A STATISTICAL MODEL OF FLOW LEVELS IN

RIVER LOKKO - ADAMAWA STATE

BY

OLOGBOSERE, OSAYOMONRE ANTHONY

DEPARTMENT OF WATER RESOURCES AND

ENVIRONMENTAL ENGINEERING

AHMADU BELLO UNIVERSITY

ZARIA

PRESENTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE AWARD OF THE

DEGREE OF MASTER OF SCIENCE IN WATER RESOURCES

AND ENVIRONMENTAL ENGINEERING

AHMADU BELLO UNIVERSITY

ZARIA

MAY, 1993.

D E D I C A T I O N

This work is dedicated to the service of God and humanity.

D E C L A R A T I O N.

I declare herewith that the content of this research report are my original works and, that no similar works has been done in the project area to the best of my knowledge.

PREFACE TO STUDY

The purpose of this project is to present a methodology and to develop a model for estimating flow level data using data from Lokko river gauging station. It is hoped that the methodology will be applicable to other river basins. In the development of the mathematical model, efforts were made to use the fundamentals of deterministic, parametric and stochastic modelling techniques. Attempts were made at the descriptions of the internal hydrological processes of the catchment area, the complexity of the exercise is associated with the scale of the catchment areas as the descriptions of internal hydrological processes are based on the equations of fluid mechanics, therefore some generalization were made, however, a methodology is hereby presented.

At a catchment area scale of the order of a few hundred square kilometres, it is almost impossible to attempt to model the hydrologic response using the internal descriptions which are based on the equations of fluid mechanics. This is because of the great spatial variability of many of the variables concerned, for example, topography, soil properties, vegetation, etc. Therefore, resort is usually had to

study the overall system without reference to the detailed structure of the components of the processes of hydrology involved. The nature of the operation of the system is inferred from the input and output observations, and hence it is referred to as "external description". The conceptual model so developed are lumped models in which only variations in time are examined.

Hydroclimatologic processes are either continuous, meaning that there are non zero values of that variable at anytime, or intermittent, meaning that there are zero values of that variable at some time. Instantaneous precipitation, evaporation, sediment transport in rivers and some run off (usually on some small rivers with negligible underground or surface water storage), represent the typical hydrologic intermittent time series. The river Lokko flows throughout the year, meaning that the river flow level series is continuous, while the rain fall is not continuous, meaning that the rainfall is intermittent with zero values on days of no rainfall or one of insignificant amount.

An intent of this research work is to present a methodology for developing some mathematical model which

would serve as an early warning system as to the occurence of floods in the project area. And as a secondary objective, to provide some relevant information for water resources planning and development in the area.

For purposes of continuity and better understanding, the work has been presented in a number of chapters:

Chapter 1 - Introduction:

This chapter discusses floods in general and their relationship to man; it also discusses some historical flood events and defines floods in general, giving standard for its measurement in river Lokko. It discusses in detail the general characteristics of the project area, the scope and objective of the study, the sources of data and other information needed for the study, including the advantages of modelling flow level data, the approach to the problem which discuss the various deterministic, parametric and stochastic methods employed and the results obtained.

Chapter 2 - Literature Review:

The research efforts in hydrological estimation are discussed in this chapter, in addition to the limitations posed by data availability, or the approach to the problem, or computational suffistication.

Chapter 3 - Existing drainage and storage pattern:

The geomorphology of the catchment area with particular emphasis on the existing drainage network and storage pattern are examined, results obtained suggests improvements which could be made to reduce flood wave propagation rate, improve basin storage capacity, encourage infiltration into the underground reservoir, reduce sediment transport rate and ameliorate the flood situation. Useful geomorpphologic results are presented.

Chapter 4 - Basic data:

This chapter discusses the relevant basic data which include, climate, precipitation, temperature, wind, radiation, air humidity and water budget, as well as the physiography, geology, geomorphology, Land system and land use of the area.

Chapter 5 - Time series method for the estimation of stream flow data based on Box-Jenkins approach.

The methodology employed is discussed in detail and appendixes are provided where necessary. Figures and tables are provided to further elucidate the points being discussed.

Throughout, attempts will be made to present the materials in a form which will be useful not only to specialist

concerned with flood problems, but also to researchers and administrators in such diverse areas as hydrology, geology, civil engineering, water engineering, agriculture and economic planning. The detailed references presented should provide a sound basis for further studies.

Osayomonre. A. Ologbosere,
Federal Polytechnic Mubi,
Adamawa State,
Nigeria.

A C K N O W L E D G E M E N T

I wish to express my profound gratitude to the Management of the Federal Polytechnic, Mubi for her sponsorship and understanding of my need for a graduate study, and value for a staff to be trained as a specialist in water resosurces and environmental engineering. Without their financial support and release for study fellowship, it would have been difficult for me to commence this programme. Special thanks to members of the Polytechnic community for their words of encouragement, particularly the Rector, Alhaji A.M. Gerei, the Vice Rector, Engr. F.C. Udeagwu, the Registrar, Alhaji U.B. Gerei, the Director, School of Engineering, Engr. C.E. Eteng, and the Acting Head of Civil Engineering Department, Mr. A. Mbata.

I am particularly grateful to Dr. L.I. Odigie for his advice and superivision of this work, without whose knowledge and experience I could not have completed this study. Special thanks to Prof. Ogunrobi and Dr. Okuofu for their understanding and advice during my graduate studies, and all those including Engr. Douglas Osula who found time to go through this work and offered their comments which has helped in putting together this beautiful piece. My thanks also to

the typist Mrs. Saratu B. Yerima who typed the manuscript for the pains taken to ensure that the errors were kept at the minimum.

I am especially thankful to my wife, Grace for her financial and moral support, and my children Judith, Michael, Martha and Tony (Jnr) who always offered to follow me down to the campus to do all my cookings and washings. But I always disappointed them, which they never failed to express to me through their mother.

Lastly, I am especially thankful to God for everything, for without His mercies I could not have attempted the programme in the first instance, and without His grace it would have been impossible for me to complete this work. May honour and glory be to Him for ever and ever, Amen.

ABTRACT

Monthly water levels for four years in Lokko river were studied statistically using the time series theory, and an ARMA (1.1) model was identified as suitable to describe the data. An estimation model was developed, the results obtained were found to be satisfactory, except in the rainy season when the *error* could be as high as 37 percent on the average, and as high as 50 percent for a single event, due to random storm rainfall input to the streamflow which the model is unable to account for fully. A factor of safety of 1.6 is therefore recommended for the multiplication of the estimated levels for early warning and for engineering works a factor of safety of 2.2 is recommended.

TABLE OF CONTENT

Title	Page
DECLARATION.....	ii.
CERTIFICATION	iii
DEDICATION	iv
PREFACE TO STUDY	v.
ACKNOWLEDGEMENT	x.
ABSTRACT	xii
TABLE OF CONTENTS	xiii
LIST OF TABLES	xviii
LIST OF FIGURES	xix
LIST OF SYMBOLS	xx
CHAPTER 1: INTRODUCTION	1.
1.1 General	1.
1.2 Devastating effects of floods and historical perspective	3.
1.3 Definition of floods	5
1.4 The project area	7
1.5 Objectives and scope of study	11
1.6 Sources of data and other information ...	12
1.7 Advantages of modelling flow data	12
1.8 Approach to the problem	13

CHAPTER 2: LITERATURE REVIEW ..... 19.

2.1 General 19

2.2 The rational formula 20

2.3 The water - balance model ..... 21.

2.4 The linear model 21

2.5 The routing model 22

2.6 The infiltration model 22

2.7 The hydrological simulation model..... 23

2.8 The statistical model 25

2.9 The time series model 30

2.10 Concluding remarks 33

CHAPTER 3: EXISTING DRAINAGE AND STORAGE PATTERN 34

3.1 General ..... 34.

3.2 Average width of basin 37

3.3 Coefficient of basin shape 39

3.4 Average slope of the basin and flood plain 39

3.5 Time of concentration 43

3.6 Nature of river Lokko 44

3.7 Numerical description of the features

of the basin 44

3.8 Concluding remarks 51

CHAPTER 4: BASIC DATA 53

4.1 General 53

4.2 Climate 53

4.3 Precipitation 55

4.4 Run off water level 57

4.5 Temperature 62

4.6 Winds 63

4.7 Radiation 64

4.8 Humidity 65

4.9 Physiography 65

4.10 Geology 66

4.11 Geomorphology 68

4.12 Land system 71

4.13 Land use 71

CHAPTER 5: TIME SERIES METHOD FOR THE ESTIMATION OF WATER LEVELS IN RIVER LOKKO 73

5.1 Auto correlation analysis of the historical data 73

5.2 Identification of pattern in the historical data 80

5.3 Identification of stationarity in the historical data 80

5.4 Identification of seasonality in the historical data 81

5.5 Transformation of the historical data ... 82

5.6 Auto correlation analysis of the transformed data 86

5.7 Identification of pattern in the transformed data 89

5.8 Identification of stationarity in the transformed data 89

5.9 Identification of seasonality in the transformed data 90

5.10 Removing non-stationarity from the transformed data 90

5.11 Auto correlation analysis of the differenced transformed data 91

5.12 Identification of pattern in the
differenced - transformed data 94
5.13 Identification of stationarity in the
differenced - transformed data 94
5.14 Identification of seasonality in the
differenced - transformed data 95
5.15 Partial auto correlation analysis of the
difference - transformed data ..... 95.
5.16 Identifying a tentative ARMA (p,q)
model..... 99
5.17 Estimating the parameters of the
ARMA(1,1) model 99
5.18 Transformation of the estimated data 105
5.19 Model sensitivity analysis 106
5.20 Model modification 106
5.21 Model validity 114
5.22 Use of the model for estimation..... 118
CHAPTER 6: DISCUSSIONS OF RESULTS 122
6.1 General 122
6.2 On randomness of data 123
6.3 On stationarity of data 125

6.4 On identification of tentative
ARMA(p,q) model 125
6.5 On ARMA (1,1) model parameter estimates.. 126
6.6 On modification of the model 129
6.7 On validity of the model 132
6.8 On table of results 133

CHAPTER 7: RECOMMENDATIONS AND CONCLUSIONS..... 137
7.1 Recommendations 137
7.2 Conclusions 138
REFERENCES 140
APPENDIX 155

LIST OF TABLES

Page

Table 1 Average velocities of flow in a basin for concentration time derivation ----------35

Table 2 Results: Using historical data ----------74

Table 3 Student t Distribution ----------77

Table 4 Critical points of the Chi-squared (x^2) statistics ----------79

Table 5 Monthly adjustment constants, K(E_t) --110

Table 6 Results: Using generated data ----------115

LIST OF FIGURES

Figure 1 Study Area ----------6

Figure 2 General features ----------35

Figure 3 Main rivers, Natural surface reservoirs and flood plains ----------38

Figure 4 Histogram of monthly rainfall and hydrographs of monthly water levels -----59

Figure 5 Histogram of rainfall and hydrographs of daily water levels ----------61

Figure 6 Auto correlation coefficients of historical mean monthly water level against period (month) ----------77

Figure 7 Plot of a against skewness coefficient, G..87

Figure 8 Auto correlation coefficients of differenced - transformed data of monthly mean water level against period (month) ---92

Figure 9 Partial auto correlation coefficient of differenced - transformed data against period (month) ----------93

Figure 10 Mean square error versus parameters i_1, m_1 and m_{12} ----------97

Figure 11 Error at previous period (e_{t-1}) versus transformed forecast data (x_t) ----------98

Figure 12 Hydrographs of monthly water levels against period (month) ----------102

Figure 13 Auto correlation plot of the residuals -----104

ABBREVIATIONS, SYMBOLS AND DEFINATIONS OF TERMS

The symbols used in this work are the conventionally simple types and should therefore present minimum problem of recognision. However, for ease of referencing a full table with explanation/descriptions of all the symbols used is presented below:

A	Total catchment area
A_s	River Sensen catchment area
A_k	River Lokko catchment area
A_{i-1}	Average area of catchments of streams of order i-1
A(u)	Average catchment area of order u
AR	Auto Regressive Process
ARIMA	Auto Regressive Integrated Moving Average Process
ARMA	Auto Regressive Moving Average Process.
b_w	Average basin width
a, b, a_o, a_1^1, c_p	Model parameters
c	Coefficient of run off
d	Coefficient of basin shape
d_f	Number of degrees of freedom

D(u)	Drainage density
e_t	Error in X_t at time t
E_t	Error in model error at time t
E_t, K	Monthly adjustment constant
B_a	Skewness coefficient for a
hmax	Highest point in the basin
hmin	Lowest point in the basin
I	Rainfall intensity
K	Time lag or period
L_s	Length of River Sensen
L_k	Length of River Lokko
L	Average basin length
Li-1	Average length of stream of order i-1
Lsg	Length of River Senguno
Lus	Length of upper River Sensen
Luk	Length of upper River Lokko
LGR	Length of river Garwa
L(u)	Average length of stream of order u
Lof	Average length of overland flow
a	Power index in Box-Cox transformation.
$a^1{}_t$	Model estimate for a_t
a_t	Value of a at time t
m_1, m_{12}, i_1	Parameters of ARMA or AR and MA process
x^2 - test	Chi - square test

e^2	Variance
m	Year mean of monthly adjustment constant
MA	Moving Average Process
MARMA	Multiple Auto Regressive Moving Average Process
n	Number of data in time series
P	Probability of exceedence in random sampling

Q	Run off or discharge
R_b	Bifurcation ratio
N_i	Number of streams of order i
R_L	Stream length ratio
R_a	Stream area ratio
r_L	Ratio of stream length ratio to bifurcation ratio.
g	Auto correlation function
r	Auto correlation estimates
r_k	Auto correlation coefficient at k lag time
$r^1{}_k$	Auto correlation coefficient of transformed data at time k
$r^{11}{}_k$	Auto correlation coefficient of transformed - differenced data, x^{111},

	at time k
s	Average basin slope
s_F	Average flood plain slope
t	Time of concentration
u	Catchment order
v	Average velocity of flood flow
x_t	Transformed observation at time t
$\bar{x}$	Mean of transformed data
x^1	First differenced series of transformed data.
x^{11}	Seasonal series of transformed data

x^{111}	Product of first differenced and and seasonal series.
x^T_t	Transformed series
x^F_t	Forecast series
$\bar{Y}^D$	Mean of original observation
Y_t	Original observation at time t
Y_{t+k}	Original observation at time t + k
$F^Y t$	Forecast data at time t

CHAPTER 1: INTRODUCTION.

1.1 GENERAL

From earliest times floods have been one of the main causes of interaction between man and his environment, indicating the natural events of floods and man's inability to control them. This is due mainly to the fact that the flood plains generally attract economic development and settlement, and man in his quest for survival would naturally look for such a place to settle.

In arid and semi-arid areas, flood plains could be found as extensive agricultural farmlands because they are rich in mineral nutrients required by plants for their growth. And in rugged areas they often represent the only extensive tract of flat land for easy cultivation and communication. The proximity to water usually make them atttactive for agriculture, and often attract industrial establishment and trade links. This situation is well noted of Lokko village which though has no notable industry is an agricultural settlement and a trade centre, as well as providing a link between Jimeta-Yola the capital of Adamawa state and little Gombi in the northern part of the state and, through to Maiduguri in Borno state. If such a link is distrupted due to flooding, the trade and communication link between Borno state, the northern and the southern parts of Adamawa state would be distrupted resulting in economic hardship.

Man has been responding to floods since long before the days of Noah (Genesis 7), they are still not fully understood by the hydrologists, engineers, agriculturists, planners and politicians. Worse still is it to the general public who finds flood as an unexpected, often inexplicable and always a traumatic experience. Man has therefore come to see flood as a natural phenomenon which must occur so that some other natural effects, beneficial or destructive can take place.

The purpose of this work therefore is to seek for ways of reaping the beneficial effects and avoiding the destructive effects. This can be done with the present day scientific knowledge as to predict when flood would occur, with some reasonable degree of accuracy, and such knowledge should be transferable.

It is because of the place flood has come to occupy in human existence that in the early sixties major attempts were made in various countries to improve on the knowledge of flood hazards and of the possible responses to it. In the United states of America (U.S.A) a task force was set up which reported in 1966 calling for a more imaginative reponse to flooding than building flood embankments and reservoirs; In Leningrad, United Soviet Socialist Republic (U.S.S.R) in 1967 an international symposium on "Floods and their computation" was jointly convened by the International

Association of Hydrological sciences, UNESCO and the World Meteorological Organisation, and it attracted over 100 papers from 23 countries. The need to know the devastating effects of floods, and be able to forecast flood will form the subject of discussion in the following sections.

1.2 DEVASTATING EFFECT OF FLOODS AND HISTORICAL PERSPECTIVE

The earliest known devastating effects of floods dates back to the time of Noah (Genesis 7) in eastern Turkey, near Mount Ararat, when God directed Noah to construct a boat large enough to take him and his household, as well as enough food and pairs of every living thing on earth and that within seven days He (God) was going to send rain that would destroy all living things that were not in the boat. It is also reported that the rain fell for fourty days and fourty nights before it stopped and it took one hundred and fifty days before the flood was abated, and a further twenty-one days before bare ground could be seen. Only Noah and his household was saved.

Similarly, we are told by Lambert and Millard (1969) of the Babylonian flood saga in which King Atra-hasis was only given seven days by the god Enki to pull down his reed house and build a boat with the materials. The boat was loaded with his possessions, animals and birds; As soon as they were

aboard, the flood came and, apart from Atra-hasis, his passengers and family, we are told that the entire human race was destroyed. The flood we are further told lasted for seven days and seven nights. These are examples of the earliest floods on records, it could be noted that in both cases seven days advance warning was given and no need was required for the construction of dams, levees, etc. It seems therefore that the early warning system is the most desirable, by observing the rise of the stage above the maximum permissible limit of 1.30m above datum.

Further flood events can be found on account of 1887 flooding of Hwang HO in China which claimed over 900,000 lives; A tropical cyclone struck the shores of Bengal in November 1970 and devastated the emerging nation of Bangladesh leaving over 6,000 people dead on just one island when the protecting levees broke; June 1972, Hurricane Agnes caused the single largest natural disaster, in terms of damage, in the history of the U.S.A. with estimated loss approaching N30 billion; August 1974, disastrous flooding inundated almost half of Bangladesh leaving over five million people homeless. In Nigeria it has dealt its deadly blows on both urban and rural dwellers, the government have not been left out in feeling the impact of flooding. According to Oranugo (1990), the peasant farmer's half acre plot and the

nations economic nerve centre the oil fields, the slump area in Maroko village and an army barraacks in Lagos, the down south town of Port-Harcourt and the up north town of Daura have all tasted the devastating effects of floods; In addition the residents of Uwa street and environs in Benin City, in 1976 were nearly sacked from their natural habitats; the Ogunpa flood of 1980 claimed about 100 lives; in Kano, 1988, flood destroyed about 23,000 houses and claimed about 80 lives; in the same year, nearly 61,000 tonnes of cash crops were washed away in Kwara, Bauchi and Sokoto states. In September 1989, flood rendered 30,000 inhabitants of Lokko village and some neighbouring villages in Adamawa state homeless, and in 1990 flood devastated farm lands and houses in Fufore village in Adamawa state, to mention but a few, which it is hoped is enough to inform the reader that in some places "Flood" is synonymous to disaster or death.

1.3 DEFINITION OF FLOODS

Before attempting to define floods, a few words about what floods are would not be out of place. Floods are natural occuring range of stream flow conditions. These natural events are very often mistaken for natural disaster; floods are not natural disaster, but are man made in that man has put himself at risk by developing flood plains for settlement, agriculture and industry, and by building roads,

bridges and railways lines in floodable positions, due to ignorance or economy.

