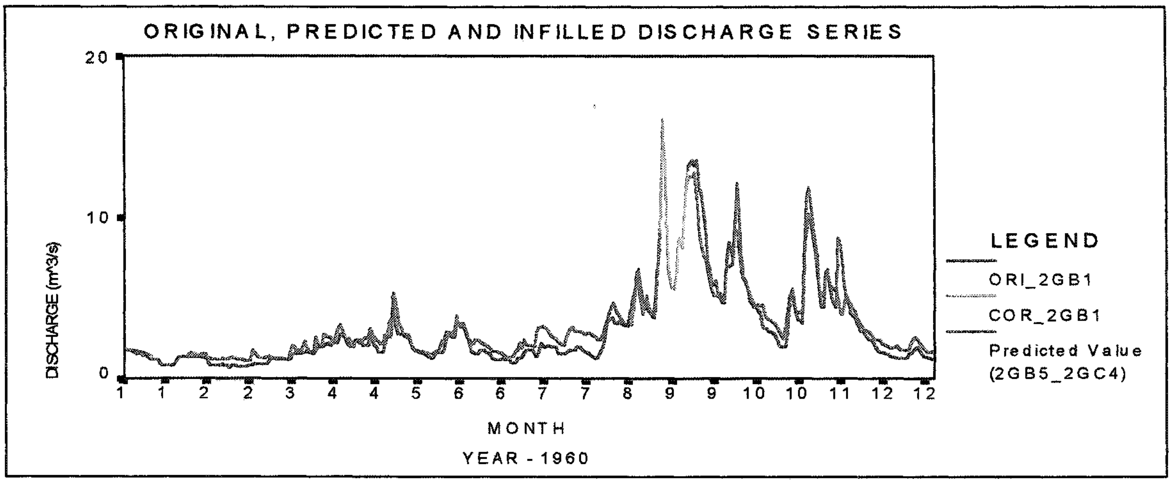

Figure 5.3a: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1960

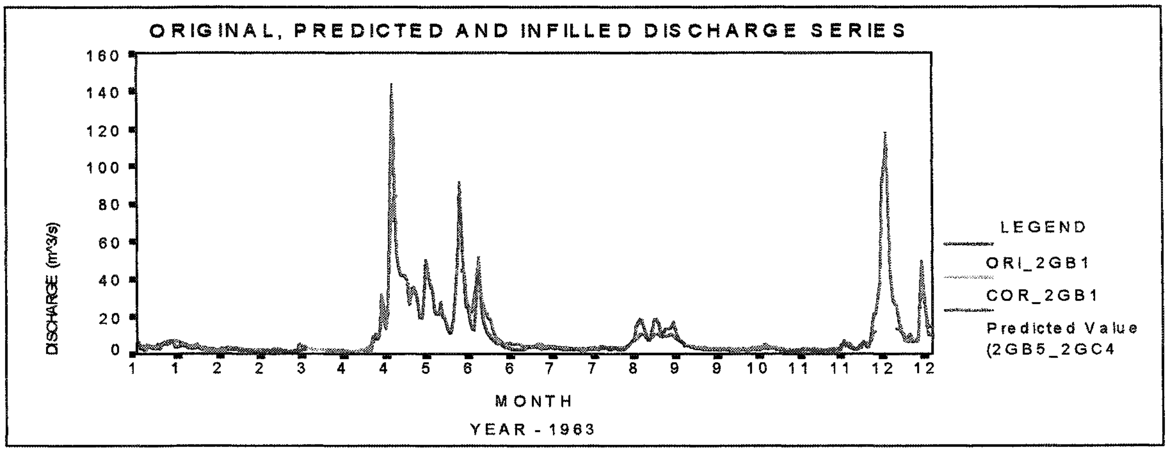

Figure 5.3b: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1963

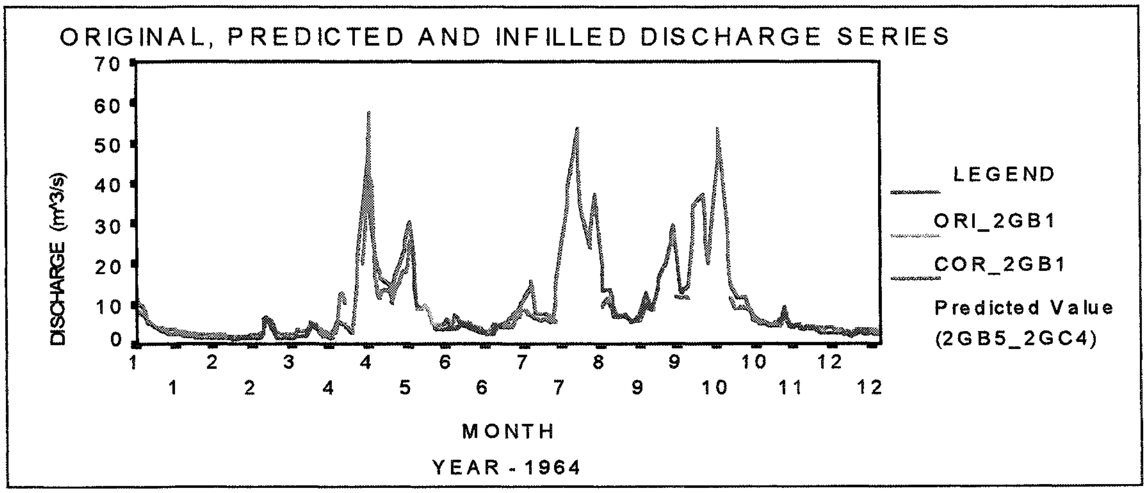

Figure 5.3c: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1964

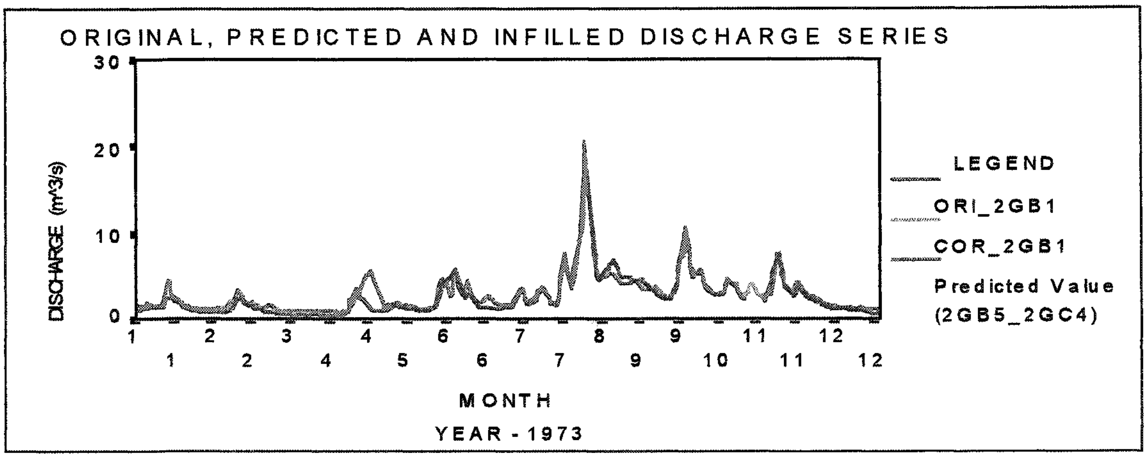

Figure 5.3d: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1973

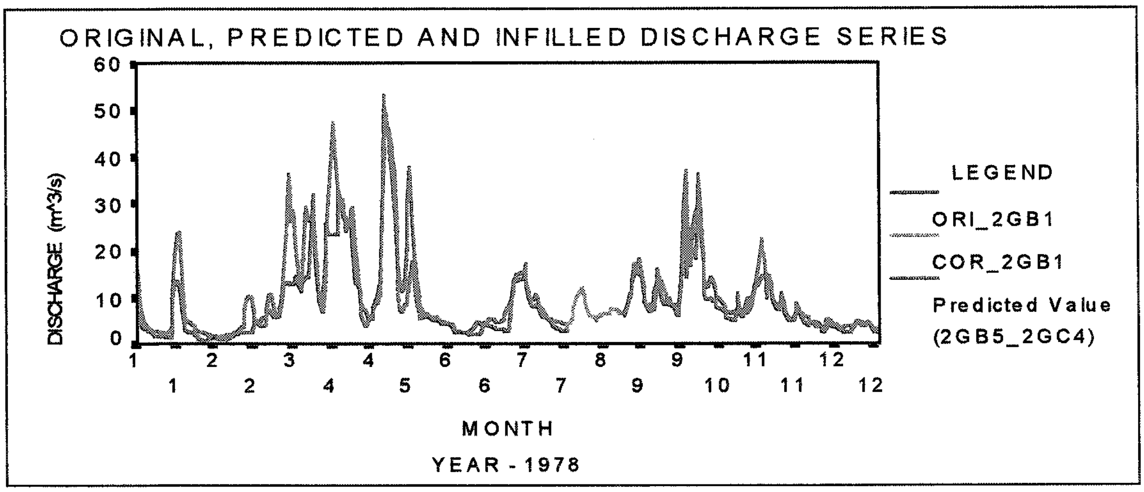

Figure 5.3e: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1978

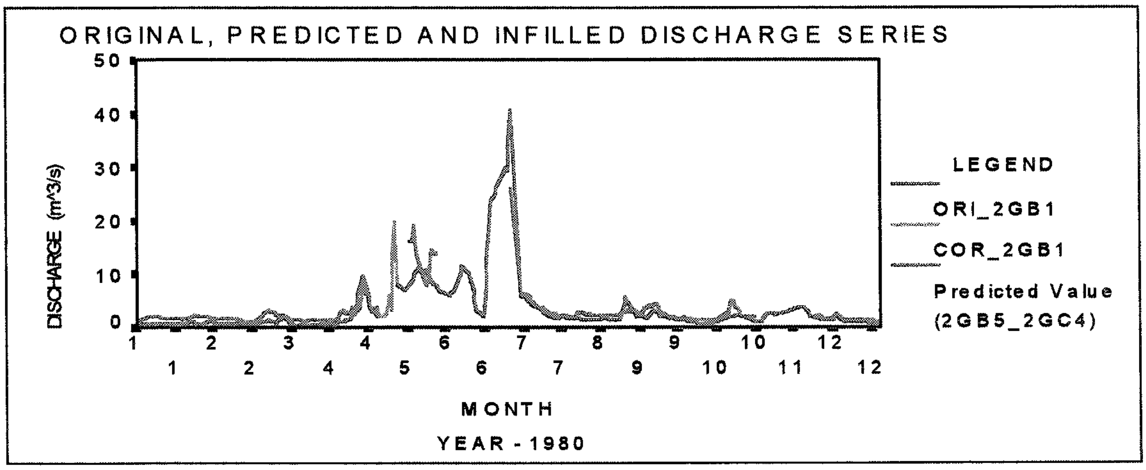

Figure 5.3f: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1980

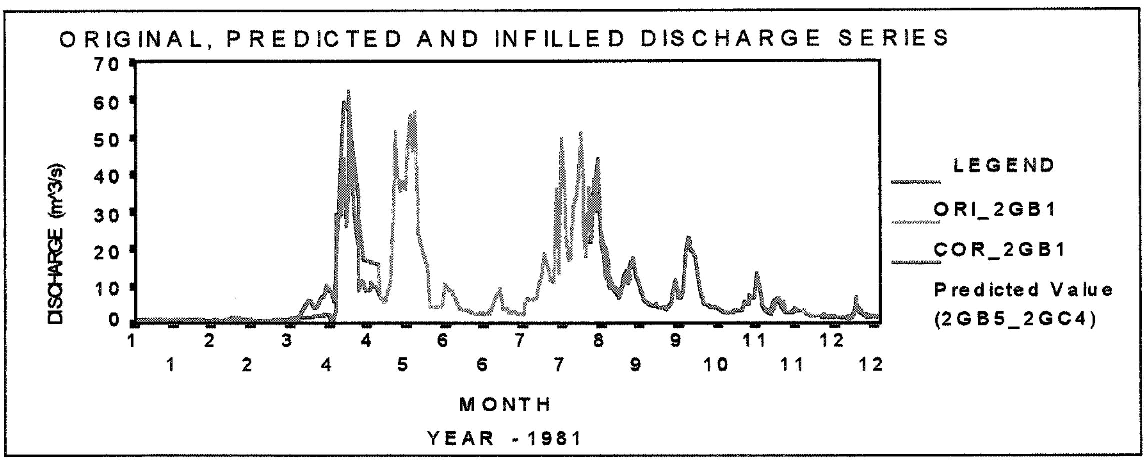

Figure 5.3g: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1981

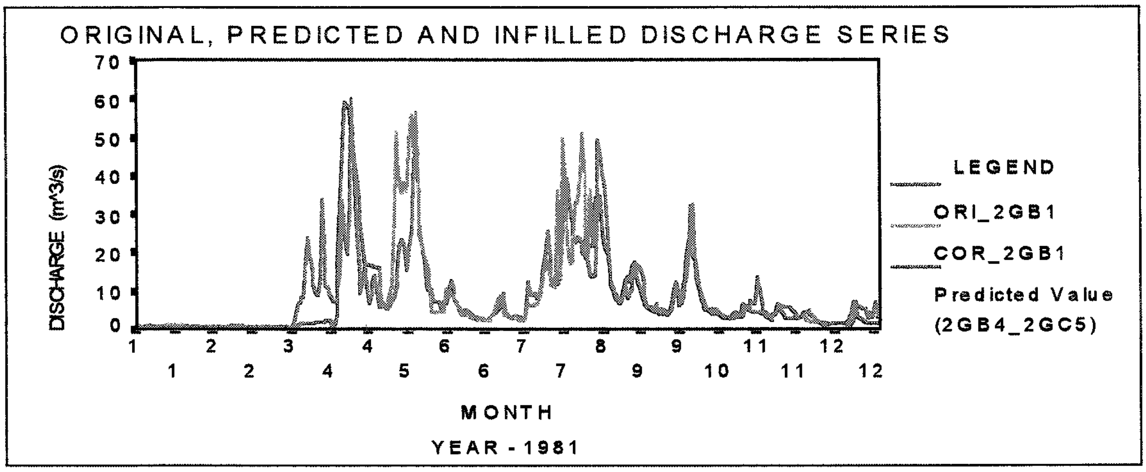

Figure 5.4: Measured, Predicted (2GB4 and 2GC5) and Infilled discharge series for the year-1981

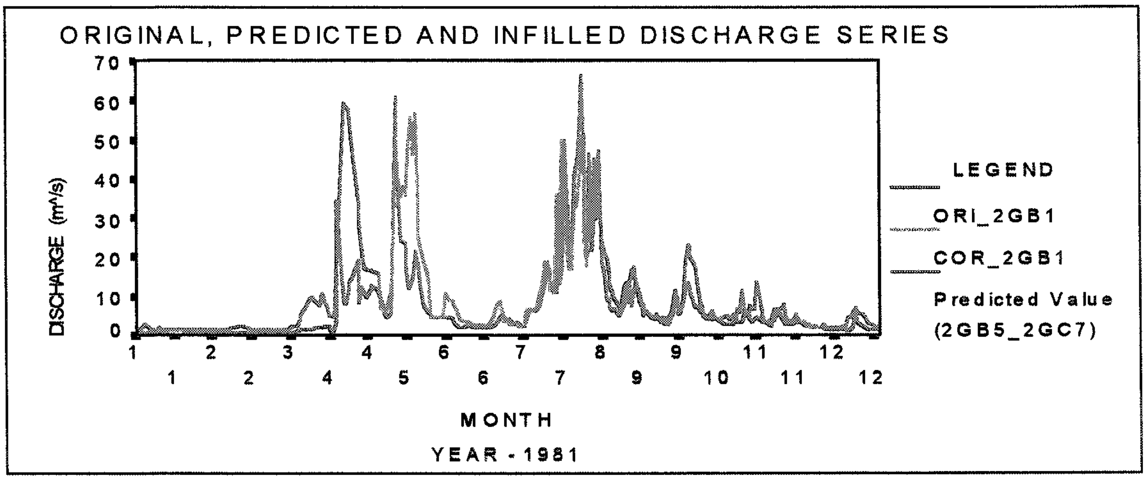

Figure 5.5: Measured, Predicted (2GB5 and 2GC7) and Infilled discharge series for the year-1981

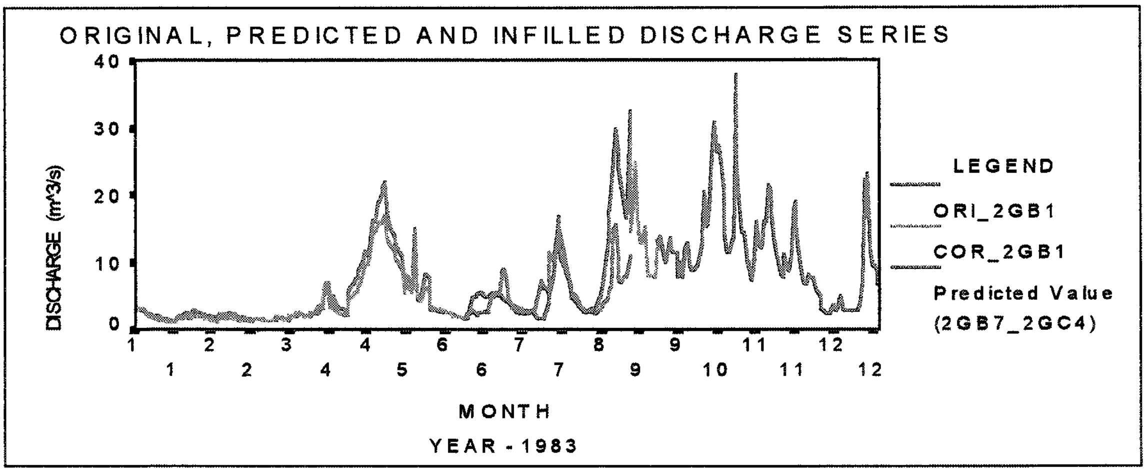

Figure 5.6: Measured, Predicted (2GB7 and 2GC4) and Infilled discharge series for the year-1983

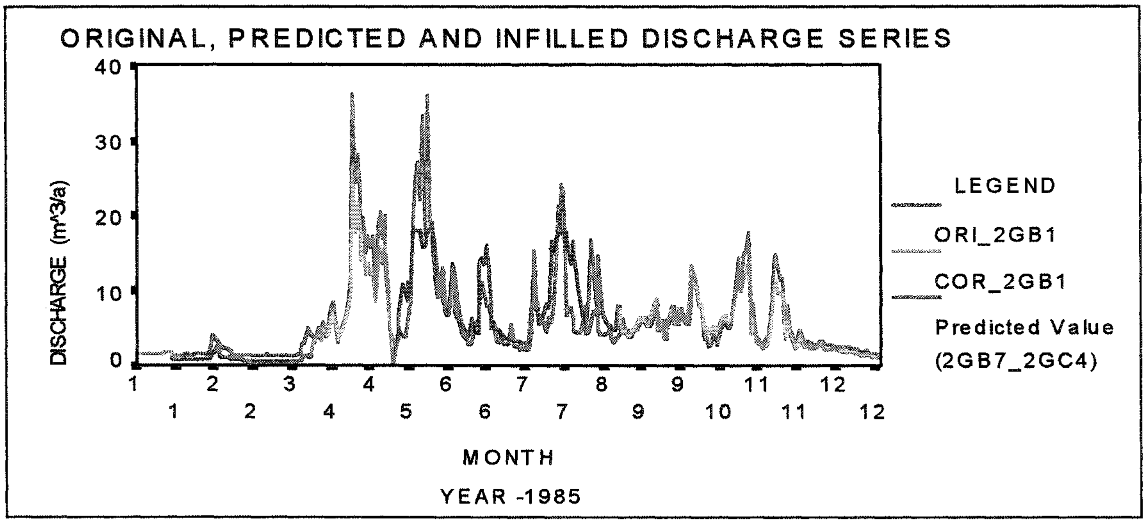

Figure 5.7: Measured, Predicted (2GB5 and 2GC4) and Infilled discharge series for the year-1985

5.13 Conclusions

- Most of the gauging stations were shown a good correlation with 2GB1 station. The maximum correlation observed between the 2GC4 and 2GB1 stations (R^2=0.85) for the Turasha sub catchment. The minimum correlation was found with the 2GB3 station (R^2=0.44) which is located far away from the 2GB1 station.