To state a definition for flood is extremely difficult, partly because floods are complex phenomena and partly because they are viewed differently by different people, however, any meaningful definition of flood must incoporate the notions of damage and inundation. Some of such definitions were proposed by Chow,V.T. (1956), who defined flood as a relatively high flow which overtaxes the natural channel provided for the runoff; and Rostvedt (1968), proposed that a flood is any high stream flow which overtops natural and artificial banks of a stream. However, because the channel geometry of a stream vary throughout its course there is no single bankfull level everywhere above which the river is in flood and below which it is not in flood. In a strictly hydrological sense, therefore, a flood may be any relatively high water level (or stage) or discharge above a selected water level or flood discharge. For the purposes of this work, therefore, a flood in Lokko basin will be defined as a selected water surface profile in Lokko river channel to be defined by a datum of 96.280m (1.300m on gauge meter), at the stage measurement station on the Lokko bridge abutement, which when exceeded will cause the body of water to rise to overflow land which is not normally submerged at the datum

level.

1.4 THE PROJECT AREA

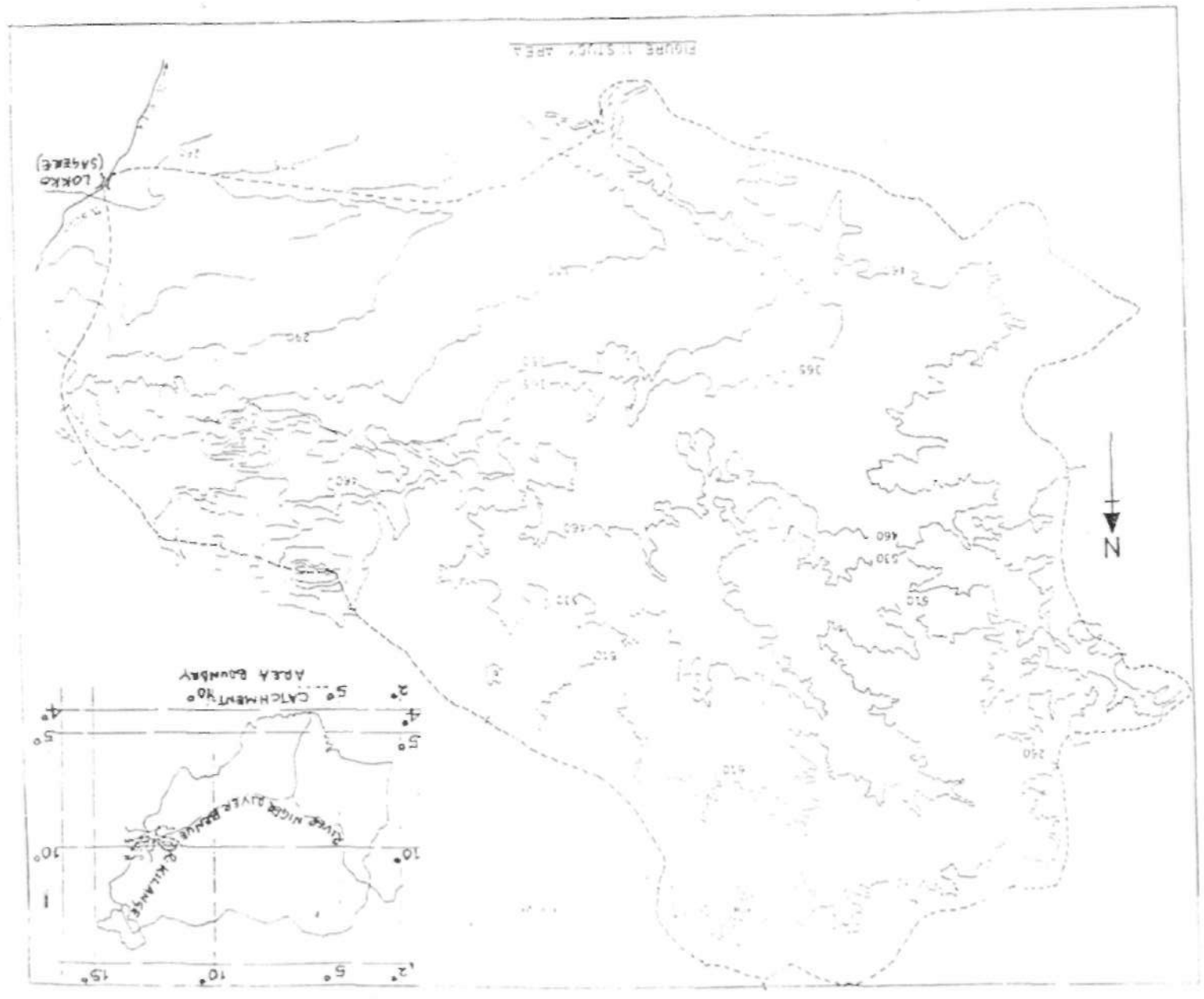
The study area (figure 1) which lies between longitude 12°15'E and 12°36'E and latitude 9°44'E and 10°09'N is the catchment area of river Lokko.

It is located to the North-West of Lokko village, which is situated at about 60 kilometres north of Jimeta Yola, the capital of Adamawa State, and about 4 kilometres south of Song, the headquarters of Song Local government. The River Lokko is one of the tributaries of Kilange river and drains into the River Benuein the south east.

The study area covers about 86,200 hectares and has a population of over 105,000 people. The south eastern part of the catchment area is a flood plain which drain towards the Lokko village; one of the largest villages in the area, Dumne, is located south of the flood plain. The village is located at approximately $9^{o}47'N$ $12^{o}24'E$. A minor road passes through the village linking it with another village, Laro, to the west of Dumne, thus making transportation along that reach to Lokko possible.

Plantations can be found at Laro and Dumne and, along the flood plain. In a fairly wide strip north of the aforesaid road can also be found scattered vegetation with some farming at Lokko village. Further North and running in

FIGURE 1: STUDY AREA



the North-West and South-East direction and along the Western part of the catchment area are scattered hills and peaks, criss-crossed by numerous streams, some of these streams drain directly into Lokko River, while others drain first into River Sensen, which also drain into River Lokko. However, most of the area drain directly into River Lokko. The lower portion of River Sensen from latitute $9^{o}53'N$ down to latitute $9^{o}47'N$ is bordered by alluvial fan with scattered cultivation..1s2

None of the rivers has significant perenial flow, but there is small dry season flow in Lokko River which is too small to support irrigation farming, without some storage facility. P.F.P(Nig) et al (1980) estimated the average annual run-off of River Sensen to be 37 million cubic metres, from synthetic run-off data for measured rainfall data of the River Kilange basin, therefore, the average annual run-off of Lokko River was estimated to be 70 million cubic metres, as it receives about two times the rainfall effect directly considering the uniform rainfall distribution within the climatic band, Cocheme et al (1967) enclosing the study area. The Lokko River receives most of this flow during the wet season, while the rest is from baseflow contributions. The early wet season flows are extremely unreliable and flashy because of the soil moisture deficit resulting from the dry

season, such that the river flow lag behind the rainfall. The groundwater resources have not been investigated and so no information is available but indications are that the water table is fairly shallow in the flood plain hence the dry season flow in River Lokko.

Most households own two "wet" fadama plots and two "dry" upland plots and nearly 90 percent own some "wet land" reflecting the importance placed upon farming. The major crops cultivated in the area presently is sorghum which is also the staple diet of the inhabitants. Maize and rice are crops of next importance. Other crops are groundnuts, vegetables, cowpeas, millet and cassava.

The main highway from Yola to Maiduguri passes through Lokko, Song, little Gombi and Michika. Off this highway is virtually inaccessable except for track roads which are not motorable. No town or village in this catchment area has electricity, pipe borne water or telephone, the nearest town with electricity is Song, which is supplied by the Rural Electrification Board (REB), which was commissioned in 1981. Water wells are extremely lacking in the project area except for temporary excavations in river or stream channels.

Farm produce are sold weekly in local markets in Lokko, Song and Dumne. Marketing will improve with the provision of improved road network system.

The catchment area has no single rainfall measurement station, the closest to it being located at Song, about 4 kilometres North of Lokko village where a rain gauge measurement station is located; No flow measurement is done here but at Malabu on River Kilange to which River Lokko is a tributary, therefore, flows in River Lokko was synthsized from available records of flows in River Kilange and River Sensen, P.F.P.(Nig.) et al.(1980)

1.5 OBJECTIVES AND SCOPE OF THE STUDY

The main objective of this study is to develop appropriate simple and flexible mathematical models for the flow from River Lokko based on monthly data. The one month was chosen as the length of the discrete time interval because it contains more information about the run-off processes than the annual time series. Generally, information on the processes of hydrology are lacking in the study area, attempt therefore was made to use available data to develop the needed mathematical models.

The parameters of the mathematical models so developed were made of functions of a position within the catchment area. The primary objectives of this study therefore,are:

a) to develop mathematical models for estimating the height of flood-waters in the study area, and

b) to develop mathematical models that will provide the

longest possible advance warning time of critical flood heights.

The secondary objective is to provide information for other beneficial uses.

1.6 SOURCES OF DATA AND OTHER INFORMATION

The data and other information used in this study were obtained from several sources including personal discussions with some of the residents of the study area.

The main sources of information include:

a) the hydro-meteorological unit of the Upper Benue River Basin Authority, Yola, Adamawa State, Nigeria.

b) Local Newspapers.

c) Engineering project reports

1.7 ADVANTAGES OF MODELLING FLOW DATA

Synthetic generation of flow data is not designed to produce samples that will replace the historical set of data. However, a mathematical model which is capable of generating reliable data has several advantages over the historical data. The mathematical model may be very well suited for operational procedure. The identified advantages therefore include the following:

a) The information contained in a historic set of data can be more closely examined and more completely extracted

by examining the time series patterns that are as lik
to occur in the future as the observed pattern.

b) The mathematical models can be used to gener
additional data as may be required for the purpo
of short-term or long-term forecasting.

c) Generated data are generally in compactible usable fo
whereas, historical data often have to be laborou
extracted from office files and publications, that
not necessarily data bank.

d) Sets of data and other relevant information can
available for analysis of water resources projects.

1.8 APPROACH TO THE PROBLEM

The approach to the mathematical modelling was based
the outcome of the statistical tests which was carried out
the data. The correlation coefficient was computed at zer
lag and was found not to be equal to zero, which shows tha
there is correlation among data, indicating that a time
series model would be most suitable to describe the data.
time series, the observation run across time and the
estimation of the future is based on past values of
variable and/or past errors.

Time series are generally studied in two domains,
time domain and the frequency domain. In the time domain
operation consists of measuring and identifying the elemer
of the series, trend, cyclical movement, seasonal movemer

and randomness (or residual) and, in some cases, of removing certain elements from the series; This process of seasonal adjustment consists of the identification of the seasonal element and then the removal (or filtering) of the seasonal element from the other element of the time series. While in the frequency domain, the decomposition of the time series into its component parts which are associated with frequency rather than with time are studied. In this work, the series are studied in time domain rather than frequency domain.

In the development of the time series model for the water levels in River Lokko in the time domain, the approach adopted is as listed below:

a) The auto correlation analysis of the historical data was carried out in order to determine whether the data generating the auto correlations are random, that is, if the auto correlation coefficients are not significantly different from zero, results obtained shows that the data are non-random, and this was confirmed with the use of the Box-Pierce Q-statistics, indicating that there is some pattern in the historical data.

b) The auto correlation coefficients of the historical data was plotted against time, with the 95 percent confidence limits, in order to confirm that there is actually some pattern in the data, and the nature of the pattern. It

was observed that there is some pattern and that the pattern repeats itself every six months alternately, with the occurence of positive - positive, or negative - negative peaks every twelve months, suggesting the existence of period to period as well as seasonal dependence of the data.

c) Stationarity was also sought for by examining the autocorrelation plot and it was observed that the autocorrelation coefficients somewhat fluctuate about the theoretical mean for a random auto correlated series of zero mean. It does not have a noticeable trend. It was, therefore concluded that the historical data are stationary but non-random. This conclusion was also based on the fact that the auto correlations were not significantly difference from zero and that it dropped within the 95 percent confidence limits after the second month. Since theoretically, the auto correlations of stationary data drop to zero, that is, are not significantly different from zero, after the second or third month. However, period to period differencing was performed in order to avoid the occurence of spurious autocorrelations.

d) The twelve monthly repetition of positive or negative auto correlation values suggest the presence of seasonality, and this was also filtered out through

seasonal differencing.

e) The Box-Cox transformation was performed on the historical data, and tested for randomness using the Q-statistics, the transformed data were also found to be nonrandom, the various identification stages mentioned earlier were performed and it was noted that there is some seasonal pattern which has a strong influence on the randomness of the data and stationarity was also observed, however, the fluctuation of the auto-correlation about the mean is not perfect, suggesting the possibility of spurious auto-correlations resulting in the non-randomness of the data.

f) In order to remove period to period and season to season non-stationarity, one short-differencing and one long-differencing were performed on the transformed data, and a new series was formed from the product of the two series due to one short-differencing and one long-differencing. The various tests mentioned earlier in sections 1.8(a), (b), and (c) were performed on the new series, and the new series was found to be stationary and random.

g) Partial auto correlation analysis was performed on the differenced-transformed data in order to identify an appropriate ARMA (Auto Regressive Moving Average) model

for the estimation of the water levels, an ARMA (1,1) was identified.

h) The parameters of the mathematical model were estimated for the condition which introduces the minimum mean square error (MSE), by the method of decomposition; validation of the model by the substitution of measured data revealed a general pattern of the estimated data dependence on the immediate past residual, a linear pattern was exhibited which was found to be unstable.

i) The peturbations resulting, in the instability of the system were traced to the index of the power transformation of the estimated data and the error term in the linear part of the mathematical model. Model modifications and necessary adjustments were carried out through mathematical modelling of the index, the error term and the estimation model as a whole. (Appendix 1, 2 and 3). The final mathematical model was validated and the results obtained were very satisfactory, except for the rainy season period when it is not able to account adequately for contributions due to storm events.

j) The appropriateness of the model to generate future values of the time series were also tested, through the tests for randomness, and all but one of the autocorrelations were found to lie within the 95 percent confidence limits and the computed chi-square (x^2) value

of 13.65 was also found to be less than the theoretical chi-square (Q-statistics) value of 63.981. These two tests confirmed the appropriateness of the model to estimate future values of the time series of water levels in Lokko River.

CHAPTER 2: LITERATURE REVIEW

2.1 GENERAL

Mathematical modelling is increasingly becoming an interesting area in hydrology and many research work has been done in this area. However, most of the research efforts have not addressed the crucial issue of estimating future rainfall amount on a short-term or long-term basis satisfactorily, such that the well developed models of run-off responses to rainfall input have not yielded in most cases comparable results with observed data. Because of the difficulties involved in estimating rainfall series most present day research efforts are directed towards modelling the run-off series. Bras and Rodrique - Iturbe (1976) developed a rain fall generator by assuming that storm has a basic structure.

The social-economic problems associated with flooding has made man to seek for ways to control it. In the course of doing this mathematicians, statisticians, engineers and hydrologists have developed mathematical models based on hydrological theory or statistical theory, or a combination of both. Nash and Sutcliffe (1970) noted that it is extremely difficult if not almost impossible to estimate future flood events, simply by applying the known physical laws to the measured rainfall and boundary conditions within a catchment

area, largely because of the complexity of the boundary conditions..pm

2.2 THE RATIONAL FORMULA

In view of the difficulties encountered in estimating floods, the largely empirical or analytical approach is often adopted, in which the areal precipitation is computed, the results obtained are used in calculating the direct run off using time distributed hydrograph, and if necessary employing flood routing procedure to estimate the change in shape of the hydrograph as it moves to a selected downstream location.

The first mathematical model for estimating floods was published by Mulvaney (1850-51) who developed the rational formula, and was modified by Darcy (1856) to simply

$$Q = C\ I\ A$$

where,

Q = Run off, cu.m per day

C = Coefficient of run off

I = Rainfall intensity, m per day

A = Catchment area of the basin, m^2.

Fuller (1914), developed a similar model with the parameters also basically measures of flood and drainage basin area.

These methods are largely in error in estimating flood events because they cannot account for the boundary conditions, and are basically as inadequate in extrapolating future events..pm

2.3 THE WATER - BALANCE MODEL

According to the development of rainfall-run off model by Meyer (1915), his procedure involved water balance calculations from precipitation, evapotranspiration, and soil storage to estimate monthly and annual flow volumes; This procedure is not flexible as it depends on a good knowledge of the physics of data acquisition and it can tell the future events if only future estimates of the inputs are available.

2.4 THE LINEAR MODEL

Sherman (1932), developed the unit hydrograph theory and it was the first expression of the linear approach in flow simulation.

In view of its wide acceptance more researches were conducted on the theory of the unit hydrograph and Snyder (1938) developed the synthetic unit hydrograph theory which could be used when needed flow data were not available, and this is the nearest attempt to estimation of future data during that period.

2.5 THE ROUTING MODEL

However, not long after, Clark (1945) demonstrated the equivalence of the linear routing procedure and showed that a unit hydrograph could be constructed by the routing of the time-area diagram of the basin converted to flow units. The technique of flood routing as described above was applied for many years therefore, in an attempt to model flood wave propagation.

The above technique can be modified by a consideration of the wedge-storage concept, this approach have been well tested and found to be satisfactorily; Rosso et al (1983) applied the technique to model the propagation of flood waves throughout the middle Shaballe River plain in East Africa. Results obtained were good with about 10 percent error at low flows and 5 percent error at peak flows. However, the performance of this model is based on a good choice of space and time steps which are usually very difficult to select, this coupled with the continuous calibration of the Muskingum parameter makes the model inflexible and unsuitable for estimating future events.

2.6 THE INFILTRATION MODEL

Horton (1933), barely a year after Sharman developed the unit graph theory developed the infiltration theory of ground water flow contribution to stream flow. The infiltration process outlined by Horton proved less immediately useful

than the unit hydrograph concept. However, the presentation offered for the first time a quantitative basis for considering the process by which rainfall is converted to run-off that feed the streams.

Going by present day technology, this method would yield very good results, but the development of the relevant data base would be too expensive in view of the problems associated with data acquisition. Another major problem with the consideration of the infiltration process is that when treated as a function of time, it did not properly account for changes in the infiltration capacity of the soil. These difficulties led Cook (1946) to the simplification of the infiltration approach by introducing infiltration indices which merely expressed an average loss rate.

2.7 THE HYDROLOGICAL SIMULATION MODEL

With the coming and availability of digital computers in mid 1950s, more daring research efforts were made, that Rockwood (1969), developed an elaborate simulation program for flood forecasting based on run off response to rainfall input; the results proved the method to be successful. Unit hydrograph analysis also expanded into "Linear Hydrograph Analysis", but this perpetuates the limitations of the linear assumptions in the unit hydrograph concept and shows little prospect of developing into a useful technique, for analysis

in hydrology. Multiple correlation studies of rainfall - run off process also proved rather fruitless. The rainfall - run off process is a complex one and is not easily represented by an equation involving a linear sum of terms or by any non-linear expression.

The approach therefore, was to attempt the hydrologic simulation. This procedure seeks to describe the rainfall runoff process by a series of mathematical functions, each describing a particular portion of the process and in combination simulating the whole process. The computer facilities made this possible, in that, the manual computation which was the usual approach consumed many man-hours and risk numerous errors, due to approximations, omissions and computational mistakes due to fatigue.

The first published work on hydrologic simulation was by Linsley and Crawford (1960), and after that a large number of models have been described in literature and they could be classified as linear, routing, event and continuous water-balance models. Sugawara (1961), proposed a model consisting of a complex cascade of linear storages for simulating the entire run-off process from rainfall to stream flow, the problem however was that he could not relate the various storages to features of the natural process, and hence his model can only be fitted to a real water shed by a lengthy trial-and-error process.

One of the earliest water balance model was the Stanford Water shed Model, developed by Linsley and Crowford (1960), this model has been improved upon several times by various authors, but they all follow the general concept of the hydrologic cycle as described by Linsley et al (1958). The model has been well tested and found to be satisfactory, but the calculations involved are enormous.

Noting the importance of accurate rainfall measurement to stream flow estimation, Johanson (1971) carried out a study which indicated that if too few segments of a water shed are used, the variance of the simulated flow will be too high, but the mean will be reasonably correct. This means that the rain gauge will on occasion receive the maximum rainfall within the catchment area, and since the rain gauge catch is assumed to be the catchment average, simulated flows will be too high, and vice versa. Further studies by Johanson, showed that the errors decreased rapidly as the number of rain gauges is increased; but not linearly with catchment area, due to differences in topography.

2.8 THE STATISTICAL MODEL

Parallel to the development of the infiltration approach, there were efforts to deal with the run-off problem on a statistical basis, Kholer and Linsley (1941) developed a coaxial method based primarily on a multi-variable graphical relationship. Betson et al (1969), subsequently demonstrated

that the same type of relationship could be developed analytically. The relationship obtained gave an approximation to the initial soil moisture, and was far superior to the use of infiltration indices, however, it did not perfectly explain the variability in storm run off nor could it provide information on future events.

Before the published work of Betson et al (1969), many works were carried out by various authors to overcome some of the abvious limitations, the first of which is acquisition of accurate and continuous data; The second and controlling factor was the limitations of manual computations. With the work of Hoyt (1942), the run off process was fairly well understood. Further works were done in this area by Linsley and Ackermann (1942), and by Linsley et al (1958). Kresge and Nordenson (1955) demonstrated that these procedures could be used to extend flow data base for frequency analysis.

While research efforts were still on in areas of experimental and analytical/graphical methods of stream flow modelling, statistical methods were also being evolved and in the 1960's major achievements were made, and this was possible with the availability of digital computations. The critical constraint in stochastic (or statistical)

methods lies in the fact that stochastically generated flows are only as reliable as the parameters of the distributions used in their generation. As yet, appropriate specification for reliability are only imperfectly known. For example, Burges (1970), studied the record of Arroyo Seco in California and found that 33 years sample gave notably different parameters than the 65 years record, and Ott (1971) found that a record length of nearly 100 years was necessary on Dry Creek in California and Fisher river in North Carolina to obtain stable values of the mean and variances.

By the close of the 1960s there were basically two approaches to the development of dynamic models for flow estimation namely the one based on theory of hydrology and the other based on statistical estimation theory. Researchers from the 1970s seem to favour the statistical estimation theory since the results obtained are highly reliable and are only dependent on the data from which the estimation model was developed. Efforts were therefore made to replace most of the existing gauges in the network systems with computerized radar controlled systems that could collect and process incoming data in fractions of a second and store the needed information. This provided reliable data base from which further researches in real-time forecasting evolved.

The system based on the theory of hydrology, though is still very much in practice, it is however, not receiving much attention these days, in view of the problems involved in the data collection for its development, the complexities in the development of the model, the complex nature of the model, and the difficulties in the use of the model. What is of utmost concern is the generation of artificial sequences of a single variable for use where precipitation sequences are to be generated for routing through a rainfall-run off model, or for use where stream flow at a gauging site is to be generated to estimate the frequency with which extreme events occurs or forecast when certain events would occur.

In the light of the above and with bias to stream flow generation which is the purpose of this work, references are hereby made to the works of accredited authors in this area. The generation of annual stream flow was perharps first studied comprehensively by Hurst (1951) and (1956). He was primarily concerned with the design of storage reservoirs on the River Nile. Hurst quantified observed persistence in terms of a parameter h, the average value of which he found to be approximately 0.73, for a very large number of time series. And which he noted to be in distinct disagreement with the value of h equal to 0.5 synonymous with an independent random process. Early attempts at generating

stream flows assumed that the observed persistence was essentially short-termed and could be represented by Markovian auto correlations, Thomas and Fiering (1962). However, this assumption does not appear to agree with the findings of Matalas and Huzzen (1967).

In a significant contribution to synthetic hydrology, Mandelbrot and Wallis (1969a, b, c, d, e) contended that Hurst's findings could not be explained by Markovian auto correlation and proposed Fractional Gaussian Noise instead as a model with a suitable theoretical basis. The autocorrelation function of fractional Gaussian noise, governed by the parameter h, implies the existence of a long term memory, giving rise to the phenomenon of long term persistence, the intensity of which increases as the parameter h increases from 0.5 to 1.0. Markovian auto correlation is synonymous with a value of h of 0.5.

Fractional Gaussian noise being a process with an infinite memory, finite memory approximations to the fractional Gaussian noise are required for generating synthetic flows on a digital computer. Recent approximations to the fractional Gaussian noise include fast fractional Gaussian noise, Mandelbort and Wallis (1971). ARIMA models, Box and Jenkins (1970) as applied by O'connell (1971, 1974); filtered fractional Gaussian noise, Matalas and Wallis (1971); and the broken line process, Mejia et al (1972). Of

these approximations, ARIMA processes are relatively easily applied in generating synthetic flows.

2.9 THE TIME SERIES MODEL

The systematic development of time series analysis started with the work of Yule (1927), who showed that Wolfer's sunspot data could be represented satisfactorily by a second-order auto regressive (AR) model. Wold (1938) showed that any non deterministic stationary time series can be decomposed into a moving average of independent random variable known as "innovations" or the one-step-ahead predictor error. The estimation of parameters in AR and other linear difference equation models was considered in depth by Mann and Wald (1943); and Whittle, in a series of ground-breaking papers (1951, 1952, 1952a, 1954, 1954a, 1954b) explored various aspects of time series analysis, including spectral representation, maximum likelihood estimation of parameters, hypothesis testing, test of fit, and random process on a plane.

From 1950 to 1970 several developments took place in the field of spectral estimation and non-parametric modelling of time series, they are reported by Bartlett (1966), Hannan (1960), Blackman and Tukey (1959), Parzen (1957), and Jenkins and Watts (1968). The importance and usefulness of parametric modelling was re-emphasized by Box-Jenkins (1970),

who also showed that a number of practical non-stationary time series i.e. data when plotted will not lie parallel to the time series could be modelled by using integrated and seasonal auto regressive moving average (ARMA) models.

With the exception of the work of Quenouille (1949), the literature of time series analysis until 1970 was largely confined to scalar time series. Since 1970, however, multiple time series analysis has received increasing attention in both the statistical and control literature, Hannan (1970), Akaike (1974), Kashyap and Rao (1976). Akaike (1974) discussed developed canonical models for multi-input, multi-output (MIMO) systems, while multiple ARMA models are discussed by Hannan (1970) and by Kashyap and Rao (1976). The single-input, single-output ARMA model is identifyable, but the multiple ARMA model requires restrictions on parameters to be identified.

Jamieson and Wilkinson (1972), suggested the use of a first - order auto regressive model as a tool for short-time data estimation, but found that such a stationary model explained only 45 percent of the rainfall time series, other time series such as MA, ARMA, Box and Jenkins (1970), MARMA etc., can be used, the main limitation is that they have no parameters representing real conditions. Their parameters must be satisfactorily determined from data. Therefore, they do not represent well incidences of convective storms which

results in high flood level or flash flood effects.

Meteorologists have only imperfect and very complex deterministic models based on atmospheric conditions. They are more an exercise to study atmospheric interactions of climatic variables than rainfall predictors. Their parameters and structure would not lend themselves to efficient on-line modification resulting from observed data. It can therefore be seen that rainfall prediction is a very hard one. In an attempt to resolve the dilema Bras and Rodriques-Iturbe (1976) suggested a non-stationary, time-varying, multi-dimensional rainfall generator that preserves first - and second - order statistics in time and space, and assume that storms have a basic structure. The model was applied by Wilson (1976) and the results were found to be very satisfactory.

The first work to be reported on estimation of a time series with adaptive filtering was by Makridakis and Wheelwright (1978). Long (1975) extended the concept of the method to include mixed ARMA models that led to the generalization of the method so it could deal with all types of data and processes. In another paper by Makridakis and Wheelwright (1978), the concept was extended to sequential ARIMA models and non-stationary time series.

2.10 CONCLUDING REMARKS

The importance of modelling in water resources engineering cannot be over emphasized. McCuen (1975), in building a case for modelling pointed out that models provide a means of improving decision - making capabilities and the basis for improved operations.

CHAPTER 3: EXISTING DRAINAGE AND STORAGE PATTERN

3.1 GENERAL

The Lokko River catchment area is well connected with stream or river channels, most of which are ephemeral (fig.2). To the North-West is the Ziomo stream (Mayel Ziomo) into which flows numerous other streams, the Ziomo stream flow South wards and meet Gondor stream which flows from the North-East to South-West direction. Kuntu stream and numerousother streams from this North-West region flows into River Sensen (Mayo Sensen). The Gondor stream and numerous other streams all flow into River Sensen. The River Sensen flows initially from the North-West to the South-West direction from longitude $12^{o}18'E$ to $12^{o}26'E$ at latitude $9^{o}52'N$ after which it flows South-wards to latitude $9^{o}47'N$ when it turns and flows in the South-West direction. The River Sensen drains into River Lokko North-West of Gangir Karemi village. The River Lokko (or Lokko) flows from the Western end of the catchment area and in the South-West direction.

Flowing into River Lokko are numerous other streams which drains the hilly terrain during the wet season, but are generally dry during the dry season, and the Lokko River flows through the flood plain which runs in the East-West direction. The total stretch of the flood plain from

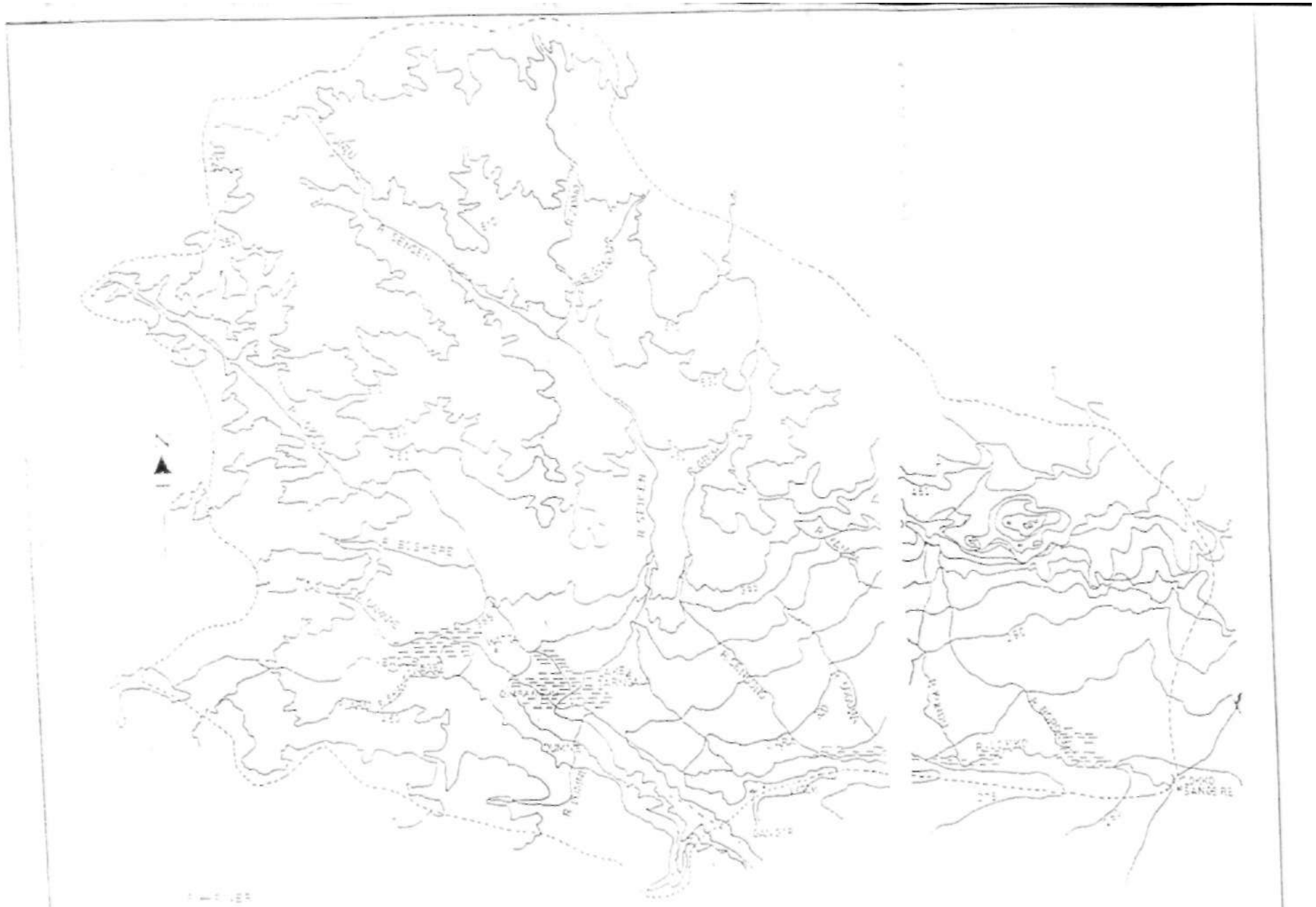


TABLE 1: Average velocities of flow in a basin for concentration time derivation.

Basin Character	Flat	Mildly rolling	Hilly	Highlands	Mountains
	Average Basin Slope				
	0.5 percent	2 percent	5 percent	10 percent	30 percent
	Velocity,m/s				
Swampy	0.07	0.08	0.30	-	-
Forest	0.12	0.20	0.50	0.80	1.20
Grassy Pastures	0.20	0.50	0.80	1.20	2.00
Gently sloping valley	0.40	0.70	1.00	1.60	2.50
Steep valley	-	-	1.20	2.20	4.00
Rocky steep cliffs	-	-	-	3.00	5.00

(Taken from Engineering Hydrology by Nemec,J. (1964))

Balubulla Baba village to Sangare Lokko (or Lokko village) is about 27 kilometers. Notable among the tributaries of river Lokko are Boshere stream, which runs Eastwards from the South-Western end, Damare stream, Loma stream and Shahera stream also run Eastwards from the South-Westerm end of the catchment area. Along with numerous other streams, some of them draining Northwards, drains into River Lokko, River Dumne drains Northwards and flow into River Lokko also in the West-East direction.

Other streams which drain into River Lokko include Gbim stream located North of the catchment area, it flows Southwards and into River Senguno, which flows in the general North West - South East direction into river Lokko. River Yamba, Sogura, Kundu, and Borongo in the South-Eastern part of the catchment area all flow Southwards in the genral North West-South East direction. There are no artificial drainagechannels in the study area as no engineering work had earlier been done except for the bridge construction; therefore the drainage pattern described so far are all natural. It will be observed that the catchment area drains generally from the North West to the South East ending at Lokko at the flow outlet from the catchment area, and this shows the general orientation of the slope of the land, needed in drainage channel design.

The natural storage pattern which are the only ones available include, the underground storage, Lake Tarngala, located some 2.5 kilometers North-East of Dumne village, the flood plain (shown in dashed lines, Figure 3) which runs in the East-West direction from Badubulla Baba to Watta village the flood plain running from Gwaragunda to Dumne in the North-West to South-East direction, the flood plain running in the East-West direction from Dara to Garin Lamba, and the adjacent low lying lands to the flood plains covering over 33 squares kilometers, these storages can be developed to slow down flood wave propagation down the slope, to effect controlled flow through the channels thereby reducing the level of inundation and for other beneficial uses.

Results of the analysis of the geomorphologic features of the basin are presented below:

3.2 AVERAGE WIDTH OF BASIN

The average width of the basin was computed from the expression

$$b_w = \frac{\text{Total catchment Area, A}}{\text{Average basin length, L}} \quad \text{.....(1)}$$

where,

Total catchment Area, A = 862km^2

Average basin length, L = 62.94km

hence

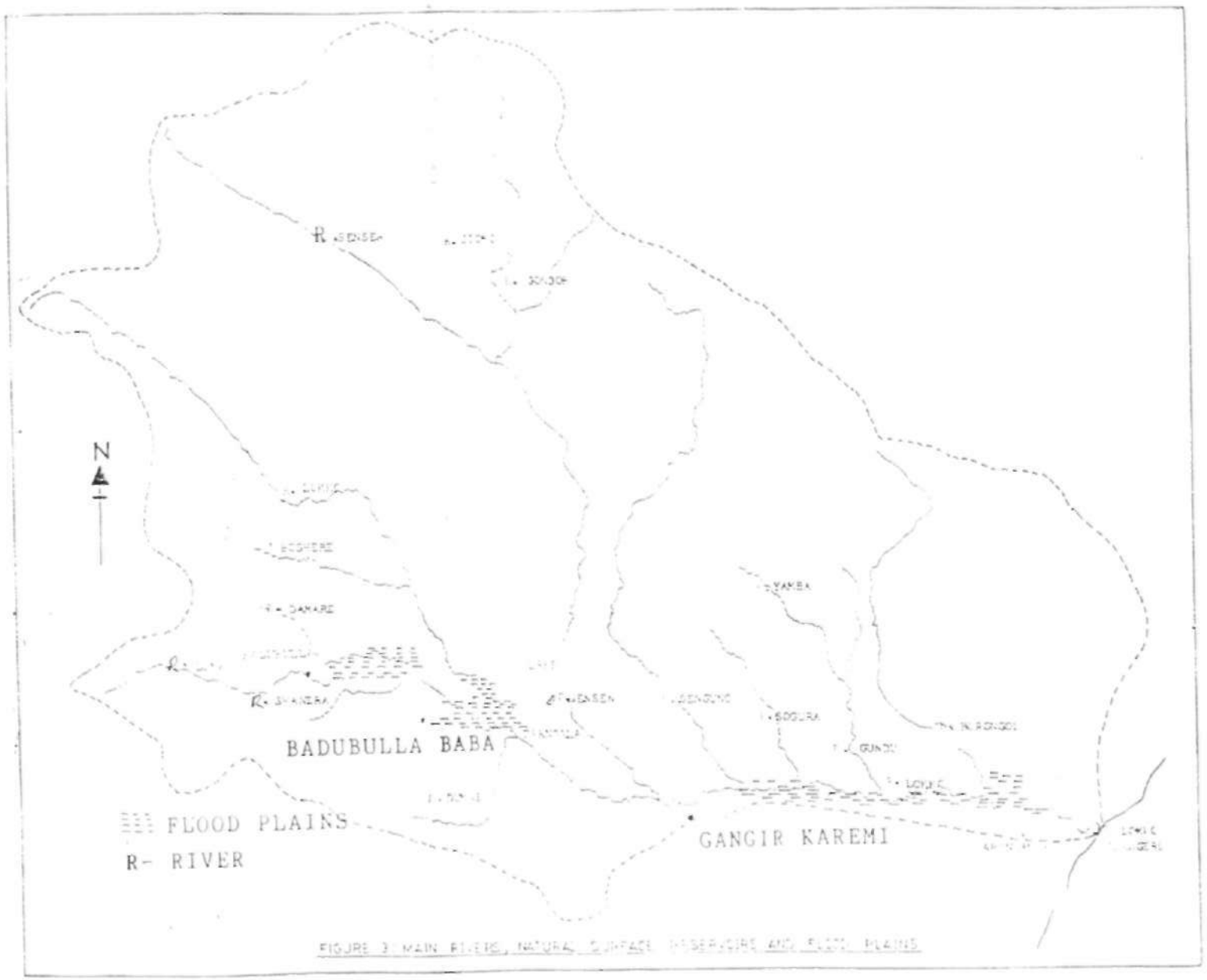


FIGURE 3 MAIN RIVERS, NATURAL SURFACE RESERVOIRS AND FLOOD PLAINS

$$bw = \frac{862 \text{ km}}{62.94}$$
$$= 13.70\text{km}$$

3.3 COEFFICEINT OF BASIN SHAPE

The coefficient of basin shape was computed from the expression

$$d = \frac{(\text{Average basin length, L})^2}{\text{Total catchment Area, A}} \quad \ldots\ldots\ldots\ldots(2)$$

$$d = \frac{62.94^2}{862} = 4.6$$

which shows that the basin is a fairly narrow one and it is prone to flash floods.