- The linear interpolation technique was applied where the data gaps are less than seven days during the base flow period.

- Most of the data gaps were infilled using the multiple linear regression analysis. The major portion (12.29% out of 22.50%) data gaps were infilled using the 2GC4 and 2GB5 stations. The second major portion of data gaps (5.26%) were infilled using the 2GB7 and 2GC4 stations. The remaining (4.93%) data gaps were infilled using other stations.

CHAPTER SIX

THE ERRORS IN DETERMINING DISCHARGE

6.1 Introduction

The accuracy or, more correctly, the error of a measurement of discharge may be defined as the difference between the measured flow and the true value. The true value of the flow is unknown and can only be ascertained (within close limits) by weighing or by volumetric measurement. An estimate of true value has therefore to be made by calculating the uncertainty in the measurement within confidence (95%) level. It should be stressed that the statistical analysis of river flow data is only applicable if the field data have been obtained by acceptable hydrometric principles and practices. Statistical analysis is an aid to improve the presentation of the hydrometric data for the user's benefit but the final quality of data depends on the hydrologist.

6.2 Nature of Errors

Errors of observation are usually grouped as random (or stochastic), systematic and spurious.

6.2.1 Randon errors

Random errors are sometimes referred to as experimental errors and the observations deviate from the in accordance with the law of chance such that the distribution usually approaches a normal distribution. They are the most important errors to be considered in stream flow (see figure 6.1).