3.4 AVERAGE SLOPE OF THE BASIN AND FLOOD PLAIN

The average slope of the basin was computed from the expression

$$s = \frac{hmax - hmin}{\sqrt{A}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

where,

h_{max} = Highest point in the basin
= 762m

h_{min} = Lowest point in the basin
= 253m

hence,

$$s = \frac{0.762 - 0.253}{\sqrt{862}} \times 100 \text{ percent}$$

= 1.73 percent

Similarly the average slope of the flood plain located in the Southern part of the catchment area was computed from expression (3) above, where

$$h_{max} = 381m$$

$$h_{min} = 253m$$

hence,

$$S_F = \frac{0.381 - 0.253}{\sqrt{28.45}} \times 100 \text{ percent}$$

$$= 2.40 \text{ percent}$$

The average slope of the flood plain is higher than that of the entire basin, because of the small total area of the flood plain relative to that of the entire basin.

The average slope of the basin of 1.73 percent identifies it as a mildly rolling basin, with rocky steep cliffs and, having an average velocity of flow of about 1.0 meter per second, (Table 1).

The average slope of the flood plain of 2.40 percent identifies it also as mildly rolling and, having an average velocity of flow of about 0.76 metre per second.

The basin slope is the combined effect of the steep slopes in the hilly portion (Figure 2), where the stream flow is more straight than in the flood plain, where the slopes are gentle and the streams (or rivers) flow is generally sinnous and slower than in the hilly portion where the flow is swift and is generally of the shooting (super critical)

variety, and usually of high erosive power, eroding the sides of the channel and resulting in the sinuosity observed. The eroded materials are deposited on the channel floor, thus resulting in the rise of the channel floor while the river eats its way into the hillside. During intense rainfall flash flood may occur in this hilly region, resulting in the transportation of this bedload and other eroded materials to the flood plain, there by reducing the natural reservior capacity of the flood plain.

3.5 TIME OF CONCENTRATION.

Using the average velocity of flow of 1.0m per second for the entire basin, (Section 3.4), the time of concentration was computed using the formula suggested by Sokolovsky (Nemec, 1964), which is widely used. The formula is given as

$$t = \frac{L}{3.6v} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

where,

t = Time of concentration

L = The stream length or average basin length
= 62.94km.

V = Average velocity of flow

hence,

$$t = \frac{62.94}{3.6 \times 1.0}$$

= 17.48 hrs

meaning that it takes about 3/4 of a day for flood to manifest from a heavy rainfall input.

3.6 NATURE OF RIVER LOKKO.

From the discussions so far, it can be concluded that River Lokko is fully mature, one can discern a sequence of three tracts. At its head is the mountain (or hilly or torrent) tract which is either a gorge or V-shaped with angles from 30 to 90 degrees, in the middle of its course is the valley tract: the gorge has opened up, the slopes are gentler and the valley is wider, this opens up to a plain tract, at the bottom of the valley where there is the flood plain, which, as its name implies, gets flooded whenever there is a high discharge of water. A cross-section of the flood plain would show fine, horizontally stratified deposits which provide storage. The rivers are braided, representing bifurcating flow (or network) with islands in between.

3.7 NUMERICAL DESCRIPTION OF THE FEATURES OF THE BASIN.

In order to describe the features of the basin in numerical terms the following parameters which are characteristics of the drainage system were used:

a) The stream order

The Strahler (1957) stream ordering system was used, which says that, when two first order streams meet, they

form a second order streams, when two second order streams meet, they form a third order stream etc. However, when an N-th order stream meets a link of order P < N, the resulting stream will have order N, that is, the lower-order link gets "Swallowed up". This statement can be put mathematically as

$$\left.\begin{aligned} N * P &= N + 1 \quad \text{if } N = P \\ N * P &= \text{Sup}(N,P) \quad \text{if } N \neq P \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(5)$$

where * stands for the meeting of the streams of order N and P, and "Sup" refers to the one of superior order. From the stream ordering, both the Lokko River and Sensen River are of order 5, and are also the rivers with the highest order in the catchment area under study. The two rivers therefore form the two main and most important rivers of the catchment, and accoring to P.F.P. (Nig.) et al(1980), are the only rivers that have some persistent flow in the study area, though not significant.

b) Bifuration Ratio

Since Horton (1945) and Strahler (1957) streams have the same bifurcation ratio, we have using Horton's law of stream numbers, that

$$R^b = \frac{N_i - 1}{N_i} \text{ for any order i} \quad \ldots\ldots\ldots\ldots(6)$$

where

R_b = Bifurcation ratio

N_i = Number of streams of order i

from which

$N_5 = 2$

$N_4 = 4$

$N_3 = 8$

Hence,

$$R_b = \frac{N_4}{N_5} = \frac{N_3}{N_4} = \frac{4}{2} \text{ (or } \frac{8}{4})$$

which says that there are two streams of lower order to every one of the next higher order.

It was therefore concluded that there are 62 streams in the study area.

c) The stream length

According to Horton's law of stream length

$$R^L = \frac{Li}{Li-1} \text{ for any order i} \quad \ldots\ldots\ldots\ldots\ldots\ldots(7)$$

where,

R_L = stream length ratio

Li-1 = Average length of stream of order i-1

For a stream of order 5,

$$L_5 = \frac{Ls + Lk}{2} \quad \text{i.e., i = 5}$$

where,

L_s = Length of River Sensen = 40.68km

L_k = Length of River Lokko = 62.94km

Therefore,

$$L_5 = \frac{40.68 + 62.94}{2}$$

$$= 51.81km$$

Similarly,

$$L_4 = \frac{L_{sg} + L_{us} + L_{uk} + L_{Gr}}{4}$$

where

L_{sg} = Length of river Senguno = 21.31km

L_{us} = Length of Upper River Sensen = 17.56km

L_{uk} = Length of upper river Lokko = 22.81km

L_{Gr} = Length of river Garwa = 17.63km

Therefore,

$$L_4 = \frac{21.31 + 17.56 + 22.81 + 17.63}{4}$$

$$= 19.84km$$

Hence, the stream length ratio is,

$$R_L = \frac{L_5}{L_4} = \frac{51.81}{19.84} = 2.61$$

Therefore, the total length of streams in the basin is 326.26 kilometers.

d) <u>The Stream Area</u>

According to Strahler's law of stream areas

$$Ra = \frac{A^i}{A^i_{-1}} \text{ for any order i} \quad \ldots\ldots\ldots\ldots\ldots\ldots(8)$$

where,

R_a = stream area ratio

A_{i-1} = Average area of catchment of streams of order i-1

Therefore,

$$A^5 = \frac{A_k + A_s}{2}$$

where,

A_k = Area of Lokko River sub-catchment

= 565.69km^2

A_s = Area of Sensen River sub-catchment

= 296.31km^2

Hence,

$$A5 = \frac{565.69 + 296.31}{2}$$

= 431km^2

Similarly,

$$A_4 = \frac{A_{sg} + A_{us} + A_{uk} + A_{Gr}}{4}$$

where,

A_{sg} = Area of Senguno River catchment = 128.50km^2

A_{us} = Area of upper Sensen River sub-catchment

= 197.2km^2

A_{uk} = Area of upper Lokko River sub-catchment

= 322.96km^2

A_{Gr} = Area of Garwa River catchment = 171.60km^2

128.50 + 322.96 + 197.20 + 171.60

$$A_4 = \frac{\qquad\qquad\qquad\qquad\qquad\qquad}{4}$$

$$= 205.065km^2$$

Hence, the stream area ratio is,

$$R_a = \frac{431}{205.065} = 2.10$$

This means that the catchment area of a stream in the study area is about 2.10 times the catchment area of a stream of the next lower order.

For a catchment of order u, where u = 5 for Lokko River,

$$L(u) \sum_{i=1} \quad \sum_{i=1} L_i = L_1 R_b^{u-1} \frac{(r_L^{u-1})}{r_L - 1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (9)$$

with,

$$r_L = \frac{R_L}{R_b}$$

where,

$$L(u) = L(5) = L_5 = 51.81km$$

$L_1 = ?$

$R_b = 2.00$

$R_L = 2.61$

Then,

$$r_L \quad \frac{R_L = 2.61}{R_b \quad 2.00} = 1.305$$

Hence,

$$(1.305^5 - 1)$$

$$51.81 = L_1.2^4. \frac{}{1.305 - 1}$$

$$= L_1.16. \frac{(3.785 - 1)}{1.305 - 1}$$

$$51.81 = L_1.16 \frac{(2.785)}{(0.305)}$$

$$= 146.1L_1$$

From which we have,

$L_1 = 0.355km$

= Average length of stream of order 1

Similarly,

$$A(u) = A_1 R_b^{u-1} \frac{(r_L^u - 1)}{r_L - 1}$$

where,

$A(u) = A(5)$ = Total catchment area

$= 862km^2$

A_1 = catchment area of order 1

$$862 = A_1.2^4 \frac{(2.785)}{0.305}$$

$$= 146.1\ A_1$$

From which we have,

$A_1 = 5.90km^2$

= Average catchment area of stream of order 1

e) Drainage Density

The drainage density was computed from the expression

$$D(u) = \frac{\sum_1^n \sum_1^n L(i)}{A(u)} = \frac{L(u)}{A(u)} \quad \ldots\ldots\ldots\ldots\ldots\ldots (10)$$

$$\frac{A(u)}{}$$

Such that,

$$D(u) = \frac{51.81\ km^{-1}}{862}$$

$$= 0.06km^{-1}$$

The low value of $D(u) = 0.06km^{-1}$ is smaller than $1.50km^{-1}$ for Sandstone, it follws therefore that the drainage basin is made up of massive Sandstone.

f) Average length of overland flow

The average length of over land flow was computed from the expression,

$$L_{of} = \frac{1}{D(u)}$$

$$= \frac{1}{0.06} = 16.67km$$

This is approximately equal to half the distance between the stream channels. Since the average width of the basin is 13.70km (section 3.2), if follows that the flood water drains directly into the stream channels, creating a network through which water flows into permanent streams, permanent in the sense that they are also fed by sub-surface and probably ground water flow.

3.8 CONCLUDING REMARKS

The average length of overland flow of 16.67km compared

to the basin width of 13.70km indicates that only a small proportion of the flood water has enough time to percolate into the soil in the lateral flow across the basin, however, in the longitudinal flow the basin length is 62.94km, with a time of concentration, of 17.48hrs, it follows that a lot of the flood waters get enough time to percolate through the massive sandstone formation before they reach the outlet.

CHAPTER 4: BASIC DATA.

4.1 GENERAL

Detailed basic data on the study area are scarce or are not available, however, attempts have been made to extract information on climate, precipitation, temperature, wind, radiation, air humidity, water budget, physiography, geology, geomorphology, land system and land use from earlier studies reports prepared by Dar al Handasah (1980), SOTESA (1980), Land Resources Division Commonwealth Overseas Development Administration (1972) and consullint (1975), that are relevant to the project area, and the data were collected from Upper Benue River Basin Development Authority, Yola.

4.2 CLIMATE

Generally, the climate of the area is highly dependent on the properties and movement of the Intertropical Convergence Zone (ITCZ), Tuley (1972), whose characteristic behaviour is still being studied by researchers.

The Northward progression of the sun during the dry season in the Northern hemisphere brings rain bearings South Westerly winds to the study area during the month of April when the first rain of the year is usually experienced. This rainfall lasts till October, during which period water for plant growth becomes available under conditions of increasing day lengths. In only two months is there an excess of rainfall over potential evaportranspiration, these are the

months of August and September. According to Cocheme and Franquin (1967), the rainfall pattern is highly correlated with latitute. The poor rainfall and its uncertainty in the early wet season, June and July, are a severe constraint to rainfall agriculture in most parts of the area.

The low-lying plains experience a slightly warmer and more humid climate and are more suitable for irrigation. For irrigated agriculture, the most important feature of the climate other than the rainy period is the hot spell centred on March when the daily maximum temperature exceeds 40^{o}C. This is believed to be due to orographic effects, as the area is hilly and of spares vegetation, heat radiation effects will be high. This temperature of 40^{o}C can adversely affect the growth and yield of crop.

Contrary to the movement of the sun during the dry season, the Southward movement of the sun during the rainy season, cause the Northernly and the North Easterly dry winds to blow from the Sahara. Beginning from late October, this dry wind is referred to as the "Harmattan"; It brings cold environment to the entire region which increases in intensity sometimes forcing the minimum temperature down to 17^{o}C which however is lower than the minimum temperature observed in some of the more Northern parts of Nigeria, which is partly because of the area's low altitute and proximity to the River

Benue.

4.3 PRECIPITATION.

Analysis of daily rainfall records for 7 years (1983-89) available at the hydro - meteorological unit of the Upper Benue River Basin Authority Yola, Nigeria, compared with reports as recorded by Tuley (1972) are in good agreement. The records shows that the annual rainfall series exhibits cyclicity, with a seasonal length of 3 years, being the period between two troughs or crests, and there is evidence of trend as the data does not fluctuate around a constant mean, either for the annual mean or for the rainy season average. The mean annual rainfall depth is 782mm, while the rainy season average in this locality is as high as 1245mm, such that high value rainy season records of 1,016 - 1,270mm is very probable. Maximum precipitation depth per annum could be up to 1550mm and could be down to a minimum of 600mm, and there is no year without rainfall.

According to Cocheme' and Franquin (1967), the precipitation pattern in the study area and environs is not with the typical longitudinal trend, dips are observed in isohyet due to general circulation and some local Southward displacement. Cocheme' and Franquin (1967), further noted that the study area which lies between longitude $12^{O}15'E$ and $12^{O}36'E$, and latitude $9^{O}44'N$ and $10^{O}09N'$ is roughly enclosed

by a more general climatic rectangular bound within longitude 10^{o}–13^{o}E and latitude 9^{o}–11^{o}N, in which lies a distinct isolated belt, of 1015mm mean annual rainfall, about 30–50km wide running East-West direction. In addition, it is beleived that localized Oragraphic and Convective effects are responsible for the recorded flood events.

The monthly rainfall series exhibit seasonality with recorded minimum peak of approximately 160mm re-occuring about every 35 months (that is approximately 3 years) bounding two maxima of about 275mm, with the peaks increasing greatly initially and then slightly from the minimum to the maximum, indicating the probable occurence of flood events to be every 3 years. The monthly rainfall mean depth is 65.15mm and it is 103.23mm for the rainy months. The peaks are generally reached in the month of August (70 percent of the time), and are occasionally reached in the months of June or July, an indication that flood events are most likely to be observed in the month of August and Seldom in the month of June or July, which might be attained in times of early rains and early dry season. In some months of intense rainfall, the total depth could be as low as 2mm in some dry wet season month. The average of rainfall months is 7.6 months, with 8 months being more common at 70 percent.

Between June 26, and July 10, in addition to the more frequent rainfalls there occur at least two heavy rains which wets the soil significantly for planting. Between July 29, and August 5, at least one intense rainfall do occur sometimes up to 70mm, it abates again until between August 16, and September 12, when at least two intense rainfall would occur in quick succession, usually more than 40mm; the floods usually appears during this period.

The precipitation which was reported by the Weekly Scope of September 3,1989, which caused devastating floods in Lokko village is insignificant the earlier rains were of September 1, of 52.9mm depth and September 2, of 2.3mm depth, others are of August 27, of 80.00mm depth and August 28, of 25.4mm depth, indicating therefore that the floods lag behind precipitation by at least 2 days. The actual flooding occured on September 1, 1989 which is earlier than the date recorded by the weekly scope news papers of September 3,1989.

4.4 RUNOFF WATER LEVEL.

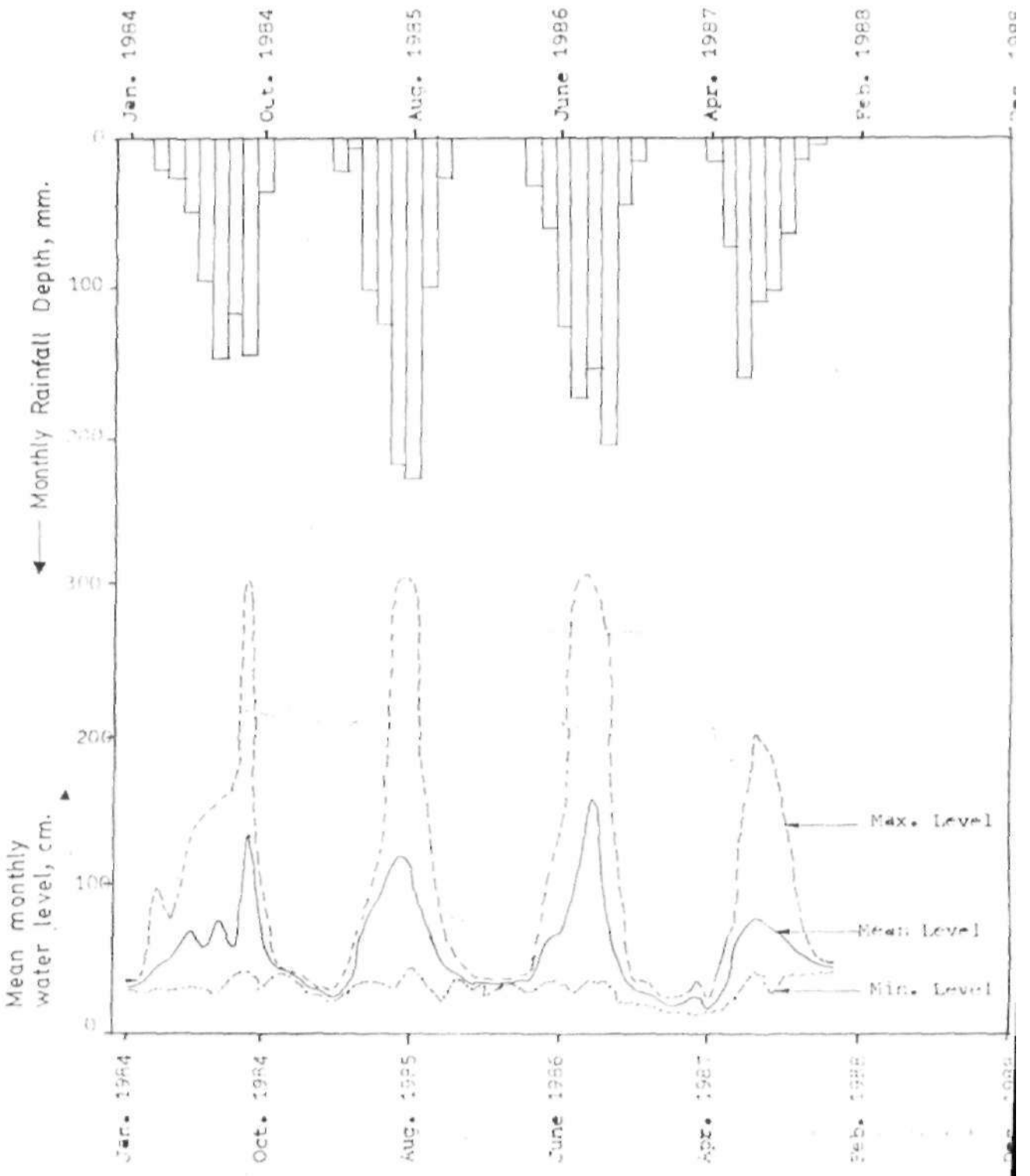
Analysis of the run off water levels for the period 1984-87, shows that, the monthly water levels vary from month to month, reaching its peak during the periods of the heaviest rainfall, which usually falls within the months of

July and August, sometimes high precipitation depth may be observed in the months of June or September (Figure 4). This figure shows the response function of the water level to rainfall input on monthly basis. It will be observed that there is a strong evidence of cyclicity in the peaks, as the flow is generally low in the months of January and February when there are no rainfall, but at the instance of precipitation the curve rises steadily, only dropping when there are dry spells and to rise as soon as rainfalls. This rise continues until it reaches its maximum, either in July, August or September, after which it drops steadily to some minimum level, as there are no more rainfall input.