6.2.2 Systematic errors

Systematic errors are those which can not be reduced by increasing the number of observations if the instruments and equipment remain unchanged. In stream flow, systematic errors may be present in the water level recorder, in the reference gauge or datum, and in the current meter. These errors may be generally small but some cases their effect may cause a systematic error in the stage-discharge relation which can serious effect on low values of discharge.

6.2.3 Spurious error

These are human errors or instrument malfunction and cannot be statistically analyzed. The observations must therefore be discarded (see figure 6.1).

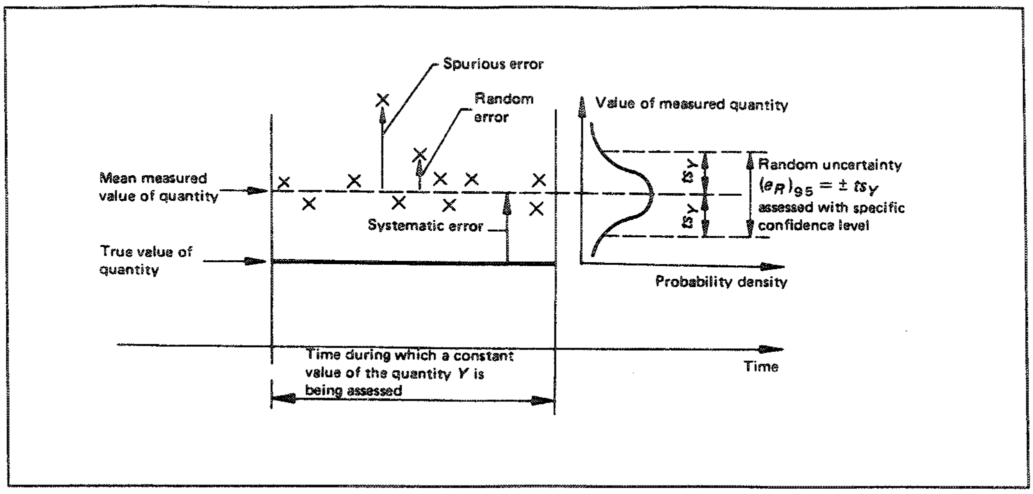

Figure 6.1: Errors in measurement

6.3 Source of Eror for Single Discharge Measurement

6.3.1 Equipment

Accurate measurement requires that measurement equipment be properly assembled and maintained in good conditions. Current meters are especially susceptible to damage when in use, as measurement must often be made when drift or floating ice is present in a stream.

6.3.2 Characteristics of measurement section

The basic characteristics of the measurement section affect measurement accuracy. If possible, the section should be deep enough to permit use of the two-point method of measuring velocity. The presence of bridge piers in or near the measurement section adversely affects the distribution of velocity. Piers also tend to induce local bed scour which affects the uniformity of depth.

6.3.3 Spacing of observation verticals

The spacing of observation verticals in the measurement section can affect the accuracy of the measurement. Twenty-five to thirty verticals should normally be used, and the verticals should be placed so that each segment will have approximately equal discharge (Operation Hydrology, Report no. 13, WMO- No. 519). However, a measurement vertical should be located fairly close to each bank and at "breaks" in depth.

6.3.4 Rapidly changing stage

When the stage changes rapidly during a discharge measurement, the computed discharge figure losses some of its reliability and there is uncertainty as to appropriate gauge height to apply to that discharge figure. Consequently, the standard procedure for making discharge measurements should be shorten when the stage is changing rapidly.

6.3.5 Measurement of depth and velocity

Inaccuracies in sounding (for automatic) or vertical position of staff and in the placement of the current meter are most likely to occur in those sections having great depth and velocities. Heavy sounding weights should be used to reduce the vertical angle made by sounding line and correction tables should be used in determining vertical distance. Where velocities are not perpendicular to the measurement section, correction should be made by measuring the cosine of the angle between the perpendicular and the direction of current.

6.3.6 Wind

Wind may affect the accuracy of a discharge measurement by obscuring the angle of the current, by creating waves that make it difficult to sense the water surface prior to sounding the depth, and by affecting the 0.2 depth velocity observations in shallow depths.

6.3.7 Ice

Reliable measurements may usually be made when measuring from ice cover if the measurement verticals are free of slush ice. Slush ice interferes with the operation of current meter rotor moreover causes difficulty in determining the effective depth of water.

6.4 Theory of Propagation of Errors

If a quantity Q is a function of several measured quantities X, Y, Z, ..., the error in Q due to errors $\delta x, \delta y, \delta z, \ldots$ in X, Y, Z, ... , respectively, is given by

$$\delta Q = \frac{\partial Q}{\partial x}\delta x + \frac{\partial Q}{\partial y}\delta y + \frac{\partial Q}{\partial z}\delta z + \qquad (6.1)$$

The first term in equation (1), $(\partial Q/\partial x)\delta x$, is the error in Q due to an error δx in X only (i.e. corresponding to $\delta y, \delta z, \ldots$, all being zero). Similarly the second term $(\partial Q/\partial y)\delta y$, is the error in Q due to an error δy in Y only. Squaring gives

$$\delta Q^2 = \left(\frac{\partial Q}{\partial x}\delta x\right)^2 + 2\frac{\partial Q}{\partial x}\frac{\partial Q}{\partial y}\delta x\delta y + \left(\frac{\partial Q}{\partial y}\delta y\right)^2 + \qquad (6.2)$$

Now the terms $(\partial Q/\partial x)(\partial Q/\partial y)\delta x\delta y$, etc., are covariance terms and, since they contain quantities which are as equally likely to be positive or negative, their algebraic sum may be conveniently taken as being either zero or else negligible as compared with the squared terms.

$$\delta Q^2 = (\frac{\partial Q}{\partial x}\delta x)^2 + (\frac{\partial Q}{\partial y}\delta y)^2 + (\frac{\partial Q}{\partial z}\delta z) + \qquad (6.3)$$

i.e. the error in Q, δQ is the sum of the squares of the errors due to an error in each variable. Now

$$\frac{\partial Q}{\partial x} = yz, \quad \frac{\partial Q}{\partial y} = xz, \quad \frac{\partial Q}{\partial z} = xy \qquad (6.4)$$

and

$$\delta Q = [(yz\delta x)^2 + (xz\delta y)^2 + (xy\delta z)^2 + \quad]^{1/2} \qquad (6.5)$$

Diving by Q = xyz

$$\frac{\partial Q}{Q} = [(\frac{yz}{xyz}\delta x)^2 + (\frac{xz}{xyz}\delta y)^2 + (\frac{xy}{xyz}\delta z)^2 + \quad]^{1/2} \qquad (6.6)$$

and

$$\frac{\partial Q}{Q} = [(\frac{\delta x}{x})^2 + (\frac{\delta y}{y})^2 + (\frac{\delta z}{z})^2 + \quad]^{1/2} \qquad (6.7)$$

where $(\partial x)/x$, $(\partial y)/y$, and $(\partial z)/z$ are fractional values of the errors (standard deviations) in X, Y and Z, and if they are each multiplied by 100 they become percentage standard deviations. Let X_Q be the percentage standard deviation of Q and

X_X = percentage standard deviation of x
X_y = percentage standard deviation of y
X_z = percentage standard deviation of z

then

$$X_Q = \pm (X_X^2 + X_Y^2 + X_Z^2 + \ldots)^{1/2} \qquad (6.8)$$

which is generally referred to as the root-sum-square equation for the estimation of uncertainties.

6.5 Analysis the Magnitude of Errors

An error analysis shows that three types of errors influence the random error of a single discharge measurement from a rating curve. They are (a) Rating curve error, (b) Water level measurement error, and (c) Error caused by ignoring all physical parameters, other than water level, that affect discharge. In this study, rating curve errors of the gauging stations were calculated in order to estimate of inflow of Malewa catchment.

6.5.1 The uncertainty in the stage-discharge relation

The equation for the stage-discharge relation may be expressed in the general form

$$Q = C\,(h + a)^{n} \qquad (6.9)$$

Where Q is the discharge, C is a coefficient, h is the stage, a is the datum correction denoting the value of stage at zero flow and n is an exponent.

In order to estimate the uncertainty in the relation, equation (6.9) is linearized by a logarithmic transformation of the form

$$\ln Q = \ln C + n \ln (h + a) \qquad (6.10)$$

The procedure is then one of estimating S_e, the standard error of estimate

S_e is calculated from

$$S_e = \left[\frac{\sum (LnQi - LnQc)^2}{N-2}\right]^{1/2} \qquad (6.11)$$

Where Q_i is the current meter observation and Q_c is the discharge taken from the rating curve corresponding to Q_i and (h + a), where $Q_c = C\,(h + a)^{n}$ and N is the number of gauging.

If the stage-discharge relation comprises one or more break points, S_e should be calculated for each segment and (N-2) degrees of freedom are allowed for each segment. At least 20 current meter observations should be available in each range before a statistically acceptable estimate can be made of S_e.

Therefore, 95% current meter observations, on average, will be contain $= t \times S_e \times 100$

Where, t is Student's t correction for the sample size at the 95% confidence level for N gauging and it is 1.96.

The uncertainty of the main gauging stations (see details in appendix-D) are shown in table 6.1.

Table 6.1: Error uncertainty of the main gauging stations

Sl. No.	Station	No. of Observation	$\Sigma(\text{Ln Qi - Ln Qc})^2$	S_e	t×Se×100
1	2	3	4	5	6
1	2GB1	25	0.395433	0.1311211	25.7
2	$2GB5_{(lower)}$	45	5.230532	0.34875	68.35
3	$2GB5_{(upper)}$	14	0.401298	0.18287	35.84
4	$2GC4_{(lower)}$	28	1.883583	0.26910	52.7
5	$2GC4_{(upper)}$	3	0.017767	0.13329	26.12

It can be observed that the uncertainty of the rating curve of the 2GB1 station is the lowest out of three gauging stations and it is highest for lower part of the 2GB5 station. The estimated uncertainty of high flow part of the 2GB5 station is seems to be much lower than low flow part of the same station but the total magnitude would be higher.

The number of observations were quite enough to estimate the uncertainty of the rating curves of the 2GB1 station and the lower part of the 2GB5 and 2GC4 stations. The number of observations is lower than required according to statistical point of view for the upper part of the 2GB5 station but it can be accepted as the high flow measurements are not so easy. Only three measurements were available for the high flow part of the 2GC4 station. So the estimated uncertainty for this part is not reliable.

6.6 Analysis the Magnitude of Inflow and the Uncertainty

The main interest of the study is to estimate the inflow from the catchment into the lake at the 2GB1 station and the uncertainty of inflow. The total yearly discharge and the percentage of yearly uncertainty were calculated from daily discharges.

6.6.1 Total yearly discharge

The discharge series of the 2GB1 station is the combination two discharge series. These are, the discharge series calculated from the rating curve and the infilled series. The total yearly discharges were calculated by adding the daily discharge as shown in table 6.2.

6.6.2 The yearly uncertainty

The uncertainty of discharge series is the combination two uncertainties. The uncertainty of the 2GB1 station calculated from rating curve and the uncertainty of the infilled series. The total yearly uncertainty was estimated as:

$$X_Q = \left[\frac{\sum (Qi)^2 (Xi)^2}{(\sum Qi)^2}\right]^{1/2}$$

Where X_Q = overall yearly uncertainty in discharge
Qi = daily discharge
Xi = uncertainty in daily discharge

The uncertainty for infilled series

The uncertainty of the infilled series was estimated as:

$$X_O = (X_Q^2 + X_P^2)^{1/2}$$

Where, X_O = Overall uncertainty of infilled series.

X_Q = estimated rating curve error at the 95% confidence level.

X_P = estimated standard error of the mean predicted value at the 95% confidence level.

6.7 Results and Discussions

The total yearly discharge, yearly uncertainty and percentage of yearly infilled data are shown in table-6.2 and the yearly discharge and errors were plotted are shown in figure-6.2. It can be observed that large variations in the yearly discharges. The maximum yearly inflow observed during the year-1961 (4424×86400 m^3±3.14%) and the minimum inflow during the year-1984 (613×86400 m^3±1.89%). The average yearly inflow from the catchment was (2486×864000 m^3±2.26%) during the years 1960-1990 (see table-6.2). The yearly infilled of missing data during the years 1960-1976 are not significant (≤5%)

Observing the graph (Figure-6.2), it seems to be a hydrological changes in the catchment. During the years 1961-1968, the yearly inflow was quite higher. After the year 1968, the yearly inflow decreases, and it continues till the year-1987. During the year-1988 the catchment repeats it's higher magnitude of flow.

The changes of yearly uncertainty is not significant. The Maximum yearly uncertainty is "3.14%" and minimum yearly uncertainty is 1.80%. The average yearly uncertainty is 2.26%. It has been observed that during the high flow years, the uncertainties are higher and during the low flow years the uncertainties are lower (see figure-6.2)

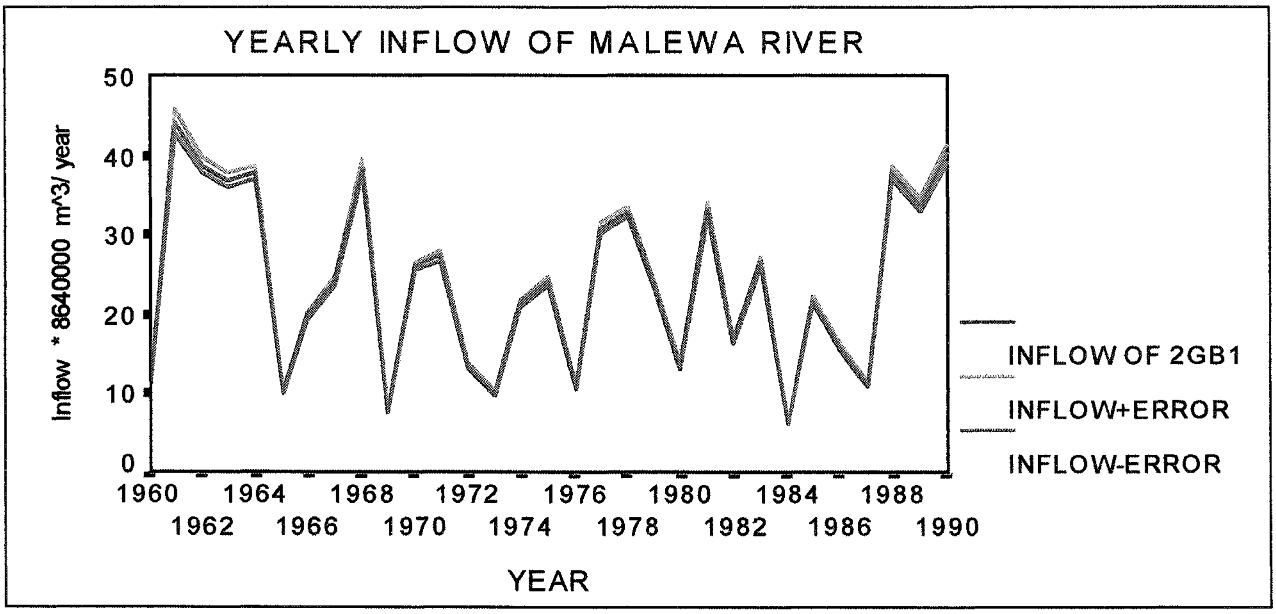

Figure-6.2: Yearly inflow and errors

6.6 Conclusion

- The uncertainty of the 2GB1 station was the lowest (25.7% at the 95% confidence level) of the three gauging stations. The uncertainty of the lower part of the rating curves was higher [$2GB5_{(lower)}$ =0 68.35% & $2GC4_{(lower)}$ =0 52.7%] compared to the upper part of the rating curves [$2GB5_{(upper)}$ = 35.84% & $2GC4_{(upper)}$ =0 26.12%].

- The variation of yearly uncertainty of inflow of the Malewa river is not significant. The yearly uncertainty varies from 3.14% to 1.80%. The average yearly uncertainty was 2.26% during the years 1960 - 1990.

- The variation of yearly inflow of the Malewa river is significant. The maximum inflow occurred (4424×86400 m^3 ±3.14%) during the year 1961 and minimum inflow occurred (613×86400 m^3 ± 1.89%) during the year-1984. The average yearly inflow was "2486×86400 m^3±2.26%" during the years' 19960-1990.

Table-6.2: Results of yearly inflow, yearly uncertainty and yearly infilled data (in percentage)

Total Volume=Un it*86400				Percent data infilled					
Year	Vol. M^3	% error	%2GB1	%(2gb5+2gc4)	%(2gb7+2gc4)	%(2gb5+2gc7)	%(2gb4+2gc5)	%(2gb3+2gc5)	%(2gb5+2gB7)
1	2	3	4	5	6	7	8	9	10
1960	1128	1.83	96.44	3.56					
1961	4424	3.14	100.00						
1962	3878	2.31	100.00						
1963	3683	2.70	95.07	4.93					
1964	3780	1.98	97.81	2.19					
1965	1015	1.82	100.00						
1966	1976	2.25	100.00						
1967	2421	2.33	100.00						
1968	3844	2.51	100.00						
1969	776	2.01	100.00						
1970	2599	1.88	100.00						
1971	2735	2.25	100.00						
1972	1341	1.93	100.00						
1973	1000	1.84	97.26	2.74					
1974	2134	1.95	100.00						
1975	2424	2.31	100.00						
1976	1069	2.22	97.26	2.74					
1977	3079	2.10	82.19	17.81					
1978	3293	1.82	92.33	7.67					
1979	2327	1.99	93.42	6.58					
1980	1337	2.51	97.54	2.46					
1981	3321	2.43	66.02	21.92	0	11.51	0.55	0	0
1982	1663	1.80	97.26	2.47	0	0	0	0.27	0
1983	2658	1.90	51.78	34.25	10.14	0	2.19	0.27	1.37
1984	613	1.89	88.25	0	2.19	4.37	3.28	1.91	0
1985	2185	1.84	49.32	35.62	10.14	4.93	0	0	0
1986	1582	2.66	0.00	99.45	0.55	0	0	0	0
1987	1101	1.92	0.00	90.41	0.55	9.04	0	0	0
1988	3785	2.49	0.00	46.72	34.7	1.91	16.67	0	0
1989	3379	2.64	0.00	0	40.55	0	50.68	8.77	0
1990	4039	2.60	0.00	0	63.56	0	28.22	8.22	0

Average = 2486 2.26

SUMMARY AND CONCLUSIONS

The Malewa catchment is a mountainous sub catchment of Naivasha basin, situated in the Rift Valley Province, Kenya. It has an area of 1428 km^2 and the main source of fresh water to the lake Naivasha. A tropical climate and relative cool conditions are experienced in the study area. The lake Naivasha area is influenced by rainshadow from the surrounding high lands.

Based on the results of the study, the following general conclusions were obtained:

1. The rainfall and discharge records are consistent. The spatial distribution of rainfall within the catchment is not uniform.

2. There is a significant relationship between the average rainfall and altitude (correlation coefficient R^2 = 0.77). The annual precipitation is more in the higher elevation.

3. The available rainfall stations are far away from the catchment and almost nocorrelation between the rainfall and stream flow (correlation coefficient R^2_{max}=0.50 & R^2_{min}=0). These rainfall stations do not represent the aereial rainfall to the catchmen. So that the rainfall data were not used to predict inflow of the catchment.

4. Most of the gauging stations have significant correlation with 2GB1 station. The maximum correlation is shown with the 2GC4 station (correlation coefficient R^2 = 0.85) and with the 2GB5 station (correlation coefficient R^2=0.81). There is minimum correlation observed between the gauging stations 2GB3 and 2GB1 (correlation coefficient R^2=0.44), because the gauging station 2GB3 is located far away from the gauging station 2GB1.

5. Most of the data gaps were infilled by using a multiple linear regression technique. To avoid the biasness, the gauging stations were selected from both sub catchments of the Malewa river. The major portion (12.29% out of 22.50%) of the data gaps was infilled using "2GB5 + 2GC4" stations. The second major portion (5.26%) of the data gaps were infilled using "2GB7+2GC4" stations. The remaining (4.93%) data gaps were infilled using other gauging stations of the catchment.

6. The quality of discharge data of three main gauging stations are better during the years 1960-1970.

7. The rating curves of main three gauging stations were reconstructed and it was found that the reconstructed rating curves significantly differ from supplied rating curves.

8. The optimization rating curve coefficients are possible for two rating curves by using three rating curves at the same time by the developed optimization procedure. The rating curve coefficients were optimized by assigning weight and without assigning weight. It was found that the optimization with assigning weight produced reliable rating coefficient for low flow conditions and without assigning weight produced reliable rating coefficient for high flow conditions.

9. The random uncertainty of the rating curves for main three gauging stations were calculated. It was found that the uncertainty of 2GB1 station was the lowest (25.7% at the 95% confidence level) of the three gauging stations. The uncertainty of the lower part of the rating curves was higher [$2GB5_{(lower)}$ =0 68.35% & $2GC4_{(lower)}$ =0 52.7%] compared to the upper part of the rating curves [$2GB5_{(upper)}$ = 35.84% & $2GC4_{(upper)}$ =0 26.12%].

10. The variation of yearly uncertainty of inflow of the Malewa river is not significant. The yearly uncertainty varies from 3.14% to 1.80%. The average yearly uncertainty was 2.26% during the years 1960 - 1990.

11. The variation of yearly inflow of the Malewa river is significant. The maximum inflow occurred (4424×86400 m^3 ±3.14%) during the year 1961 and minimum inflow occurred (613×86400 m^3 ± 1.89%) during the year-1984. The average yearly inflow was "2486×86400 m^3±2.26%" during the years' 19960-1990.