At the start of the next rainy season it rises again in response to the rainfall input, thus indicating the presence of seasonality, the seasonal length estimated, figure 4, is about 12 months. Some trend can also be observed in the data as the curve does not oscillate perfectly about a mean in the horizontal direction. It should be noted however, that the water level is not zero at any time of the year, which is a strong indication of ground water contributions.

P.F.P (Nig).et al (1980) noted that only Rivers Sensen and Lokko have some flow during the dry season indicating the presence of a fairly high ground water table in the basin, as Rivers Sensen and Lokko are the only major rivers in the basin.

FIGURE 4 : HISTOGRAM OF MONTHLY RAINFALL



HYDROGRAPH OF MONTHLY WATER LEVEL

Throughout the period under investigation, the minimum water level observed in Lokko river at the gauging station was about 17.4cm (0.174m), while the maximum was about 300cm (3.000m). Figure 4, also shows the plot of the maximum water level and minimum water level observed for the periods, 1984 to 1987. The mean monthly water level was computed to be 55.13cm (0.5513m), with a standard deviation of 30.17cm (0.3017m), this gave a 95 percent confidence interval of 55.13 ± 8.78cm, that is, 46.35 cm to 63.90cm to the mean value, indicating large variations in the mean values of the monthly run off water level data.

Daily water level is generally low from January through March, figure 5, of the order of 14cm (0.14m) to 33cm (0.33m) and may persist till April or May during years of late rains.

The water levels rises after the first rain which may occur in March or April, but usually fall back after a day or two as the rainfall generally is not persistent at this time. With another rainfall, the level rises again and drops, this process repeats itself until there is sufficient rainfall to recharge the rivers at more regular intervals, this usually occurs in July and lasts till September, during this period there is a wide variation in the daily water level, sometimes reaching peak values of 300cm (3.000m) as well as dropping down to values as low as 30cm (0.300m); All the streams and tributaries during this period generally have some flow,

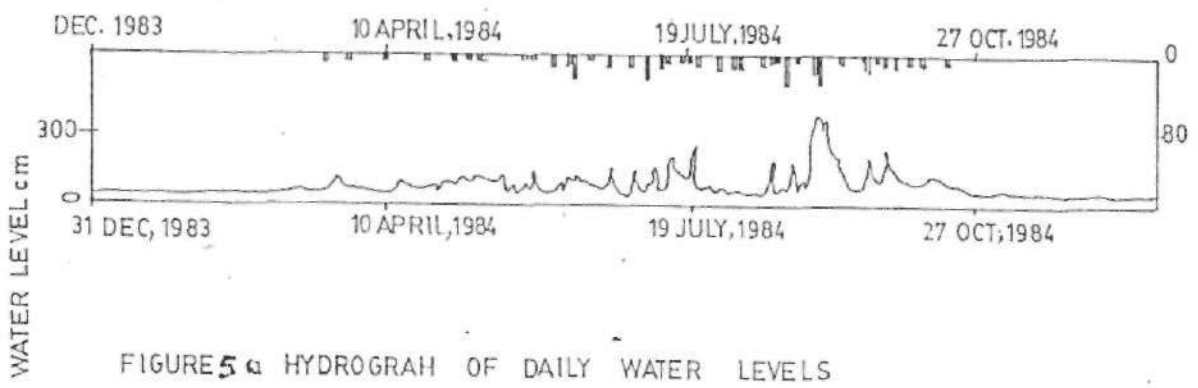


FIGURE 5a HYDROGRAH OF DAILY WATER LEVELS

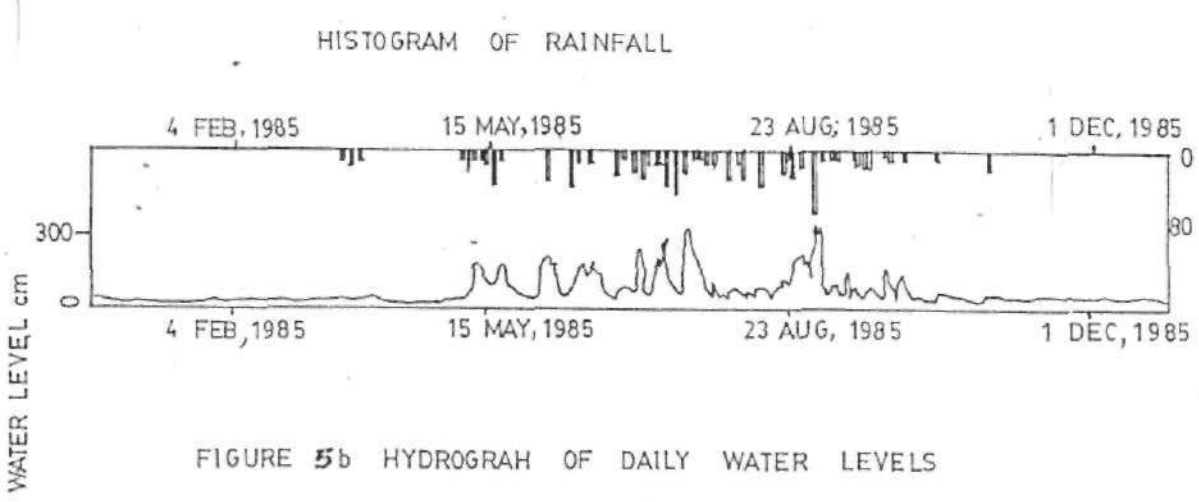


FIGURE 5b HYDROGRAH OF DAILY WATER LEVELS

Figure 5: Rainfall and River water Level Record Charts

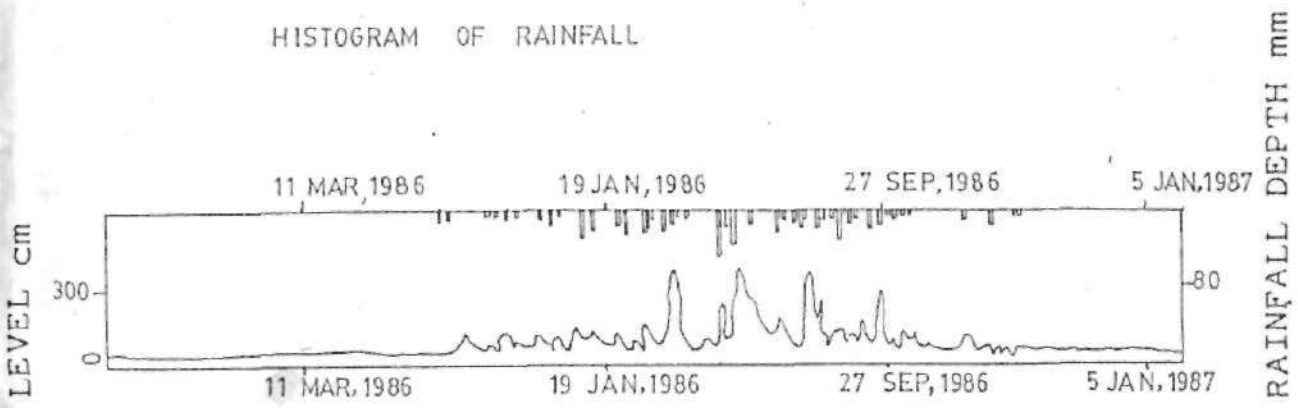


FIGURE 5c HYDROGRAH OF DAILY WATER LEVELS

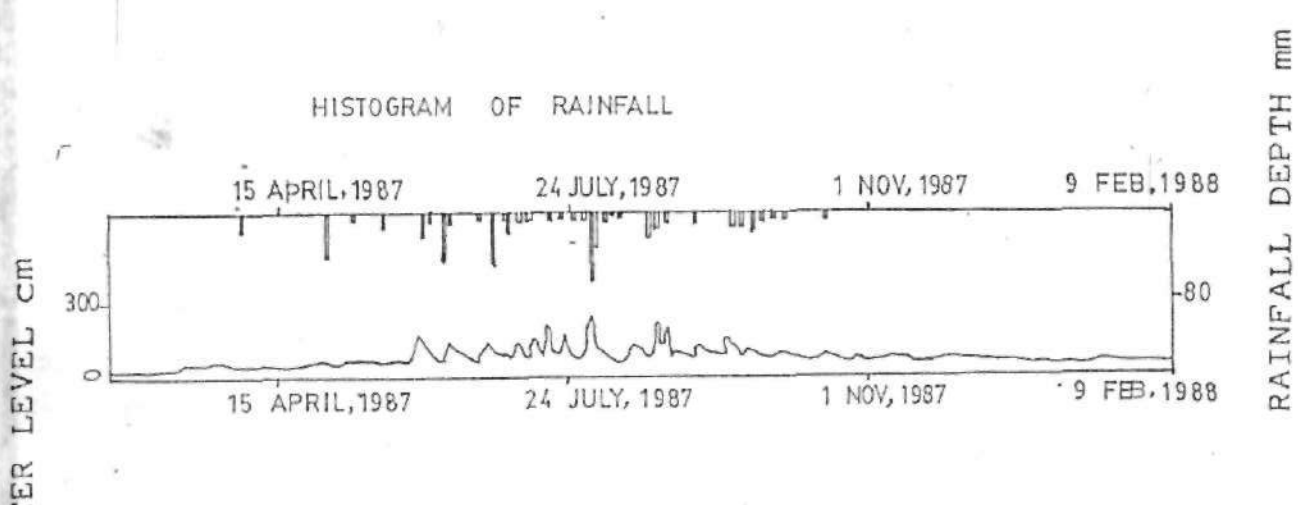


FIGURE 5d HYDROGRAH OF DAILY WATER LEVELS

indicating a rise in water table during the rainy season.

Measurement of some water level (at least 14cm (0.14m)) everyday of the year is a strong indication of the presence of a high water table, contributing directly to the flow during the wet season. However, the flow could be in the reverse direction (that is, from the river into the aquifer during the wet season), this depends on the relative water level in the rivers to that of the ground water table.

During the dry season, the entire flow can be assumed to be from the underground storage, however, this would need to be investigated in another study. The daily mean water level for the period of 1984 to 1987 was computed and found to be 43.9cm (0,437m) with a standard deviation of 43.37cm (0.4339m); giving a 95 percent confidence limit of 43.90 ± 0.06cm (or 43.84cm to 43.96cm). The large value of the standard deviation obtained is an indication of the randomness in the observed data.

4.5 TEMPERATURE

Limited data on the temperature regime of the study area are available; According to Klinkenberg et al (1963) in their studies on temperatures from recording stations in Gombi 38km North of Lokko, and the Upper Benue River Basin Development Authority temperature recording station; the broad trend of the temperature regime shows that temperatures are highest in April when it goes up to 31° to 33°C and lowest in January

when it comes down to 14° to 18°C. March to May are uncomfortably hot, June to October are warm with the rains, November to February are cold with monthly mean falling to within 20° to 22.5°C.

Night temperatures of the order of 10°C are sometimes recorded in this area during the cold period, and on high altitudes such as the Ghaanda hills lower temperatures may be experienced following the theory of temperature lapse rate or vertical temperature gradient, which in the troposhere according to Wiesner(1970) approximates to a decrease of 3.5°F for every 304.8m or 6°C per kilometer. In April mean maximum temperature over the area are in excess of 39°C while mean temperatures are of the order of 31°C. During the rains an intermediate regime with lower or more uniform temperatures occur.

4.6 WINDS

Strong winds often associated with squalls and storms do occur, Dorrel (1947). The dry season winds tend to blow more directly from the North while the wet season wind is more Westerly and even North Westerly oriented and commences earlier with the Southerly movement of the Intertropical Convergence Zone (ITCZ) from October to April, during the period of October to February, the area is exposed to very dry winds blowing from the sahara, the harmattan, often

carrying a thick haze of windborn, conspicuously diametaceous dust (Digelstone, 1958; McKewn, 1958).

The rationale of the rainfall and temperature pattern falls from the above discussions on the origin, direction and variations in the local wind pattern. Local anticyclone effects modify the generalized South West direction of the wet season wind in the study area. The more Westerly or North Westerly winds creates a pattern in the wet season of up-wind exposure and down-wind shadow effects, fluctuating around the critical 900-1000 isohyets. The high temperatures in the area are probably due to low altitude and the effects of the land to the North blocking the Northerly orientated wind.

4.7 RADIATION

The hours of bright sunshine data available are for Maiduguri, Bauchi, Ibi, Yola and Potiskum, according to Cocheme' and Franguin (1967), the data shows that, radiation rises to a maximum in the months of November to February and falls to a minimum in August, when the cloud cover is highest. Between February and May the radiation amount falls and recovers considerably, this is probably due to increased rain cloud and storm conditions as the Intertropical Convergence Zone (ITCZ) retreats Southward.

4.8 HUMMIDITY

Accoring to Cocheme' and Franquin (1967), the humidity follows the change of the seasons, with the saturation deficit reaching a peak value in the pre-rains hot period in March and April, and falling to a minimum in September in the mid-rainy season. The humidity follows a simple inverse daily relationship with temperature reaching minimum values at night and maximum ones at the height of the day.

4.9 PHYSIOGRAPHY.

The study area is located in the Southern part of the Chad formation of quaternary rocks and are of the older origin, a boundary therefore exist between them. Apart from the basement plains which lie North of Mubi, the Mandara mountains on the extreme Eastern boarder with the Camerouns to the right of river Yedsaram and North of Mubi, this two fold geological division roughly coincide with the divide between the North-Easterly flowing rivers which drain to Lake Chad and the South, and South Westerly flowing tributaries of the Benue.

The South Eastern part of the study area consist almost entirely of flat and very gently undulating plains descending gradually from about 600m from the North West through the Ghaanda hills where it rose over 900m and down to about 240m in Lokko village on the shores of River Lokko. The Uba

plains to the North of the study area are drained by the Hawal and Yedsaram river system; And in the South some areas (including the study area) are drained by the Kilange river which flows through the Ghaanda hills into the Benue river.

The Ghaanda hills consists of a series of gently undulating partially dissected hills between 900 and 300m above sea level. Further South of the study area lies the Benue valley and consists of a series of gently undulating plains rising from between 90-155m along the Benue river itself to 150-240m adjacent to the Ghaanda hills.

4.10 GEOLOGY

There are four main rock types within the study area and its environs, Bawden (1972):

a) Metasediments, gneiss and migmatite;

b) Granite;

c) Basalt; and

d) Ancient Uuvium

a) Metasediments, gneiss and migmatite

The metasediments of the basement complex are a group of highly metamorphosed sedimentary rocks consisting mainly of gneisses and magmatites; Areas of schists and quartzite also occur and small occurences of amphibolite gabbro and

pegmatite are widely distributed.

b) Granite.

The oldest rocks of the basement complex have been subjected to at least two major orogenic cycles, Oyewoye (1964). Rock of the older granite suite were formed during the second cycle, which extended from late precambrian to lower palaezoic, Jacobson et al. (1963).

The older granite shows all stages of granitisation and magmatic activity. These have a considerable range in structure, texture and mineralogy, and their contact relationships with the metasediments type are very reliable, Carter et al. (1963) and Nigeria, Geological Survey (1964). The distribution of the older granites form the massive inselbergs in the Ghaanda hills.

c) Basalts

Volcanic rocks are widely distributed in the study area, and they occur as volcanic plugs and small complex lava flow of which the Song volcanic complex is the largest. The age of this volcanic epoch has not been precisely determined but it seems likely that it extended from the mid-Tertiary to recent, Carter et al. (1963). The volcanic rock type consist of olivine basalt flows, with small quantities of pyroclastic rocks, Nigeria, Geological Survey (1964).

d) Ancient alluvium

Geologically, some areas underlain by these ancient alluviaI deposits are subject to occasional flooding at the present and therefore also contain recent alluvium. The areas boardering the present flood plains of the Lokko river is of the older alluvial deposit. The recent alluvium occurs along most of the water courses and are of thin and discontinuous sand overlaying cretaceous sediments of the Bima sandstone.

4.11 GEOMORPHOLOGY

There is a close relationship between the land forms and the underlying rocks that are observable, and hence with the soils derived from these rocks.

a) Igneous and Metamorphic rocks, mainly of the basement complex:

The wide variety in the lithology of the basement rocks results in a considerable variation in their resistance to erosion, and there are many minor changes in landscape throughout the basement plains which can be directly related to this variation in lithology.

The migmatites and gneisses generally underlie more or less dissected undulating plains. The relief is fairly uniform with only occasional hills rising above the level of

the gently sloping interfluves. Upstanding areas which do occur in these plains are generally isolated rocky hills (inselbergs) or low elongated ranges often with a serrated outline. The orientation of these low ranges usually follows quartzite ridges, dykes or fault lines.

The older granite found in the South, and South West of the project area are commonly associated with hilly and highland areas. They form groups of iselbergs, massive steep sided hills and mountain ranges. The Ghaanda hills consist of groups of iselbergs between which are relics of more gentle terrian at the same altitude as parts of the Uba plains. The hills contain large areas of bare rock and shallow stony soils and in general the regolith is thin and discontinuous.

b) Basalt

Basalt plugs and cones occur, at Song, the lavas flow from steepsided cones with well preserved craters. They followed existing valleys which were partially infilled by basalt, but have been subsequently exhumed, so that new basalt flows border many of the stream lines and the underlying basement rocks form the interfluxes.

c) Recent alluvium.

Recent river alluvium occurs along the lower part of the Lokko River for a stretch of about 13km West of Lokko town. It varies in extent from thin discontinuous sands in the smallest streams to thick broad sands on the Lokko River itself. The sands form levels and point bars, levees predominate and the streams are generally fast flowing. Clays within this pattern are also associated with backswamps and abandoned meanders.

The Southern edge of the Ghaanda hills is marked by a sharp concave break of slope with the pediment which descends gradually to gently undulating plains. The break of slope coincides with the boundary between basement complex rocks and cretaceous rocks. Generally, pediments are formed over both the basement complex and cretaceous sandstones in the study area.

This pediment is primarily an area of sediment removal and is largely free from colluvial deposits. It is floored by a thin cover of Bima sandstone over the basement, and as erosion continues the original surface on which the sandstone was deposited is being exhumed. Pediments are formed on basement complex rocks around most of the inselbergs, where the inselbergs are scattered the pediment forms only a small peripheral apron with only locally developed fans.

4.12 LAND SYSTEM.

The land system is defined as "an area or group of areas throughout which can be recognised a recurring pattern of topography, soils and vegetation". The components of this pattern are land facets and the system is defined on its constituent facets and their inter-relationships.