RECOMMENDATIONS

For estimation of inflow, the hydrological data play the main role. As a result of lake of rainfall data in the catchment the infilling of data gaps based on discharge data. The installation of at least three meteorological station (Wanjohi already exits) within the catchment will be helpful for the further hydrological studies.

The infilled data gaps should be checked by rainfall-runoff model if possible by collecting the rainfall data from Wanjohi station.

The rating curves of the gauging stations were constructed long-ago. The rating curves should be updated.

The staff of the 2GB5 station should be replaced close to the measuring section as the shape of the measuring section significantly differs from staff locate section.

The barrier of measuring structure of the 2GB1 station should be removed so that rating curve could provide satisfactory result.

Some places of the study area faces severe erosion problem. The presence of rills and gullies are the evidence of erosion especially arable bare lands in the mountain slops and overgrazing areas. Strip cropping should be practiced in the slop of the mountains and controlled grazing and conservation practice should be maintained in the erosion prone areas.

REFERENCES:

Amritkar, R.E. and Kumar, P.P., 1995, Interpolation of missing data using nonlinear and chaotic system analysis, Journal of Geophysical Research, Vol. 100(D2), pp. 3149-3154.

Ausa, B. T.,1996, A study on river discharge variation of river, Malaga, Spain, ITC Msc Thesis, Enschede, The Netherlands.

Dymond, J.R. and Christian, R., 1982, Accuracy of discharge determind from the rating curve, Hydrological - Sciences - Journal, Vol. 27(4) pp. 493-504.

Elwan, M.Y., Shahin, M. and Hall, M.,T., 1996, Synthetic generation of missing runoff data for the White Nile, river Sobat and Bahr el-Jebel, Stochastic Hydraulics. Proc. Symposium, Mackey, Australia.

Gyau-Boakye, P., Schultz, G.A., 1994, Filling gaps in runoff time series in West Africa, Hydrological - Sciences - Journal, Vol. 39(6). pp. 621-636.

Reginald, W.H., 1995, Stream flow Measurement, Second Edition, E & FN Spon, London, UK.

Richardson, J.L., and Richardson, A.E., 1972, History of East Africa Rift Lake and its climatic implications, Ecological monograps, Vol. 42(4). pp. 499-534.

Sefe, F.T.K., 1996, A study of the stage-discharge relationship of the Okavango River at Mohembo, Botswana, Hydrological - Sciences - Journal, pp. 97-116.

Stuttard, M. and others, 1995, Final Report, Monitoring Lakes in Kenya: An Environmental Analysis Methodology for Developing Countries.

Viessman, W., Lewis, G.l., 1996, Introduction to Hydrology, Fourth Edition, HarperCollins College Publishers, New York.

Walton, T.L., Fill in missing data in unvariate coastal data, Journal of applied Statistics, Vol. 23(1), pp. 31-39.

Wonnacott, H.T., Wonnacott, R.J., 1972, Introductory Statistics, Second Edition, John Wiley & Sons. Inc., New York.

World Meteorological Organization, 1980, Operational Hydrology, Report No. 13, Volume I & Volume II, Geneva Switzerland.

APPENDICES

TABLE 3.1-A: MONTHLY RAINFALL

Monthly Rainfall (mm) **NAIVASHA D.O** **Code : 9036002** **The year 1962 excluded due to one month data**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
1960	55.4	28.1	139.9	106.3	39.7	7.9	3.8	89.6	56.1	30.4	65.6	45.6	**668.4**
1961	0	0	26.2	96.5	130.2	8.7	0	49.6	50.5	102.8	334.7	117.3	**916.5**
1963	33.8	38.1	39.3	212.5	131.7	39.7	2.4	25.2	10.5	28.9	87.4	152.6	**802.1**
1964	9.2	52.3	105.9	165.8	10.8	19.9	102	8.1	37.5	89.3	55.2	54.1	**710.1**
1965	34.8	14.9	24.1	110.6	56	32.1	36.6	34	18.1	54.2	66	39.5	**520.9**
1966	20.5	45	73.6	206.2	28.2	12.8	5.6	98.8	87.9	51.1	112.9	5.4	**748.0**
1967	14.6	7.5	50.6	108.9	119.4	46.1	84.4	52.1	34.4	99.1	57	24.9	**699.0**
1968	1.8	113.7	85.5	251.4	13.8	59.1	3.6	2.5	38.9	55.5	84.1	55.4	**765.3**
1969	47	27.8	67.8	25.3	126.6	19.9	14.3	18.1	29.5	25.7	62.6	42.9	**507.5**
1970	82.7	35.4	110.3	110.4	82.2	58.5	26.4	11.6	84.3				**601.8**
1971	41.7	2.3	6.3	64	173.8	11.7	39.3	112.8	21.8	137.7	32.7	49.4	**693.5**
1972	18.6	106.6	17.6	18.6	46.3	92.3	20.7	42.3	32.1	92.5	63.9	10.1	**561.6**
1973	57.2	58.2	8	127.1	84.2	33.5	20.5	29.9	55.9	35	52.1	13.6	**575.2**
1974	13.4	16.9	32.1	223.1	67	75	70.3	44.5	46	30.2	40.8	49.4	**708.7**
1975	1.7	21.4	25.8	127.2	54.8	20.4	80.3	50.6	61.8	81.4	15.7	24.5	**565.6**
1976	7.8	29.8	16.9	69	46.1	67.8	48.4	75.8	47.6	7.5	23.2	46	**485.9**
1977	30.9	24.5	30.6	205	130.6	20.1	67.1	41.8	18.8	78.5	109.5	72.6	**830.0**
1978	76.3	54.2	222.4	125.1	73.7	15	10.9	77.1	79.4	44.6	32.5	98.2	**909.4**
1979	77.7	121.2	58.8	70.7	0	63.4	28.2	49.6	47.6	21.1	74.6	0	**612.9**
1980	39.1	6.5	0	125.6	153.8	67.5	2.2	7.9	32.2	20	93.4	9.7	**557.9**
1981	0.3	24.5	178.7	142	65.7	31.5	54.8	44.8	29.6	41.9	59.5	35.6	**708.9**
1982	24.7	20.8	6.2	118.3	111.3	20.7	8.2	57.4	45.9	92.6	170.7	65.1	**741.9**
1983	35.5	33.8	8.3	84.3	38.2	99.7	39.6	46.2	58.4	92	59.3	38.9	**634.2**
1984	3.2	4.5	3	64.2	8.3	0	51.1	42.1	6.9	52	100.3	101.9	**437.5**
1985	17.8	63.8	66.7	147.1	84.8	22.4	29.1	11.2	51	16.6	21.2	26.7	**558.4**
1986	40.8	39.9	40.6	161	40.1	36.7	7.6	48.6	27	41.9	70	55.8	**610.0**
1987	8.2	19.3	45.8	67.8	118.2	158.8	54.6	42	11.5	18.4	57.8	11.4	**613.8**
1988	38.7	9.8	80.6	121.4	90.6	16	37.9	62.3	27	17.2	41.5	15.5	**558.5**
1989	95.8	23.8	38.9	91	116.6	37.5	80.1	85	56.8	79.4	47.1	99.5	**851.5**
1990	53.4	99.7	130.1	190.7	75.9	14.5	19	77.6	46.6	101.2	63.5	23.9	**896.1**
Average	**32.8**	**38.1**	**58.0**	**124.6**	**77.3**	**40.3**	**35.0**	**48.0**	41.7	**56.5**	**74.3**	**47.8**	**668.4**

TABLE 3.1-B: MONTHLY RAINFALL

Monthly Rainfall (mm) **Lake Nakuru Met. Station** **code 9036261**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
1977	60.1	17.5	22.8	220.8	175.3	78.6	193.1	91.8	58.9	65.6	157.7	74.6	**1216.8**
1978	57.1	123.2	188.7	146.5	117.8	54.5	97.6		128.3	103.4	78.2	65.5	**1160.8**
1979	58.5	165.6	75.4	120	98.3	91.4	81.9	127.2	53.8	29.1	45.2	13.8	**960.2**
1980	28.6	0.9	36.4	102.3	342.1	85.2	26.1	58.6	18.2	36.9	87.4	10.3	**833**
1981	0.3	26	109.9	137.6	165.3	77.6	137.7	169.1	63.1	72	38.3	11.7	**1008.6**
1982	2.6	15	14.5	144.5	188.9	18.4	49.9	191.3	49.2	141.4	186.5	66.7	**1068.9**
1983	14.8	84.6	15.5	87.8	68.5	65.9	67.1	142.8	142.9	96.4	82.8	68.6	**937.7**
1984	3.4	5.9	11.2	65.4	32	67	87.6	36	69.4	79.4	62.1	63.6	**583**
1985	27.8	39.4	96.8	297.5	95.3	123.7	88.5	69.5	27.9	23.6	73.5	8.5	**972**
1986		1.1	21	171.9	94.3	163.3	111.5	127.1	76.2	46.2	41	72.7	**926.3**
1987	23.5	22.4	35.7	80.2	81.9	174.3	20.5	47	27.1	39.1	87.9	4.1	**643.7**
1988	71.3	6.5	43.3	213.4	318.1	80.3	120.5	91.8	125.3	68.4	57.5	46.3	**1242.7**
1989	29	80	69.5	107	82	59.4	115.6	77.1	175.2	138.8	93.5	115	**1142.1**
1990	38.7	74.1	158.1	199.5	137.4	18.8		57.7	43.4	70.9	64.7	52	**915.3**
Average	**49.4**	**45.8**	**90.5**	**210.2**	**156.4**	**48.7**	**193.1**	**74.8**	**51.2**	**68.3**	**111.2**	**63.3**	**1066.1**

Monthly Rainfall (mm) **Gilgil W\S Station** **Code 450**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
1982				21	124.8	57.1	48.6	75.9	35.2	82.6	114	40.3	**599.5**
1983	13.4	29.4	29.2	81.6	59.7	98.7	25.1	102.5	86.9	124.6	75.3	79.3	**805.7**
1984	10	8.2	0.7	44.7	22.9	33.7	69	57.4	10.6	114.2	97.1	36.5	**505**
1985	7.4	39.3	52.7	143.3	50.7	37.3	84.5	16.1	3.5	31.3	53.7	8.8	**528.6**
1986	4.3	3.8	43	130.3	47.5	69.4	66.8	33.9	45.2	26.5	41.4	39.8	**551.9**
1987	23.3	20.5	23.5	90.3	143.7	122.7	47.3	8.9	4.7	38.5	126.7	36.2	**686.3**
1988	48	16.4	86.7	245.1	81.8	41.7	48.3	78	80.3	90.1	78.8	11.3	**906.5**
1989	20.4		39.4	98.2	44.3	38.5	124.8	68.7	79	127.3	123.6	125.3	**889.5**
1990	26.3	74	154.4	207.5	30.8	20.4	54.4	40.3	33.9	40.9	37.9	36.4	**757.2**
Average	**19.1**	**27.4**	**53.7**	**118**	**67.4**	**57.7**	**63.2**	**53.5**	**42.1**	**75.1**	**83.2**	**46**	**692.2**

TABLE 3.1-C: MONTHLY RAINFALL

Monthly Rainfall (mm) **Milmet station** **Code 345**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
1970	182.1	0	121	312.5	135.9	17.5	91.3	129.1	0	0	0	19.3	**1008.7**
1971	0	0	0.1	21.5	17.8	12	0	18.7	12.2	3.5	3.2	9.1	**98.1**
1972	0.2	10.3	3.9	3	27.7	17.3	9.4	9.7	4.1	10	16.4	1.9	**113.9**
1973	35.9	57.6	0	48.7	368.9	47.2	174.2	143.8	171.6	75.6	56.2	20.7	**1200.4**
1974	0	12.3	160.2	117	126.1	42.7	108.3	175.4	141.3	123.5	43.2	8.2	**1058.2**
1975	4.8	12	4	144.7	239.2	115.2	151.2	218.6	108.2	123.1	45.4	17.4	**1183.8**
1976	1.7	6.1	5.1	192.5	185.7	108.2	130.6	141.5	134.1	30.9	64.3	5.5	**1006.2**
1977	30.7	25.6	28.4	315.5	218.1	103.3	71.9	70.8	44.9	96.4	275	81.4	**1362**
1978	35.4	178.8	204.7	147.2	136.4	114.9	216.3	110.4	108.6	200.5	57.1	103.1	**1613.4**
1979	139.4	107.8	109.1	133.5	127.1	68.8	116.9	133.8	0	18.3	102.3	1	**1058**
1980	39.1	0	32.9	227	227.8	35.4	24.8	88.7	41.3	24	137	3.5	**881.5**
1981	0	29.3	164.1	201.4	119.5	96.4	174.6	205.1	143.3	72.6	53.5	0	**1259.8**
1982	19.5	34.7	7.2	222.1	230.8	39.7	54	0	0	0	0	0	**608**
1983	22.1	38.6	10.1	111	227.1	110.4	147.2	207.4	121.1	181.2	58.9	104.4	**1339.5**
1984	0	0	78.1	27.1	15.6	68.6	79.3	61.5	41.1	75.4	124.6	32.3	**603.6**
1985	12.1	100.8	141	188.6	185.1	101	138.1	86	82.8	22.6	53.8	0	**1111.9**
1986	0	3.2	16	173.7	200.3	168	106.7	168.6	56.7	58.4	46.6	25.3	**1023.5**
1987	20.5	54.2	29.9	193	155.5	146.7	41.1	56.5	67.3	0	0	0	**764.7**
Average	**29**	**52.1**	**74.7**	**176.4**	**167.6**	**95.7**	**106.4**	**108.1**	**64.3**	**68.1**	**82.6**	**31.9**	**1056.9**

Monthly Rainfall (mm) **Lake Nakuru National Park Station** **Code 581**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	TOTAL
1977	64.5	22.3	33.7	205	112.2	57	125.7	82.5	37.1	65.5	162	108.8	**1076.3**
1978	73.2	91.8	122.2	164.5	85.2	63.3	59.3	87	110.1	77.6	24.6	45.2	**1004**
1979	62.8	169.2	47.9	101.1	84.8	82.6	60.5	80.6	56.3	22.3	63.2	12.4	**843.7**
1980	41.7	2.7	39.7	141	327.3	108.8	25.1	81	32.3	39.7	83.4	8.7	**931.4**
1981		19.2	98.8	134.8	146.5	69.5	100.2	171.9	86.2	59.3	40.4	29.8	**956.6**
1982	5.1	17.4		160.6	97.2	48.8	42	172	32.6	150.9	98.5	82.8	**907.9**
1983	36.7	43	3.8	83.7	48.4	69.9	60.9	116.9	192	68.8	80.5	54.2	**858.8**
1984	4.1	10.2	12.3	106.9	21.1	49	80.5	52.7	78.8	45.2	69	56.3	**586.1**
1985	23.4	31.1	82.1	173.9	99.7	113.7	67.8	49.5	28.3	14.8	58.6	12	**754.9**
1986	0.6	1.8	18.2	122.5	116.5	127.2	60	54.7	98.1	60.3	57.7	30.9	**748.5**
1987	18.1	10.7	35.5	80.7	143.2	121.8	17.8	46.8	28.8	41.5	133.5		**678.4**
1988	73.4	17.4	45.6	203	183.7	121.3	92	69.1	159.6	89.2	29.6	42.4	**1126.3**
Average	**36.7**	**36.4**	**49.1**	**139.8**	**122.2**	**86.1**	**66**	**88.7**	**78.4**	**61.3**	**75.1**	**44**	**872.7**

TABLE 3.2-A: NUMBER OF RAINY DAYS

No of rainy days **NAIVASHA D.O.** **The year 1962 excluded due to one month data**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1960	7	5	14	18	5	2	2	7	5	5	7	8
1961	0	0	8	13	14	4	0	13	12	21	22	11
1963	14	10	11	20	24	5	3	5	4	9	20	15
1964	5	14	15	18	6	6	10	5	8	14	14	14
1965	9	4	9	18	11	5	11	10	8	9	18	13
1966	8	11	13	18	6	6	4	14	12	8	13	3
1967	4	4	13	20	16	10	10	6	9	14	20	9
1968	1	17	16	21	10	9	3	6	7	14	21	11
1969	10	9	14	11	16	7	4	4	10	5	17	8
1970	19	8	15	19	16	8	5	6	12	0	0	0
1971	9	1	5	20	17	6	11	18	7	8	11	16
1972	11	14	5	9	17	15	6	10	9	15	22	6
1973	13	12	4	13	8	9	8	5	16	13	15	4
1974	5	6	10	23	10	14	14	10	13	11	14	14
1975	3	6	9	17	13	6	11	12	15	12	6	6
1976	4	11	8	14	14	13	14	13	6	8	10	9
1977	10	11	13	24	18	7	15	12	6	13	22	13
1978	9	10	20	25	10	8	10	15	11	17	14	16
1979	11	19	7	20	0	12	6	7	8	11	13	0
1980	7	2	0	16	23	12	2	5	7	7	17	4
1981	1	3	16	15	13	5	8	7	7	5	11	8
1982	4	4	4	12	13	3	2	9	9	18	23	9
1983	6	12	3	13	6	8	6	3	10	10	9	8
1984	2	3	2	11	3	0	9	11	4	7	15	8
1985	3	7	8	17	10	4	4	4	6	6	9	5
1986	7	3	8	12	7	9	5	6	7	10	12	9
1987	4	5	6	11	9	12	5	6	4	5	9	3
1988	5	2	7	15	6	6	6	9	4	4	10	6
1989	8	4	10	8	9	5	7	7	10	12	8	7
1990	6	11	15	19	8	1	4	11	4	9	11	6
Average	**13**	**14**	**20**	**38**	**25**	**15**	**13**	**17**	**14**	**17**	**21**	**14**

TABLE 3.2-B: NUMBER OF RAINY DAYS

No. of rainy days **Lake Nakuru Met. Station** **code 9036261**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	9	10	14	25	21	21	21	13	17	12	27	13
1978	8	15	22	27	15	14	20	0	22	23	16	12
1979	19	18	13	20	20	15	13	17	18	11	17	10
1980	5	2	12	19	27	19	7	14	12	15	23	6
1981	1	7	18	19	20	13	25	19	18	20	13	8
1982	1	5	3	21	17	6	14	18	10	20	24	11
1983	5	9	5	22	19	15	19	23	22	21	19	11
1984	3	2	3	16	9	15	23	18	10	21	20	8
1985	4	4	12	27	18	19	17	19	16	11	14	3
1986	0	2	11	21	24	23	15	15	13	12	17	11
1987	5	5	8	17	17	17	9	12	10	15	18	5
1988	14	7	12	25	20	14	17	21	22	15	15	11
1989	7	9	14	15	16	10	24	20	15	23	19	14
1990	5	13	22	25	15	8	0	15	9	17	18	13
Average	**6**	**8**	**12**	**21**	**18**	**15**	**16**	**16**	**15**	**17**	**19**	**10**

No. of rainy days **Gilgil W\S Station** **Code 450**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1982	0	0	0	8	19	12	12	21	10	19	25	15
1983	2	4	3	9	7	9	11	14	15	22	19	14
1984	3	5	2	20	7	4	15	20	8	16	22	9
1985	4	4	13	26	21	11	14	15	5	13	14	5
1986	4	3	9	24	15	15	17	8	5	13	9	7
1987	7	3	8	11	11	10	4	1	2	8	13	3
1988	8	2	7	16	8	9	13	14	14	15	19	8
1989	2	0	9	16	10	6	16	12	10	18	18	15
1990	7	10	21	22	8	3	9	11	5	11	10	7
Average	**4**	**3**	**8**	**17**	**12**	**9**	**12**	**13**	**8**	**15**	**17**	**9**

TABLE 3.2-C: NUMBER OF RAINY DAYS

No. of rainy days **Milmet station** **Code 345**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1970	15	0	14	19	12	9	12	20	0	0	0	6
1971	0	0	1	16	20	13	0	22	13	11	10	6
1972	1	14	3	3	23	16	11	12	7	13	15	4
1973	8	9	0	11	17	14	14	18	18	8	8	2
1974	0	1	13	20	19	11	21	20	15	12	8	4
1975	1	1	1	15	22	20	17	22	18	14	10	2
1976	1	3	4	13	19	17	17	17	15	5	8	5
1977	9	6	5	21	28	16	12	14	9	19	27	14
1978	7	9	17	19	17	14	26	15	15	23	12	6
1979	15	17	10	12	19	13	13	14	0	6	16	1
1980	2	0	4	13	23	16	13	16	8	8	15	3
1981	0	4	14	21	15	11	21	25	20	9	11	0
1982	3	5	3	21	29	14	16	0	0	0	0	0
1983	5	7	2	12	26	14	21	25	24	28	10	12
1984	0	0	16	10	1	18	17	19	10	18	13	7
1985	4	3	12	24	24	16	18	18	13	9	9	0
1986	0	1	7	20	25	18	16	17	15	9	14	6
1987	3	5	5	15	24	17	10	14	8	0	0	0
Average	**4**	**5**	**9**	**17**	**21**	**15**	**17**	**16**	**11**	**12**	**12**	**4**

No. of rainy days **Lake Nakuru National Park Station** **Code 581**

YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1977	7	7	8	18	15	11	14	10	9	12	21	12
1978	8	13	17	23	11	10	14	18	16	14	7	5
1979	13	13	9	15	13	10	9	8	10	6	9	3
1980	3	1	7	18	20	17	7	13	8	15	18	4
1981	0	5	19	18	17	13	24	17	12	15	15	8
1982	2	5	0	21	17	7	13	19	8	22	21	11
1983	4	8	1	19	12	12	14	15	20	19	22	15
1984	4	5	3	16	6	9	14	18	10	25	19	9
1985	6	6	11	17	20	18	13	17	12	11	13	5
1986	2	3	4	21	20	20	14	10	11	11	14	9
1987	5	5	4	14	17	16	5	10	11	12	14	0
1988	13	7	8	21	16	12	14	12	17	12	14	10
Average	**6**	**7**	**8**	**18**	**15**	**13**	**13**	**14**	**12**	**15**	**16**	**8**