The area around the Ghaanda hills have fairly well defined complex land system. In the Lokko foot slopes the land system is well defined, however, in the Song volcanic complex are found a complex and compound land system. It is essentially a mapping unit to include local lava flows and the areas of adjacent basement rocks influenced by the basalt lavas.

The area contains areas of good agricultural soils, Bawden et al. (1968), and are also suitable for grazing on eradication of tsetse fly. Mixed farming could be encouraged and dry season grazing could take place. The flat areas have potentials for cotton and other heavy soil crops but care must be taken in managing the Halomorphic soils.

4.13 LAND USE

There are settled farmlands at Lokko and in all other settlements along the flood plains in the project area, this however does not include areas that are excessively drained

or climatically dry such that water is limiting to crop growth and/or human habitation. Traditionally the soils are worked by the hoe cultivator and rainfall should be of sufficient intensity and duration to permit the regularly successful growth of the typical rainfall crops of the area.

The shifting cultivation are confined to the more stony basement plains, the fact that the soils are harder to work, make the plains less attractive to the cultivator.

Grazing are done both during the dry season and the wet season. In the wet season grazings are normally done in dunefields and sand plains North of the 635mm isohyet where rainfall is a limiting factor. While in the dry season grazings are done in the clay plains and the alluvial complex, the available area here is largely reduced by cultivators. However, the Halomorphic areas are largely free of cultivation.

CHAPTER 5: TIME SERIES METHOD FOR THE ESTIMATION OF WATER LEVELS IN RIVER LOKKO

In consideration of the estimation model which would best describe the data, the auto correlation at zero lag time was computed and found not to be equal to zero, that is,

$$_{o} = r_{o} \neq 0$$

Where,

$_{o}$ = auto correlaation function at zero lag

r_{o} = auto correlation estimate at zero lag

Therefore, it was concluded that a time series model would be most suitable, since if $r_{o} = 0$, a stochastic model would have describe the data best.

In developing the estimation model the following procedure was adopted:

5.1 AUTO CORRELATION ANALYSIS OF THE HISTORICAL DATA

The historical data used in the analysis were the mean monthly data for the period January, 1984 to December, 1987 and are tabulated in column 7 of table 2; column 1 of the same table refer to period. Period 1 refers to the month of January, 1984; similarly, period 12 refers to December 1984, period 24 refers to December, 1985, period 36 refers to December, 1986, and so on.

TABLE 2: RESULTS USING HISTORICAL DATA

Period t	Month	ϵ_t=K E_t	β^1_t	β_t	Y_t	Actual	Forecast, F^Y_t			Error	%
							Lower Bound	Mean	Upper Bound	e_t	Error
1	Jan. '84	.45463	.00000	-6.64766	-6.64766	31.9	30.7	30.8	30.9	1.1	3.4
2	Feb.	-.07349	-.13187	-6.70105	-6.83292	32.8	33.8	33.9	34.0	-1.1	3.4
3	Mar.	-.02096	-.60159	-6.74880	-7.35039	46.0	44.1	44.3	44.5	1.7	3.7
4	Apr.	-.04505	-.88220	-6.77161	-7.65381	50.4	51.6	51.8	52.0	-1.4	2.8
5	May.	-.36742	-1.30429	-6.83009	-8.13438	69.2	66.1	66.4	66.7	2.8	4.0
6	Jun.	-.32292	-1.60647	-6.80475	-8.47122	58.9	78.6	79.0	79.3	-20.1	34.1
7	Jul.	-.32837	-1.85390	-6.89602	-8.74992	75.2	90.8	91.3	91.7	-16.1	21.4
8	Aug.	-.17273	-1.87860	-6.92430	-8.80290	58.6	93.3	93.8	94.2	-35.2	60.1
9	Sep.	.16722	-1.48558	-6.94991	-8.43549	129.8	77.3	77.7	78.0	52.1	40.1
10	Oct.	.64139	-.51668	-6.97313	-7.48981	64.4	47.3	47.5	47.7	16.9	26.2
11	Nov.	.16037	.06317	-6.99421	-6.93104	41.1	35.7	35.7	35.8	5.4	13.1
12	Dec.	-.10832	.23425	-7.01338	-6.77913	38.9	33.0	33.0	33.1	5.9	15.2
13	Jan. '85	.45463	.76685	-7.03082	-6.27135	31.1	25.3	25.4	25.5	5.7	18.3
14	Feb.	-.07349	.76988	-7.04671	-6.27683	28.9	25.4	25.5	25.5	3.4	11.8
15	Mar.	-.02096	.70402	-7.06120	-6.35718	23.8	26.4	26.5	26.6	-2.7	11.3
16	Apr.	-.04505	.53601	-7.07442	-6.53841	36.8	29.0	29.1	29.2	7.7	20.9
17	May.	-.36742	-.02082	-7.08649	-7.10731	76.1	38.9	39.1	39.2	37.0	48.6
18	Jun.	-.32292	-.53376	-7.09752	-7.63128	93.3	50.9	51.1	51.3	42.2	45.2
19	Jul.	-.32837	-1.01107	-7.10760	-8.11867	116.1	65.4	65.7	66.0	50.4	43.4
20	Aug.	-.17273	-1.24661	-7.11682	-8.36343	113.7	74.2	74.5	74.8	39.2	34.5

TABLE 2 (Contd)

21	Sep.	.16722	-1.03131	-7.12527	-8.15658	75.1	66.7	67.0	67.3	8.1	10.8
22	Oct.	.64139	-.20300	-7.13300	-7.33600	47.4	43.8	44.0	44.1	3.4	7.2
23	Nov.	.16037	.27086	-7.14008	-6.86922	41.7	34.4	34.6	34.7	7.1	17.0
24	Dec.	-.10832	.36537	-7.14657	-6.78120	34.5	32.9	33.0	33.1	1.5	4.3
25	Jan. '86	.45463	.84483	-7.15253	-6.30770	31.4	25.8	25.9	26.0	5.5	17.5
26	Feb.	-.07349	.81248	-7.15799	-6.34551	31.6	26.3	26.4	26.5	5.2	16.5
27	Mar.	-.02096	.72414	-7.16300	-6.43886	32.6	27.6	27.7	27.8	4.9	15.0
28	Apr.	-.04505	.54265	-7.16760	-6.62495	32.7	30.4	30.5	30.6	2.2	6.7
29	May.	-.36742	.03485	-7.17182	-7.13697	60.4	39.5	39.7	39.8	20.7	34.3
30	Jun.	-.32292	-.46756	-7.17569	-7.64325	67.8	51.3	51.5	51.7	16.3	24.0
31	Jul.	-.32837	-.94502	-7.17925	-8.12427	100.7	65.7	66.0	66.3	34.7	34.5
32	Aug.	-.17273	-1.17671	-7.18251	-8.35922	152.7	74.1	74.0	74.7	78.3	51.3
33	Sep.	.16722	-.96986	-7.18552	-8.15538	90.9	66.6	66.8	67.1	14.1	17.4
34	Oct.	.64139	-.15222	-7.18827	-7.34049	40.8	43.9	44.0	44.2	-3.2	7.8
35	Nov.	.16037	.31027	-7.19080	-6.88053	28.4	34.6	34.8	34.9	-6.4	22.5
36	Dec.	-.10832	.39465	-7.19313	-6.79848	25.4	33.2	33.3	33.5	-7.9	31.1
37	Jan.'87	.45463	.86568	-7.19527	-6.32959	19.4	26.1	26.2	26.3	-6.8	35.1
38	Feb.	-.07349	.82674	-7.19724	-6.37050	18.7	26.6	26.7	26.8	-8.0	42.8
39	Mar.	-.02096	.73348	-7.19905	-6.46557	25.4	28.0	28.1	28.2	-2.7	10.6
40	Apr.	-.04505	.54847	-7.20071	-6.65224	17.4	30.8	30.9	31.0	-13.5	77.6
41	May.	-.36742	.03825	-7.20225	-7.16400	30.2	40.1	40.3	40.4	-10.1	33.4
42	Jun.	-.32292	-.47001	-7.20365	-7.67366	66.9	52.1	52.4	52.6	14.5	21.7
43	Jul.	-.32837	-.94505	-7.20495	-8.15000	78.4	66.6	66.9	67.2	11.5	14.7
44	Aug.	-.17273	-1.18855	-7.20614	-8.39469	72.7	75.5	75.9	76.2	-3.2	4.4
45	Sep.	.16722	-.98344	-7.20723	-8.19067	64.8	68.0	68.3	68.6	-3.5	5.4
46	Oct.	.64139	-.16586	-7.20824	-7.37410	55.7	44.7	44.9	45.0	10.8	19.4
47	Nov.	.16037	.29843	-7.20917	-6.91074	48.7	35.2	35.3	35.4	13.4	27.5
48	Dec.	-.10832	.38499	-7.21002	-6.82503	46.6	33.6	33.8	33.9	12.8	27.5

The auto correlation coefficient, $r_k 1$ at different periods were computed using the general equation given as

$$r_k = \frac{\sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y}_t)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

where,

K = period
Y = Mean of original observation = 55.125cm
n = Number of data in series

Y_t = Original observation at time t

Y_{t+k} = Original observation at time t plus k period.

r_k = Auto correlation coefficient at K period.

The auto correlation coefficients obtained for the various lag times were,

$r_o = + 1.000$	$r_1 = + 0.654$	$r_2 = + 0.239$
$r_3 = - 0.064$	$r_4 = - 0.396$	$r_5 = - 0.427$
$r_6 = - 0.568$	$r_7 = - 0.501$	$r_8 = - 0.353$
$r_9 = - 0.080$	$r_{10} = + 0.248$	$r_{11} = + 0.473$
$r_{12} = + 0.572$	$r_{13} = + 0.473$	$r_{14} = + 0.263$
$r_{15} = - 0.009$	$r_{16} = - 0.264$	$r_{17} = - 0.383$
$r_{18} = - 0.403$	$r_{19} = - 0.342$	$r_{20} = - 0.204$
$r_{21} = - 0.082$	$r_{22} = + 0.164$	$r_{23} = + 0.318$
$r_{24} = + 0.267$	$r_{25} = + 0.219$	$r_{26} = + 0.129$
$r_{27} = + 0.014$	$r_{28} = - 0.097$	$r_{29} = - 0.178$
$r_{30} = + 0.205$	$r_{31} = - 0.190$	$r_{32} = - 0.081$

r_{33} = + 0.023 r_{34} = + 0.066 r_{35} = + 0.077

r_{36} = + 0.075 r_{37} = + 0.062 r_{38} = + 0.040

r_{39} = + 0.020 r_{40} = 0.000 r_{41} = - 0.027

r_{42} = - 0.026 r_{43} = - 0.016 r_{44} = - 0.003

r_{45} = + 0.005 r_{46} = + 0.008 r_{47} = + 0.004

These auto correlation values are plotted in figure 6, against the corresponding periods (months). As shown by Anderson (1942), Bartlett (1946), Quenoville (1949), and others, the auto correlation coefficients of a random series have a sampling distribution that can be approximated by a normal curve with zero mean and standard error $1/\sqrt{n}$. This information was used to determined whether some auto correlation coefficients comes from a population whose value is zero at K period. Since n is 48, the standard error is $1/\sqrt{48} = 0.144$. This means that 95 percent of all sample based on the historical data auto correlation coefficients must lie within a range specified by the mean plus or minus 2.016 standard errors (fiducial or confidence limits) for the data to be random.

The value of 2.016 was obtained from the student t - distribution table 3. Thereffore, the data can be concluded to be random if the calculated auto correlation coefficients are within limits.

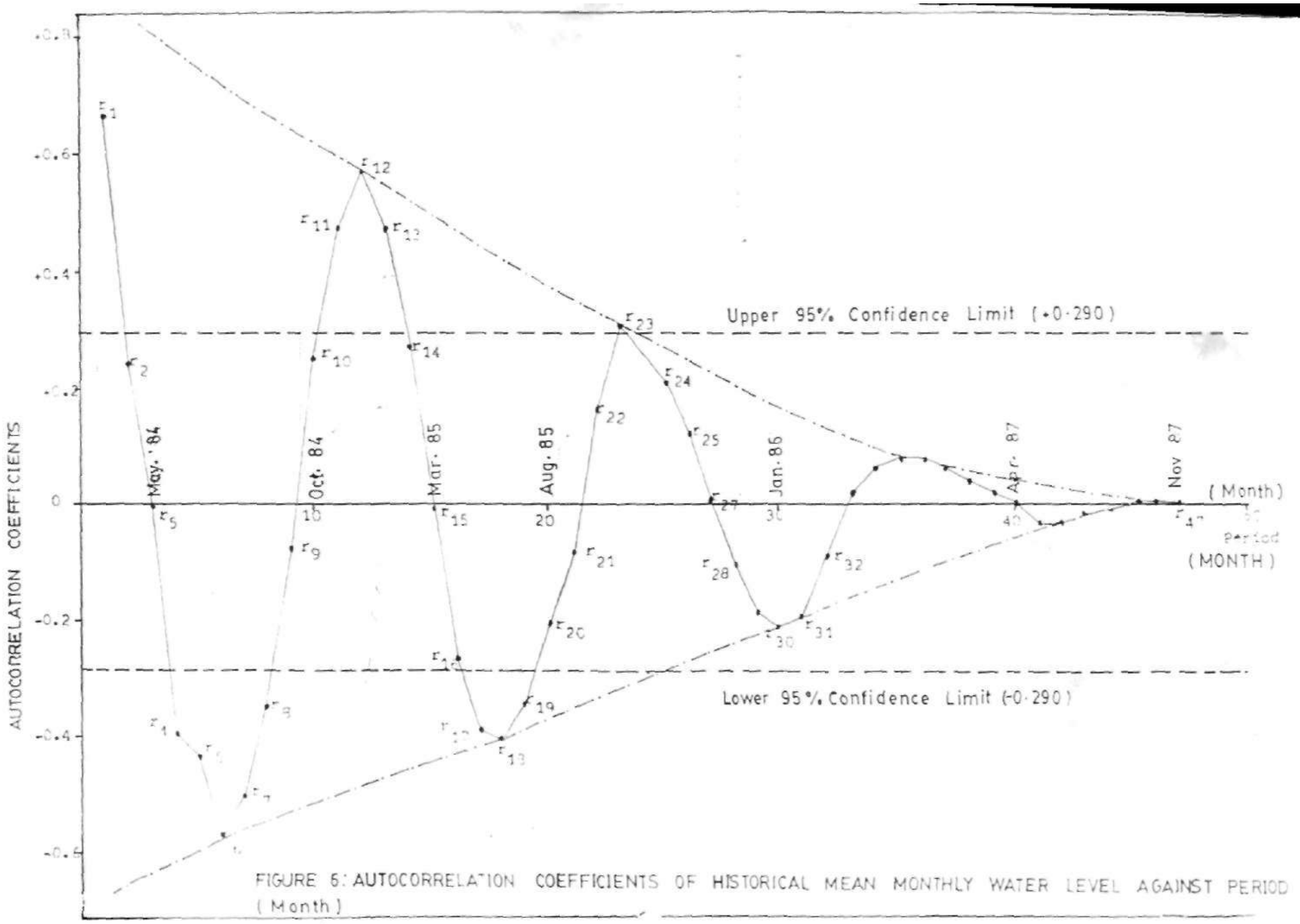


FIGURE 6: AUTOCORRELATION COEFFICIENTS OF HISTORICAL MEAN MONTHLY WATER LEVEL AGAINST PERIOD (Month)

TABLE 3: Student t Distribution.

df \ p	0.50	0.10	0.05	0.02	0.01
1	1.000	6.34	12.71	31.82	63.66
2	0.816	2.92	4.30	6.96	9.92
3	0.765	2.35	3.18	4.54	5.84
4	0.741	2.13	2.78	3.75	4.60
5	0.727	2.02	2.57	3.36	4.03
6	0.718	1.94	2.45	3.14	3.71
7	0.711	1.90	2.36	3.00	3.50
8	0.706	1.86	2.31	2.90	3.36
9	0.703	1.83	2.26	2.82	3.25
10	0.700	1.81	2.23	2.76	3.17
11	0.697	1.80	2.20	2.72	3.11
12	0.695	1.78	2.18	2.68	3.06
13	0.694	1.77	2.16	2.65	3.01
14	0.692	1.76	2.14	2.62	2.98
15	0.691	1.75	2.13	2.60	2.95
16	0.690	1.75	2.12	2.58	2.92
17	0.689	1.74	2.11	2.57	2.90
18	0.688	1.73	2.10	2.55	2.88
19	0.688	1.73	2.09	2.54	2.86
20	0.687	1.72	2.09	2.53	2.84
21	0.686	1.72	2.08	2.52	2.83
22	0.686	1.72	2.07	2.51	2.82
23	0.685	1.71	2.07	2.50	2.81
24	0.685	1.71	2.06	2.49	2.80
25	0.684	1.71	2.06	2.48	2.79
26	0.684	1.71	2.06	2.48	2.78
27	0.684	1.70	2.05	2.47	2.77
28	0.683	1.70	2.05	2.47	2.76
29	0.683	1.70	2.04	2.46	2.76
30	0.683	1.70	2.04	2.46	2.75
35	0.682	1.69	2.03	2.44	2.72
40	0.681	1.68	2.02	2.42	2.71
45	0.680	1.68	2.02	2.41	2.69
50	0.679	1.68	2.01	2.40	2.68
60	0.678	1.67	2.00	2.39	2.66
oC	0.674	1.64	1.96	2.33	2.58

(Taken from forecasting: Methods and Applications by Makridakis,S. and Wheelwrights,S.C. (1978))

$$-2.016(0.144) \leq r_k \leq +2.016\ (0.144)$$

$$-0.290 \leq r_k \leq +0.290$$

Figure 6, shows that the auto correlation coefficients of the historical data at the various periods. The two dotted lines perpendicular to the vertical axis are the upper and lower 95 percent confidence limits for a random series of the historical data and this correspond to - 0.290 to + 0.290. Note that, not all the auto correlation coeficients lie within these limits, which means that the data are non-random. This was confirmed with the x^2 - Test developed by Box - Pierce (1970), known as the Box - Pierce Q - statistics. The Q - statistics is computed as

$$x^2 - \text{test} = n\sum_k r_k^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (13)$$

where, m is the largest period included.

According to the test, if the computed value of the x^2 -test is less than that from the table of the value of x^2 - Statistics, table 4, that is x^2 - critical, it follows that the data generating the auto correlations are random; otherwise, the data are non-random. The computed x^2 - Test was 147.570 and was found to be greater than the x^2 - critical of 63.981. It was therefore concluded that the data are non-random.