***** This macro for spreadsheet *****

♦♦ This macro checks the missing gaps and create the continuous series (date) and prints 9999 for Gaps ♦♦

```
Sub DataGaps2()
Worksheets("sheet2").Select
Cells(1, 1).Value = "Date"
Cells(1, 2).Value = "Discharge"
k = 2
For i = 2 To 11324
   DateGap = (Worksheets("sheet1").Cells(i + 1, 1).Value) -
(Worksheets("sheet1").Cells(i, 1).Value)
   CurrentDate = (Worksheets("sheet1").Cells(i, 1).Value)
If DateGap > 1 Then
   Cells(k, 1).Value = (Worksheets("sheet1").Cells(i, 1).Value)
   Cells(k, 2).Value = (Worksheets("sheet1").Cells(i, 2).Value)
   k = k + 1
    For j = 1 To CInt(DateGap - 1)
      Cells(k, 1).Value = CurrentDate + j
      Cells(k, 2).Value = 9999
      k = k + 1
   Next j
Else
   Cells(k, 1).Value = CurrentDate   '(Worksheets("sheet1").Cells(i, 1).Value)
   Cells(k, 2).Value = (Worksheets("sheet1").Cells(i, 2).Value)
      k = k + 1
End If

Next i
End Sub
```

***** This DBASE macro for water level *****

```
sele 1
use gA5
sele 2
use reform
zap
```

```
n=1

DO WHILE n<=28
go top
do while .not. eof(1)
sele 1
 mstat=station
 myear=year

if date< 10
 mday='0'+ltrim(str(date))
else
mday=ltrim(str(date))
endif
if ((n+1)/2) <10
mmonth='0'+ltrim(str((n+1)/2))
else
mmonth=ltrim(str((n+1)/2))
endif


XX=ltrim(field(n+3))
 if &XX>0
  mwl=&XX
sele 2
  append blank
  replace station with mstat, year with ltrim(str(myear))
  replace day with mday
  replace month with mmonth
  replace time with '460'
  replace wl with mwl
 endif

 sele 1
yy=ltrim(field(n+4))
 If &yy>0
   mwl=&yy
   sele 2
   append blank
   replace station with mstat, month with mmonth
   replace year with ltrim(str(myear)), day with mday
   replace  time with '960', wl with mwl
 endif

 sele 1
```

```
if &XX=0 .and. &yy=0
sele 2
append blank
replace station with mstat, month with mmonth, year with ltrim(str(myear)),day with
mday, time with '000'
endif


sele 1
skip
enddo
n=n+2
enddo
```

***** The DBASE macro for Rainfall and discharge *****

```
sele 1

use R9036261
sele 2
use reform
zap

n=1

 DO WHILE n<=28
 go top
 do while .not. eof(1)
 sele 1
  mID=ID
  myear=year
 **mday=date

 if day< 10
  mday='0'+ltrim(str(day))
 else
  mday=ltrim(str(day))
 endif
  if n<10
   mmonth='0'+ltrim(str(n))
 else
```

```
 mmonth=ltrim(str(n))
 endif


 XX=ltrim(field(n+3))
mRAIN=&XX

  sele 2
  append blank
  replace ID with mID, year with ltrim(str(myear))
  replace day with mday
  replace month with mmonth

  replace RAIN with mRAIN

sele 1
skip
enddo
n=n+1
enddo
```

***** This DABSE macro for discharge ,rainfall and water level data *****
***** This macro for correcting the leap-year and months <31 *****

```
close data
use reform

sele reform
dele for (month='02' .or. month='04' .or. month='06' .or. month='09' .or. month='11') .and.
val(day)=31
dele for month='02' .and. (val(day)=30 .or. val(day)=29)
recall for mod(val(year),4)=0 .and. month='02' .and. val(day)=29
pack

INDEX ON RTRIM(YEAR)+rtrim(MONTH)+rtrim(DAY) TO rR3627

REINDIX
copy to f9036261
close all
```

***** This macro for spreadsheet *****

♦♦ Macro for filling the missing value (rainfall) data ♦♦

```
Sub level1()
Worksheets("sheet1").Select
For I = 2 To 3938
For j = 4 To 27
P = Cells(I, j)
If P = "" Then
Worksheets("sheet2").Cells(I, j).Value = "9999"
Else Worksheets("sheet2").Cells(I, j).Value = P
End If
Next j
Next I
End Sub
```

Multiple Linear Regression Analysis

This technique was applied to test the combined effects of the different independent variables on the dependent variable. In this procedure, any variables suspected to effect the dependent variable "Y" was included in the analysis. For "k" independent variables, $X_1, X_2, \ldots\ldots\ldots X_k$ the functional form of the multiple linear regression model is

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots \quad \ldots \ldots \ldots \ldots + \beta_k X_k \quad \ldots \ldots \; 4.2$$

where the βi's are the partial regression coefficients associated with each Xi and α is the interception of the line on the Y-axis (the value of Y when all the Xi's have zero values).

Estimation procedure for Multiple Linear Regression Analysis

The multiple linear regression parameters α, β_1, β_2, X_2, β_3 .... + β_k, of the equation (4.2) were estimated by using least square's method. The solution was facilitated by a Statistical computer program SPSS. The usual ways of carrying out the estimation process are as follows.

Assume that the are presented in the following tabular format;

Dependent Variables	Independent Variables X_1	X_2	X_k
Y_1	X_{11}	X_{21}	X_{k1}
Y_2	X_{12}	X_{22}	X_{k2}
.		.. ..	.. ..
..		.. ..	.. ..
Y_n	X_{1n}	X_{2n}	X_{kn}

a) Compute the corrected sum of squares and cross products for all possible pair-combinations of k+1 variables as follows:

	Y	X_1	X_2	X_k
X_1	$\Sigma x_1 y$	Σx_1^2	$\Sigma x_1 x_2$	$\Sigma x_1 x_k$
X_2	$\Sigma x_2 y$	$\Sigma x_1 x_2$	Σx_2^2	$\Sigma x_2 x_k$
. . .	. . .	. . .	. . .	
X_k	$\Sigma x_k y$	$\Sigma x_1 x_k$	$\Sigma x_2 x_k$	Σx_k^2

Where: $\Sigma x_1 y = \Sigma (X_{1j} - X_{m1}) (Y_{1j} - Y_{m1})$

$\Sigma x_1^2 = \Sigma (X_{1j} - X_{m1})^2$

$\Sigma x_1 x_2 = \Sigma (X_{1j} - X_{m1}) (Y_{2j} - Y_{m2})$, etc.

$j = 1, 2, \quad n.$

b) Construction the following k equations

$$\beta_1 \ \Sigma x_1^2 + \beta_2 \ \Sigma x_1 x_2 + \ ..... + \beta_k \ \Sigma x_1 x_2 = \Sigma x_1, y$$

$$\beta_1 \ \Sigma x_1 x_2 + \beta_2 \ \Sigma x_2^2 + \ ..... + \beta_k \ \Sigma x_2 x_k = \Sigma x_2 y$$

$$\begin{matrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & .. & .. \end{matrix} \quad . \ . \ . \quad (4.2)$$

$$\beta_1 \ \Sigma x_1 x_k + \beta_2 \ \Sigma x_2 x_k + \ ..... + \beta_k \ \Sigma x_k^2 = \Sigma x_2 y$$

and solve for β_1, β_2, ..., β_k.. Any procedure for solving for k independent equations for k unknowns can be used. It is obvious that as k becomes large, the task of funding solutions to the equations formed in step (b) become more difficult. However, with the use of computer and available Statistical Software (SPSS), calculations are not difficult.

c) Compute the interception, "α", using the formula:

$$\alpha = Y_m - \beta_1 x_{m1} - \beta_2 x_{m2} - \ldots - \beta_k x_k, \qquad (4.3)$$

Where Y_m, x_{m1},.........., x_{mk} are the arithmetic means of Y, x_1,.........., x_k respectively.

d) The estimated regression equation is:

$$Y_{(est.)} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \qquad (4.4)$$

The uncertainty in the stage-discharge relation

The uncertainty of 2GB1 station .

The uncertainty of 2GB1 station has been shown in table 1

Table 1: Error in 2GB1 station

Observation	H (meter)	H+a (meter)	Qi (m^3/s)	Ln Qi (=Yi)	Qc (m^3/s)	LnQc (=Yc)	Diff.squar. (Yi-Yc)^2
1	2	3	6	7	8	9	10
1	1.26	1.284	45.06	3.807995	42.371904	3.7464855	0.003783
2	2.37	2.394	127.82	4.850623	126.84169	4.8429398	5.9E-05
3	1.62	1.644	68.062	4.220419	65.46222	4.1814732	0.001517
4	2.16	2.184	106.4	4.667206	107.91658	4.6813585	0.0002
5	1.78	1.804	85.108	4.443921	77.086767	4.3449316	0.009799
6	1.54	1.564	65.338	4.179574	59.959805	4.0936744	0.007379
7	1.22	1.244	34.63	3.54472	40.07628	3.6907846	0.021335
8	1.34	1.364	45.325	3.813859	47.127763	3.8528623	0.001521
9	1.44	1.464	45.495	3.817602	53.377205	3.9773838	0.02553
10	1.17	1.194	32.421	3.478806	37.284743	3.6185842	0.019538
11	1.11	1.134	30.885	3.430271	34.050423	3.5278425	0.00952
12	1.04	1.064	27.779	3.32428	30.438329	3.4157026	0.008358
13	0.81	0.834	17.324	2.852093	19.826944	2.9870418	0.018211
14	0.56	0.584	9.218	2.221158	10.589858	2.3598968	0.019248
15	0.11	0.134	0.622	-0.47482	0.7937943	-0.230931	0.05948
16	0.18	0.204	1.44	0.364643	1.6632312	0.5087622	0.02077
17	0.5	0.524	8.094	2.091123	8.7503737	2.1690964	0.00608
18	0.18	0.204	2.065	0.72513	1.6632312	0.5087622	0.046815
19	0.47	0.494	8.739	2.167796	7.8879285	2.0653336	0.010499
20	0.57	0.594	8.855	2.180982	10.911078	2.3897786	0.043596
21	0.12	0.144	0.751	-0.28635	0.9009932	-0.104258	0.033158
22	0.21	0.234	2.058	0.721735	2.1175001	0.7502362	0.000812
23	0.12	0.144	1.017	0.016857	0.9009932	-0.104258	0.014669
24	0.32	0.344	4.04	1.396245	4.1720201	1.4284004	0.001034
25	0.2	0.224	2.193	0.78527	1.9608305	0.6733681	0.012522