TABLE 4: Critical points of the Chi-squared (x^2) statistics

df \ P	0.050	0.025	0.010	0.005
1	3.84146	5.02389	6.63490	7.87944
2	5.99147	7.37776	9.21034	10.5966
3	7.81473	9.34840	11.3449	12.8381
4	9.48773	11.1433	13.2767	14.8602
5	11.0705	12.8325	15.0863	16.7496
6	12.5916	14.4494	16.8119	18.5476
7	14.0671	16.0128	18.4753	20.2777
8	15.5073	17.5346	20.0902	21.9550
9	16.9190	19.0228	21.6660	23.5893
10	18.3070	20.4831	23.2093	25.1882
11	19.6751	21.9200	24.7250	26.7569
12	21.0261	23.3367	26.2170	28.2995
13	22.3621	24.7356	27.6883	29.8194
14	23.6848	26.1190	29.1413	31.3193
15	24.9958	27.4884	30.5779	32.8013
16	26.2962	28.8454	31.9999	34.2672
17	27.5871	30.1910	33.4087	35.7185
18	28.8693	31.5264	34.8053	37.1564
19	30.1435	32.8523	36.1908	38.5822
20	31.4104	34.1696	37.5662	39.9968
21	32.6705	35.4789	38.9321	41.4010
22	33.9244	36.7807	40.2894	42.7956
23	35.1725	38.0757	41.6384	44.1813
24	36.4151	39.3641	42.9798	45.5585
25	37.6525	40.6465	44.3141	46.9278
26	38.8852	41.9232	45.6417	48.2899
27	40.1133	43.1944	46.9630	49.6449
28	41.3372	44.4607	48.2782	50.9933
29	42.5569	45.7222	49.5879	52.3356
30	43.7729	46.9792	50.8922	53.6720
40	55.7585	59.3417	63.6907	66.7659
50	67.5048	71.4202	76.1539	79.4900
60	79.0819	83.2976	88.3794	91.9517
70	90.5312	95.0231	100.425	104.215
80	101.879	106.629	112.329	116.321
90	113.145	118.136	124.116	128.299
100	124.342	129.561	135.807	140.169

(Taken from forecasting: Methods and Applications by Makri-dakis,S. and Wheelwrights,S.C. (1978))

5.2 IDENTIFICATION OF PATTERN IN THE HISTORICAL DATA

Owning to chance, the auto correlation values will be slightly different from those theoretically expected, necessitating the use of confidence limits as explained above. It can be seen from figure 6, that the auto correlation values r_1, r_4, r_5, r_6, r_7, r_8, r_{11}, r_{12}, r_{13}, r_{17}, r_{18}, r_{19} and r_{23}
are outside the random limits, indicating that there is some pattern in the historical data. Note also that this pattern repeats itself after about six (6) periods (months), with the positive - positive, or negative - negative, at 12 periods (months) intervals, although the auto correlation value corresponding to period 36 is within the random limits, it however, suggests a reoccurence of the yearly pattern, accounting for the non-randomnes of the data.

In order to eliminate the pattern in the series, the power transformation proposed by Box and Cox (1964) has been used because of its high reliability; Details are given in section 5.5 below.

5.3 IDENTIFICATION OF STATIONARITY IN THE HISTORICAL DATA.

Stationarity means that there is no growth or decline in the data, that is, the data when plotted must lie parallel to the time axis, the data may however fluctuate around a constant mean, independent of time. This characteristics

of stationarity in the data can be identified by examining the plot of the auto correlation coefficients, figure 6.

Theoretically, the auto correlation coefficients of stationary data drops to zero after the second or third time lag, while for a non-stationary series they are significantly different from zero for several time periods and generally show a trend going diagonally from r_1 at an angle to the vertical across the auto correlation axis, as the period increases.

Figure 6, shows that the auto correlations somewhat fluctuate about the theoretical mean for a random auto correlated series of zero mean, it drops to zero after the second value, and it does not have a noticeable trend, therefore, it was concluded that the historical data are stationary but non-random, hence, there is no need to perform differencing in order to remove non - stationarity. The significant auto correlations r_1, r_4, r_5, r_6, r_7, r_8, r_{11}, r_{12}, r_{13}, r_{17}, r_{18}, r_{19} and r_{23} imply the existence of some pattern other than trend, this pattern would therefore be sought for from seasonality.

5.4 IDENTIFICATION OF SEASONALITY IN THE HISTORICAL DATA.

Seasonality can be defined as a pattern that repeats itself over fixed intervals of time, and this can easily be

identified by examining the auto correlations that are significantly different from zero, however, the data must be stationary for such indentification to be possible as non-stationarity introduce spurious auto correlations. Since the historical data are stationary it is easy to identify from figure 6, a clear seasonal pattern of about twelve periods (months) duration as r_{12} and r_{23} have high positive values, indicating dry season, and r_6 and r_{18} have high negative values, indicating wet seasons.

The periodic character of seasonality can be seen by the fact the $r_1 > r_6 > r_{12} > r_{18}$ and all four are significantly different from zero, actually the auto correlation $r_{12} = +0.57$ suffices to indicate seasonality with a length of twelve periods (months). The additional fact that r_{24} is large confirms it. It was noted however, that there are seasonal effects on the peaks, it smoothens them out exponentially and there tend to be some contraction to the left.

Before further analysis of the data could be done it was necessary to transform them in order to eliminate non-randomness.

5.5 TRANSFORMATION OF THE HISTORICAL DATA.

The power transformation proposed by Box and Cox (1964) was adopted because of its reliability as mentioned earlier and because of its better transformation than other

techniques. This is because this transformation induces simultaneously, normality, constancy of variance and linearity.

In performing this transformation a family of transformation from Y_t to $Y_{\beta t}$ was sought for, where

$Y_t \equiv$ historical data

$Y_{\beta t} \equiv$ transformed data such that,

$$Y_\beta = \frac{(x^\beta - 1)}{\beta} \text{ for } \beta \neq 0 \quad \text{.....(14a)}$$

or $Y = \ln x$ for $= 0$

where,

β = power tranformation (or index)

Equivalently,

$$Y = \begin{cases} x^\beta & \text{for } \beta \neq 0 \\ \ln X & \text{for } \beta \neq 0 \end{cases} \quad \text{.....(14b)}$$

and looked for a value of β that induces simultaneously normality, constancy of variance and linearity.

Theoretically, the optimal value of opt = β should have a coefficient of skewness, G = 0, such that the 95 percent confidence interval on the skewness of the transformed data.

$$G_\beta = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_{\beta i} - Y_n)^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(Y_{\beta i} - Y_\beta)^2\right]^{3/2}} \quad \text{.....(15)}$$

lies in the interval

$$G - 2f\sigma/n \leq G \leq G + 2f\sigma/n \quad \ldots\ldots\ldots\ldots\ldots\ldots (16)$$

where σ^2 = variance of the transformed data

for β = 0.05 G = 0.192

β = 0.15 G = 0.299

β = 0.25 G = 0.391

β = 0.35 G = 0.513

β = 0.50 G = 0.659

In order to find βopt (the optimal value of β), which correspond to G = 0, we write a relationship for β and G. Since β and G are generally linearly related, we have

$$\beta = a + b\,G$$

where,

a = Intercept on the β - axis

(corresponding to G = 0)

b = slope of the line of best fit

The line could have been drawn using the data above, but a better approximation was obtained using the simple linear regression (method of least squares), such that

$$b = \frac{n\Sigma \beta G - \Sigma\beta\, \Sigma G}{n\Sigma G^2 - (\Sigma G)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (18)$$

and

$$a = \frac{\Sigma\beta}{} - b\frac{\Sigma G}{}$$

$$\frac{n \quad n}{\text{.....}} \quad \text{(19)}$$

where, n = 5

	β	G	βG	G^2
	0.05	0.192	0.00960	0.036864
	0.15	0.299	0.04485	0.089401
	0.25	0.391	0.09775	0.152881
	0.35	0.513	0.17955	0.263169
	0.50	0.659	0.32950	0.434281
Σ	1.30	2.054	0.66125	0.976596

Therefore,

$$b = \frac{5 \times 0.66125 - 1.30 \times 2.054}{5 \times 0.976596 - (2.054)^2}$$

$$= \frac{0.63605}{0.664064}$$

$$b = 0.9578$$

hence, the slope of the line of best fit, b = 0.9578, that is, it is almost at an angle of 45 degrees to the horizontal. Similarly,

$$a = \frac{1.30}{5} - 0.9578 \frac{2.054}{5}$$

$$a = -0.13347$$

hence, the intercept of the line of best fit, with the β-axis is, a = - 0.13347, and this is the value of β at

which the skewness coefficient, G, is approximately equal to zero. Figure 7, shows the plot of β against G.

The computed G, corresponding to $\beta = -0.13347$ is

$G_\beta = -0.019$, with a confidence interval of

$$-0.307 \leq G_\beta \leq +0.269$$

Hence, the original data can be represented by the expression

$$X_t = Y_t^{-0.13347} \quad \text{..... (20)}$$

It was therefore required to seek for the estimation model which would fit the transformed data, and later perform the necessary re-transformation.

5.6 AUTO CORRELATION ANALYSIS OF THE TRANSFORMED DATA

The auto correlation coefficients, r^1_k, at different periods were computed using the general equation, but by replacing Y_t by X_t, such that

$$r^1_k = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^{n} (X_t - \bar{X}_t)^2} \quad \text{..... (21)}$$

where,

K = Period

$\bar{X}$ = Mean of transformed data

= 0.598

n = Number of data in series

X_t = Transformed data at time t

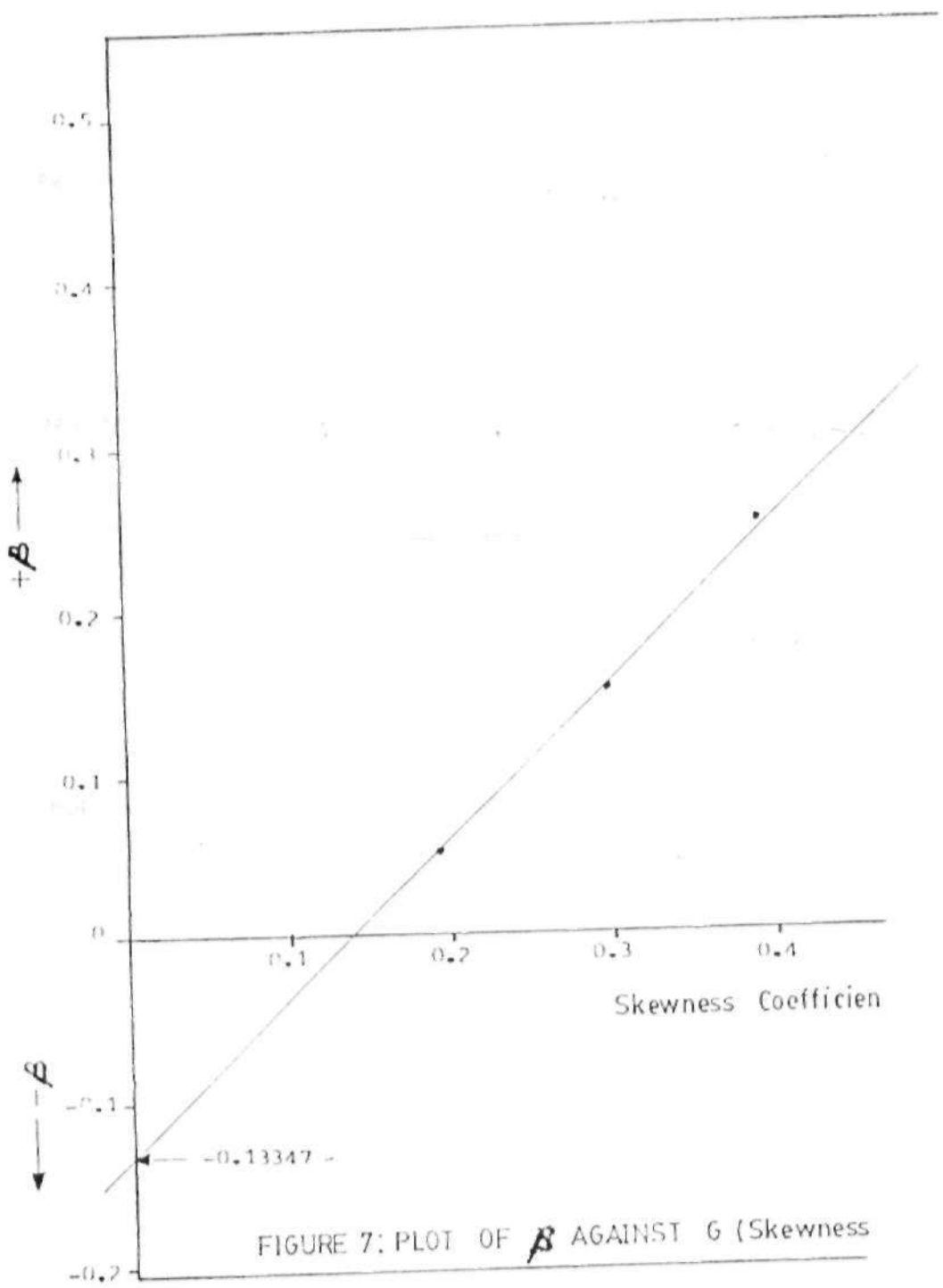


FIGURE 7: PLOT OF β AGAINST G (Skewness

X_{t+k} = Transformed data at time t plus k periods and the results are as listed below.

$r_1 = + 0.764$ $r_2 = + 0.461$ $r_3 = - 0.035$

$r_4 = - 0.299$ $r_5 = - 0.520$ $r_6 = - 0.658$

$r_7 = - 0.443$ $r_8 = - 0.418$ $r_9 = - 0.208$

$r_{10} = + 0.212$ $r_{11} = + 0.435$ $r_{12} = + 0.560$

$r_{13} = + 0.511$ $r_{14} = + 0.338$ $r_{15} = + 0.051$

$r_{16} = - 0.136$ $r_{17} = - 0.404$ $r_{18} = - 0.431$

$r_{19} = - 0.379$ $r_{20} = - 0.261$ $r_{21} = - 0.068$

$r_{22} = + 0.121$ $r_{23} = + 0.270$ $r_{24} = + 0.326$

$r_{25} = + 0.308$ $r_{26} = + 0.203$ $r_{27} = + 0.054$

$r_{28} = - 0.112$ $r_{29} = - 0.204$ $r_{30} = - 0.255$

$r_{31} = - 0.267$ $r_{32} = - 0.175$ $r_{33} = - 0.047$

$r_{34} = + 0.024$ $r_{35} = + 0.063$ $r_{36} = + 0.097$

$r_{37} = + 0.100$ $r_{38} = + 0.091$ $r_{39} = + 0.071$

$r_{40} = + 0.019$ $r_{41} = - 0.019$ $r_{42} = - 0.026$

$r_{43} = - 0.023$ $r_{44} = - 0.015$ $r_{45} = - 0.006$

$r_{46} = - 0.0008$ $r_{47} = - 0.0004$

It was observed that not all the auto correlation coefficients lie within the confidence limits which means that the data are still non-random after the power transformation, in fact the x^2 - test value of 197.224 was found to be greater than the x^2 - critical value of 63.981, which goes to confirm that the data are still non-random. The value of 197.224 for the transformed data is noted to be

higher than the value of 147.57 for the historical data, which means that the historical data was closer to randomness than the transformed data. However, in order to ensure minimum skewness in the distribution, the transformed data was used for all future analysis.

5.7 IDENTIFICATION OF PATTERN IN THE TRANSFORMED DATA.

The auto correlation plot of the transformed data was found to be very similar to the auto correlation plot of the historical data, with the auto correlation coefficients, r_1, r_2, r_4, r_5, r_6, r_7, r_8, r_{11}, r_{12}, r_{13}, r_{14}, r_{17}, r_{18}, r_{19} and r_{24} outside the random limits, indicating that there is some pattern in the transformed data. Note that the pattern also repeats itself after every six periods alternately, and at every twelve periods we have positive peaks or negative peaks, which suggests a seasonal pattern of twelve periods which has a strong influence on the randomness of the data.

5.8 IDENTIFICATION OF STATIONARITY IN THE TRANSFORMED DATA.

The auto correlation dropped to zero after the third period, with $r_3 = + 0.035$ and $r_4 = - 0.299$ which suggest stationarity in the data; however, the fluctuation of the auto correlation about the mean is not perfect, it is observed to decay exponentially, as with the case of the

historical data. This may have resulted in spurious auto correlations resulting in the non - randomness of the data. Therefore, it will be necessary to perform the first differencing (one short term) to remove period - to - period non-stationarity from the transformed data; which introduces non-randomness.

5.9 IDENTIFICATION OF SEASONALITY IN THE TRANSFORMED DATA.

Seasonal pattern of twelve months periods was observed which decays exponentially with increasing period and alternating in sign, from positive to negative, after six time periods, as was the case with the historical data. Therefore, it will be necessary to perform one long-term differencing in order to remove season - to - season non - stationarity from the transformed data, which introduces non-randomness.

5.10 REMOVING NON-STATIONARITY FROM THE TRANSFORMED DATA.

To remove period - to - period non - stationarity a new series was created that consisted of the differences between successive periods:

$$X_t^1 = X_{t+1} - X_t = X_t - X_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (22)$$

The new series X_t^1 have n - 1 data.

The long - term differencing of twelve periods (months) was also performed on the transformed data, such that

$$X_t^{11} = X_t - X_{t-12} \qquad \text{..... (23)}$$

The new series $X^{"}$ have n - 12 data.

Since performing the short and long - term differencing would remove both period - to - period trends as well as season - to - season trends, the simultaneous differencing was used, such that, the new series would be

$$X_t^{"'} = X_t \cdot X_t^{"} \qquad \text{..... (24)}$$

or

$$X_t^{"'} = (X_t - X_{t-1})(X_t - X_{t-12}) - \text{..... (25)}$$

having n - 12 data.

5.11 AUTO CORRELATION ANALYSIS OF THE DIFFERENCED - TRANSFORMED DATA.

The auto correlation coefficients, $r_k^{"}$, at different periods of the new series $X_t^{"'}$ mentioned above were computed and the results are as listed below and plotted in figure 8.