SUM = 0.395433

Therefore , Standard error of estimate for 2GB1 station

$$Se_{(2GB1)} = \left[\frac{\sum (LnQi - LnQc)^2}{N-2}\right]^{1/2} = \left[\frac{0.395433}{23}\right]^{1/2} = 0.1311211 \ m^3/s$$

Where N= Number of observation

Qi = Current meter measured discharge (m^3/s)

Qc = computed discharge from rating curve (m^3/s)

Therefore, standard error of estimate for 2GB1 station at the 95% confidence level

$Se_{(2GB1)} * t = 0.13112 * 1.96 * 100 = 25.69 \cong 25.7\%$

The uncertainty of 2GB5 station (lower part):

The uncertainty of $2GB5_{(lower\ part)}$ station has been shown in table 2

Table 2: Error in $2GB5_{(lower\ part)}$ station

Observation	H	Qi	Ln Qi (=Yi)	Qc	LnQc (=Yc)	(Yi-Yc)^2
1	2	6	7	8	9	10
1	0.54	0.31	-1.17118	0.323132	-1.12969	0.001721
2	0.55	0.35	-1.04982	0.358628	-1.02547	0.000593
3	0.55	0.32	-1.13943	0.358628	-1.02547	0.012988
4	0.55	0.26	-1.34707	0.358628	-1.02547	0.103429
5	0.55	0.3	-1.20397	0.358628	-1.02547	0.031863
6	0.56	0.38	-0.96758	0.397275	-0.92313	0.001977
7	0.56	0.36	-1.02165	0.397275	-0.92313	0.009707
8	0.56	0.43	-0.84397	0.397275	-0.92313	0.006266
9	0.56	0.52	-0.65393	0.397275	-0.92313	0.072468
10	0.58	0.45	-0.79851	0.484903	-0.72381	0.00558
11	0.58	0.41	-0.8916	0.484903	-0.72381	0.028154
12	0.58	0.27	-1.30933	0.484903	-0.72381	0.342842
13	0.59	0.6	-0.51083	0.534347	-0.62671	0.013429
14	0.59	0.52	-0.65393	0.534347	-0.62671	0.000741
15	0.6	0.54	-0.61619	0.587872	-0.53125	0.007215
16	0.6	0.5	-0.69315	0.587872	-0.53125	0.026212
17	0.6	0.67	-0.40048	0.587872	-0.53125	0.0171
18	0.6	0.41	-0.8916	0.587872	-0.53125	0.129854
19	0.6	0.53	-0.63488	0.587872	-0.53125	0.01074
20	0.61	0.81	-0.21072	0.645739	-0.43736	0.051365
21	0.62	0.89	-0.11653	0.708221	-0.345	0.052197
22	0.62	0.81	-0.21072	0.708221	-0.345	0.018031
23	0.62	0.34	-1.07881	0.708221	-0.345	0.538477
24	0.64	1.54	0.431782	0.848176	-0.16467	0.355752
25	0.65	0.71	-0.34249	0.926257	-0.0766	0.070696
26	0.65	1.42	0.350657	0.926257	-0.0766	0.182551
27	0.65	0.64	-0.44629	0.926257	-0.0766	0.136666
28	0.66	1.26	0.231112	1.010167	0.010116	0.048839
29	0.67	0.54	-0.61619	1.100243	0.095531	0.506542
30	0.68	1.48	0.392042	1.196835	0.179681	0.045097
31	0.69	0.92	-0.08338	1.300309	0.262602	0.119705
32	0.69	1.84	0.609766	1.300309	0.262602	0.120523
33	0.69	2.29	0.828552	1.300309	0.262602	0.320299
34	0.69	0.79	-0.23572	1.300309	0.262602	0.248327
35	0.69	0.6	-0.51083	1.300309	0.262602	0.59819
36	0.69	1.76	0.565314	1.300309	0.262602	0.091634
37	0.7	1.36	0.307485	1.411044	0.34433	0.001358
38	0.7	1.78	0.576613	1.411044	0.34433	0.053956
39	0.7	0.82	-0.19845	1.411044	0.34433	0.294611
40	0.7	0.95	-0.05129	1.411044	0.34433	0.156518
41	0.71	1.66	0.506818	1.529436	0.424899	0.006711
42	0.71	1.37	0.314811	1.529436	0.424899	0.012119
43	0.72	1.95	0.667829	1.655893	0.504341	0.026729
44	0.72	2.18	0.779325	1.655893	0.504341	0.075616
45	0.72	0.98	-0.0202	1.655893	0.504341	0.275146

SUM = 5.230532

Therefore, Standard error of estimate for 2GB5 station (lower part)

$$Se_{(2GB1)} = \left[\frac{\sum(LnQi - LnQc)^2}{N-2}\right]^{1/2} = \left[\frac{5.23}{43}\right]^{1/2} = 0.34875\ m^3/s$$

Where N= Number of observation
Qi = Current meter measured discharge (m^3/s)
Qc = computed discharge from rating curve (m^3/s)

Therefore, standard error of estimate for 2GB1 station at the 95% confidence level
$Se_{(2GB5)} * t = 0.34875 * 1.96 * 100 = 68.35\%$ for lower part

The uncertainty of 2GB5 station (upper part):

The uncertainty of 2GB5$_{(upper\ part)}$ station has been shown in table 3

Table 3: Error in 2GB5$_{(upper\ part)}$ station

1	0.77	2.1	0.7419373	2.848709	1.046866	0.092982
2	0.79	2.24	0.8064759	2.983277	1.0930224	0.082109
3	0.82	3	1.0986123	3.190287	1.1601109	0.003782
4	0.86	2.52	0.9242589	3.475858	1.2458414	0.103415
5	0.87	4.11	1.413423	3.548947	1.2666509	0.021542
6	0.89	3.51	1.255616	3.697148	1.3075618	0.002698
7	0.96	4.98	1.6054299	4.236947	1.443843	0.02611
8	0.99	4.08	1.406097	4.478249	1.499232	0.008674
9	1	4.96	1.6014057	4.56	1.5173226	0.00707
10	1.02	5.8	1.7578579	4.725472	1.5529674	0.04198
11	1.36	8.69	2.1621729	7.931127	2.0707951	0.00835
12	1.38	8.07	2.0881535	8.142302	2.0970729	7.96E-05
13	1.49	9.29	2.2289386	9.3476	2.2351196	3.82E-05
14	1.62	10.34	2.3360199	10.86655	2.3856897	0.002467

SUM = 0.401298

$$Se_{(2GB5)} = \left[\frac{\sum(LnQi - LnQc)^2}{N-2}\right]^{1/2} = \left[\frac{0.401298}{12}\right]^{1/2} = 0.18287\ m^3/s$$

Therefore $t*Se_{(2GB5)} = 1.96 * 0.18287 * 100 = 35.84\ \%$ at the 95% confidence level for higher part

The uncertainty of 2GC4 station (lower part):

The uncertainty of $2GC4_{(lower\ part)}$ station has been shown in table 4

Table 4: Error in $2GC4_{(lower\ part)}$ station

Observation	H	Qi	Ln Qi (=Yi)	Qc	LnQc (=Yc)	(Yi-Yc)^2
1	2	6	7	8	9	10
1	0.17	0.44	-0.82098	0.4411	-0.8185	6.15E-06
2	0.18	0.35	-1.04982	0.4811	-0.731619	0.101253
3	0.18	0.47	-0.75502	0.4811	-0.731619	0.000548
4	0.18	1.47	0.385262	0.4811	-0.731619	1.247425
5	0.2	0.67	-0.40048	0.5647	-0.571471	0.029239
6	0.22	0.89	-0.11653	0.6527	-0.4266	0.096141
7	0.23	0.59	-0.52763	0.6984	-0.359033	0.028426
8	0.25	0.81	-0.21072	0.7927	-0.232293	0.000465
9	0.26	0.76	-0.27444	0.8414	-0.172678	0.010355
10	0.26	0.81	-0.21072	0.8414	-0.172678	0.001447
11	0.26	0.99	-0.01005	0.8414	-0.172678	0.026448
12	0.27	1.09	0.086178	0.8911	-0.115312	0.040598
13	0.27	1	0	0.8911	-0.115312	0.013297
14	0.27	0.91	-0.09431	0.8911	-0.115312	0.000441
15	0.27	0.84	-0.17435	0.8911	-0.115312	0.003486
16	0.28	0.89	-0.11653	0.9417	-0.060033	0.003192
17	0.29	0.75	-0.28768	0.9933	-0.006695	0.078954
18	0.3	1.18	0.165514	1.0459	0.0448357	0.014563
19	0.3	0.89	-0.11653	1.0459	0.0448357	0.02604
20	0.32	1.16	0.14842	1.1537	0.1429343	3.01E-05
21	0.33	1.05	0.04879	1.2089	0.1897072	0.019858
22	0.35	1.44	0.364643	1.322	0.2791447	0.00731
23	0.35	1.36	0.307485	1.322	0.2791447	0.000803
24	0.4	1.25	0.223144	1.6195	0.4821125	0.067065
25	0.41	1.3	0.262364	1.6814	0.5196452	0.066193
26	0.42	2.44	0.891998	1.7442	0.5562735	0.112711
27	0.46	1.56	0.444686	2.0028	0.6945506	0.062432
28	0.47	2.24	0.806476	2.0694	0.72724	0.006278

SUM = 1.883583

Therefore, Standard error of estimate for 2GC4 station (lower part)

$$Se_{(2GC4)} = \left[\frac{\sum (LnQi - LnQc)^2}{N-2}\right]^{1/2} = \left[\frac{1.88358}{26}\right]^{1/2} = 0.2691 \text{ m}^3\text{/s}$$

Therefore t * Se = 1.96 * 0.269 * 100 =52.7 % at the 95% confidence level for lower part

The uncertainty of 2GC4 station (upper part):

The uncertainty of 2GC4$_{(upper\ part)}$ station has been shown in table 5

Table 5: Error in 2GC4$_{(upper\ part)}$ station

Observation	H	Qi	Ln Qi (=Yi)	Qc	LnQc (=Yc)	(Yi-Yc)^2
1	2	6	7	8	9	10
1	0.5	2.21	0.79299	2.3676	0.8618786	0.004745
2	0.55	3.09	1.12817	3.4631	1.2421662	0.012995
3	0.83	17.98	2.88926	17.887	2.8840808	2.68E-05
					SUM =	0.017767

Therefore,

$$Se_{(2GC4)} = \left[\frac{\sum (LnQi - LnQc)^2}{N-2}\right]^{1/2} = \left[\frac{0.017767}{1}\right]^{1/2} = 0.13329 \text{ m}^3/\text{s}$$

Therefore $t * Se_{(2GC4)}$ = 1.96 * 0.13329 * 100 =26.12 % at the 95% confidence level for higher part

It can be noted that only three observation is available for 2GC4$_{(lower\ part)}$ station. The number observation is not enough for calculating reliable uncertainty .