$r_1 = + 0.398$	$r_2 = + 0.201$	$r_3 = + 0.055$
$r_4 = - 0.149$	$r_5 = - 0.204$	$r_6 = + 0.024$
$r_7 = + 0.042$	$r_8 = + 0.102$	$r_9 = - 0.017$
$r_{10} = - 0.014$	$r_{11} = - 0.043$	$r_{12} = - 0.090$
$r_{13} = + 0.096$	$r_{14} = + 0.034$	$r_{15} = + 0.067$
$r_{16} = + 0.059$	$r_{17} = + 0.016$	$r_{18} = + 0.133$
$r_{19} = + 0.157$	$r_{20} = + 0.245$	$r_{21} = - 0.056$
$r_{22} = + 0.017$	$r_{23} = + 0.003$	$r_{24} = + 0.016$
$r_{25} = - 0.008$	$r_{26} = - 0.138$	$r_{27} = - 0.114$

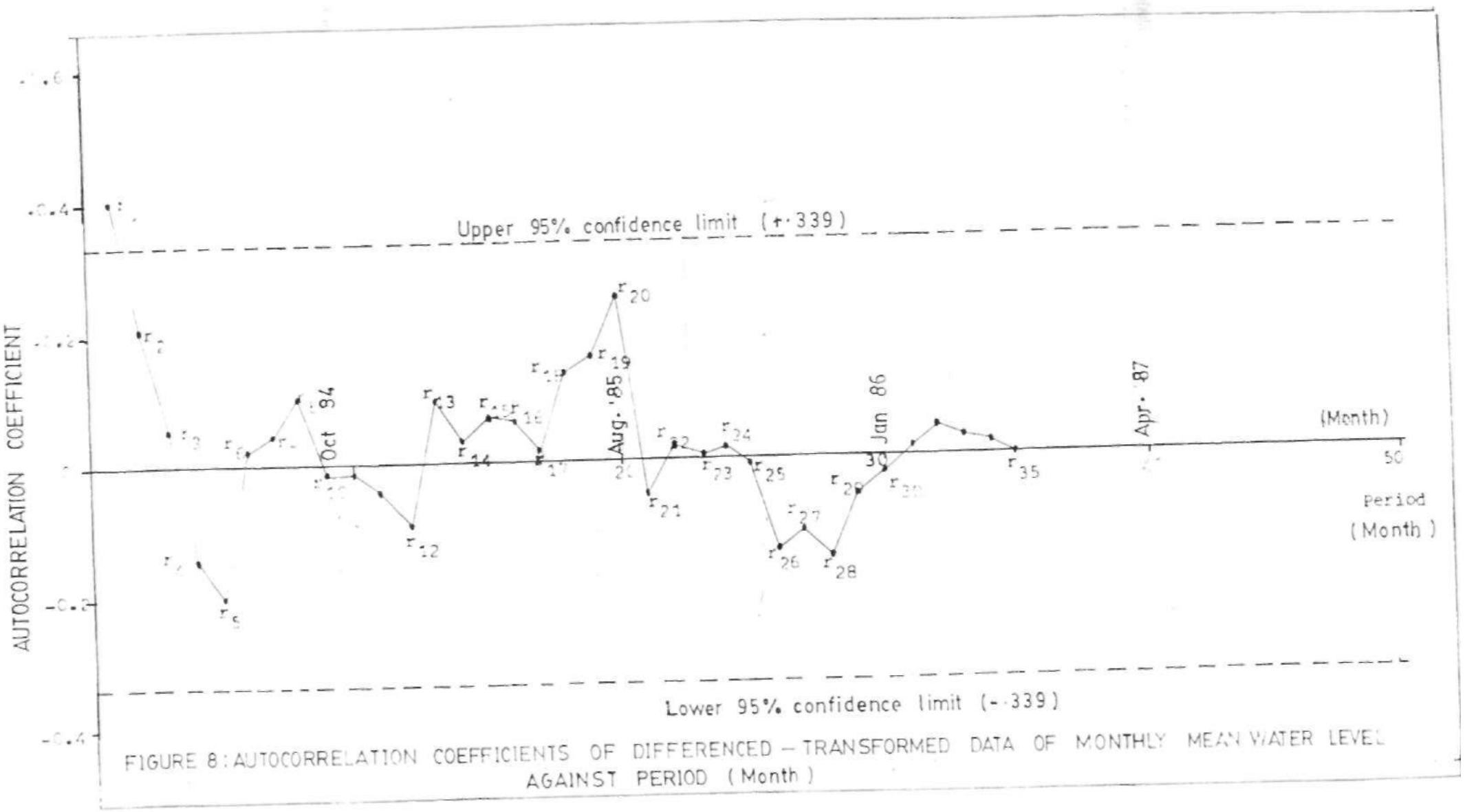


FIGURE 8 : AUTOCORRELATION COEFFICIENTS OF DIFFERENCED – TRANSFORMED DATA OF MONTHLY MEAN WATER LEVEL AGAINST PERIOD (Month)

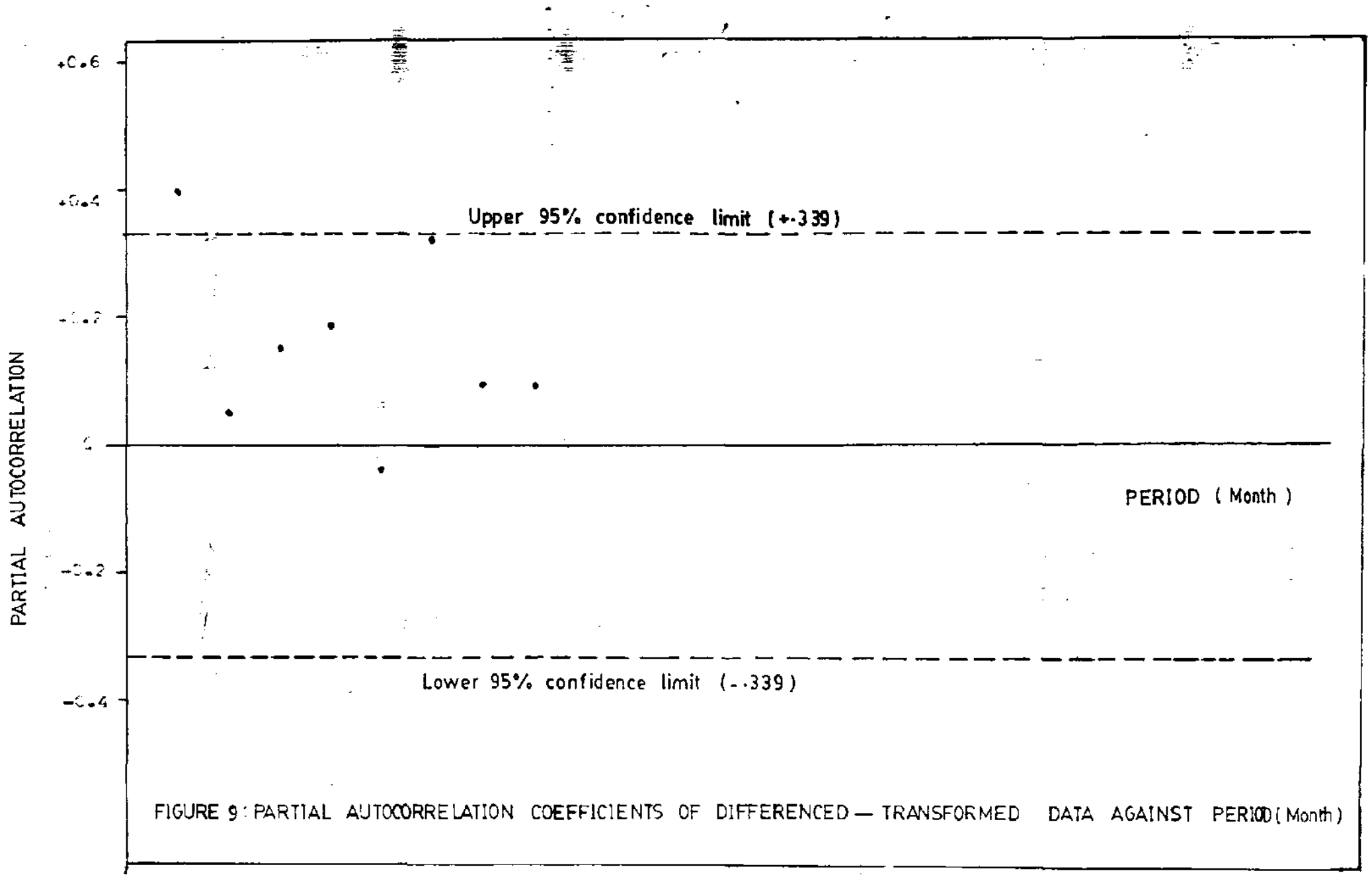


FIGURE 9 : PARTIAL AUTOCORRELATION COEFFICIENTS OF DIFFERENCED — TRANSFORMED DATA AGAINST PERIOD (Month)

r_{28} = - 0.151 r_{29} = - 0.058 r_{30} = - 0.028

r_{31} = + 0.009 r_{32} = + 0.042 r_{33} = + 0.024

r_{34} = + 0.014 r_{35} = + 0.002

It will be observed from the plot of the auto correlation coefficients of the differenced - transformed data, figure 9, that all lie within the random limits except, r_1, suggesting that the data are now random, the x^2 - test was performed and the value of 22.782 was obtained which is less than x^2 - critical of 43.781, confirming that the data are now random. The data also dropped to zero after r_3(+0.055), which means that they are also stationary.

5.12 IDENTIFICATION OF PATTERN IN THE DIFFERENCED TRANSFORMED DATA.

Figure 8, of the auto correlation plot of the differenced - transformed data shows that the sinusoidal and exponential decay pattern in the data has been eliminated, although the exponential decay can still be noted between r_1 r_{20} and r_{32}. The coefficients r_8, r_{20} and r_{32} are the auto correlations for the months of August and at peaks during the yearly circle corresponding to period of high water levels.

5.13 IDENTIFICATION OF STATIONARITY IN THE DIFFERENCED TRANSFORMED DATA.

As earlier observed the data are stationary since the

auto correlation coefficients dropped to zero after the third period, with $r_3 = + 0.055$ and $r_4 = - 0.149$. Therefore, the non - stationarity from period to period and from season to season have been removed, hence spurious auto correlation are not expected.

5.14 IDENTIFICATION OF SEASONALITY IN THE DIFFERENCED - TRANSFORMED DATA.

From the auto correlations plot, figure 8, the seasonal evidence are the peak values r_8, r_{20} and r_{32}, which occurs after every twelve period, suggesting a twelve months period, and since all the auto correlation coefficients except r_1 are within the random limits of ± 0.339, and the Q - statistics for randomness is satisfied, it follows that the difference - transformed data are random.

5.15 PARTIAL AUTO CORRELATION ANALYSIS OF THE DIFFERENCE - TRANSFORMED DATA.

The purpose of the partial auto correlation computation in time series analysis was to help identify an appropriate ARMA (Auto Regressive-Moving Average) model for forecasting. Therefore, it is only necessary to calculate the partial auto correlations until a value was obtained that is not

significantly different from zero, the number of partial auto correlations that are significantly different from zero gives the order of the AR (Auto Regressive) process, the first eight partial auto correlations,

$\phi_1 = + 0.398$ $\phi_2 = + 0.051$ $\phi_3 = + 0.152$

$\phi_4 = + 0.183$ $\phi_5 = - 0.040$ $\phi_6 = + 0.320$

$\phi_7 = + 0.093$ $\phi_8 = + 0.095$

were computed using the Yule-Walker equation given as

$$\tau_1 = \phi_1 + \phi_2\tau_1 + \phi_3\tau_2 + \ldots\ldots\ldots + \phi_p\tau_{p-1}$$

$$\tau_2 = \phi_1\tau_1 + \phi_2 + \phi_3\tau_1 + \ldots\ldots\ldots + \phi_p\tau_{p-2}$$

$$\tau_3 = \phi_1\tau_2 + \phi_2\tau_1 + \phi_3 + \ldots\ldots\ldots + \phi_p\tau_{p-3}$$

" " " " " " " " "

$$\tau_p = \phi_1\tau_{p-1} + \phi_2\tau_{p-2} + \phi_3\tau_{p-3} + \ldots + \phi_p \quad \ldots\ldots (26)$$

since the theoretical values of τ_k are not known they are replaced by their estimates r_k, where k = 1, 2, 3,....., p. It was observed that only ϕ_1 is outside the 95 percent confidence limits of ± 0.339, the cut off is 1, therefore the order of the estimation model was AR(1), this was confirmed from the fact that the auto correlation coefficients declined to zero exponentially, Figure 8.

Since for MA process, the values of the partial auto correlations does not have a cut off, rather it trails off exponentially to zero with time. Figure 9, shows that it is difficult to determine as the values of the partial auto

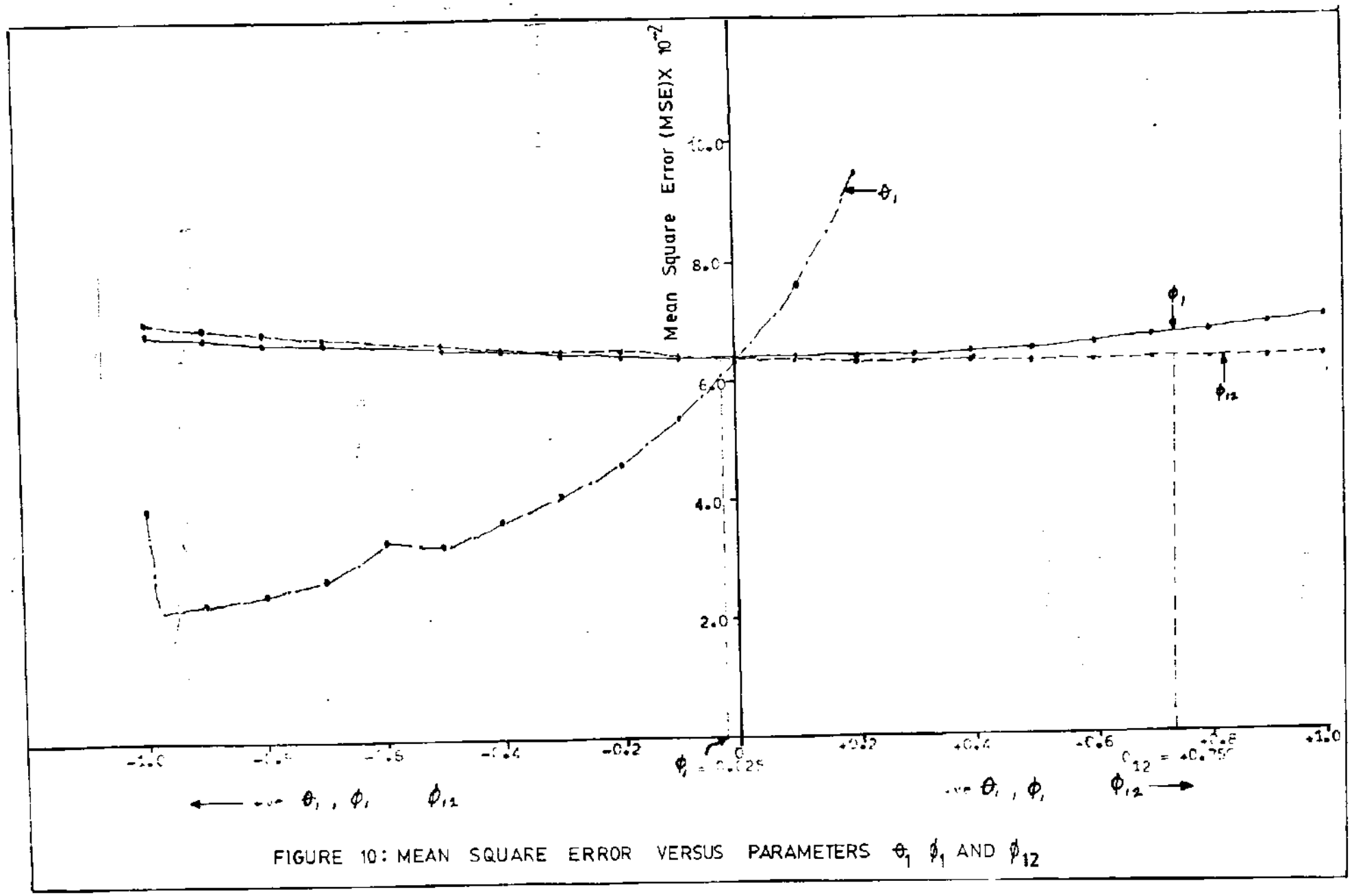


FIGURE 10: MEAN SQUARE ERROR VERSUS PARAMETERS θ_1 ϕ_1 AND ϕ_{12}

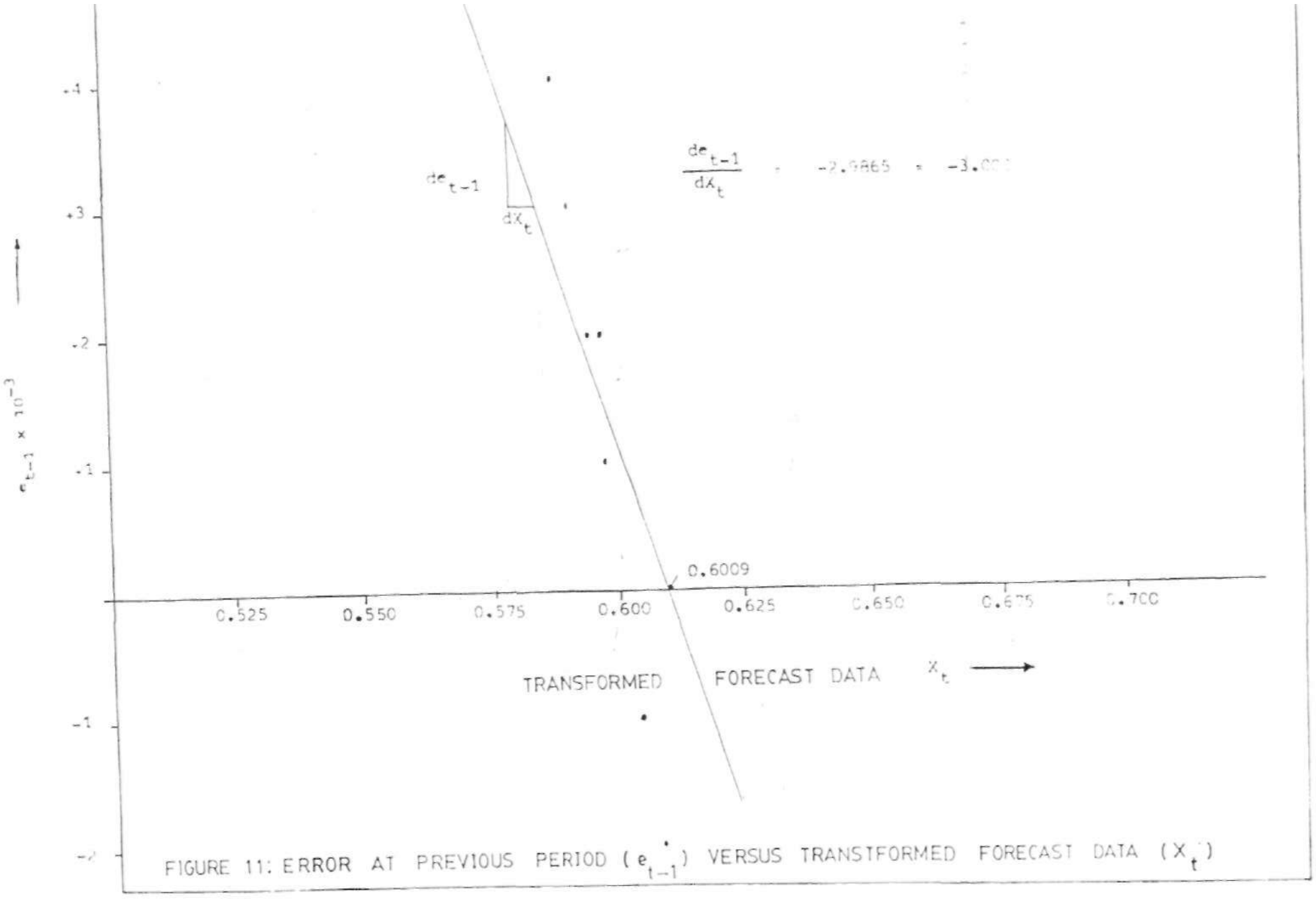


FIGURE 11: ERROR AT PREVIOUS PERIOD (e^{*}_{t-1}) VERSUS TRANSTFORMED FORECAST DATA ($X^{'}_{t}$)

correlations initially fell and rose again, however, a general exponential decay can be seen to pass through the points, therefore, it can be assumed that there is also an MA process, the order of which is 1, since only one auto correlation value is significantly different from zero. An MA(1) process was also identified.

5.16 IDENTIFICATION OF A TENTATIVE ARMA(p,q) MODEL.

In the proceeding section it was observed that present in the process are both AR (1) and MA (1), therefore according to Box and Jenkins (1970), an appropriate model to describe the process would be ARMA (1,1), where p=1 and q=1.

5.17 ESTIMATION OF THE PARAMETERS OF THE ARMA(1,1) MODEL

The general class of ARMA (p,q) process is represented by the equation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + e - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}$$

since an ARMA (1,1) has been identified as the model that would fit our data best, the model can be written as

$$X_t'' = \phi_1 X_{t-1}'' + e_t - \theta_1 e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots (27)$$

where,

$$X_{t-1}'' = (X_{t-1} - X_{t-2})(X_{t-1} - X_{t-13})$$

$$X_t'' = (X_t - X_{t-1})(X_t - X_{t-12})$$

ϕ_1 = AR parameter

θ_1 = MA parameter

e_t = error at t period

e_{t-1} = error at t-1 period

Applying the concept of backshift operator, B, then

$X_{t-1} = BX_t$

$X_{t-2} = B^2X_t$

$X_{t-13} = B^{13}X_t$

$X_{t-12} = B^{12}X_t$

$e_{t-1} = Be_t$

Therefore, eq. 27, can be rewritten as

$(X_t - BX_t)(X_t - B^{12}X_t) = \phi, (BX_t - B^2X_t)(BX_t - B^{13}X_t) + e_t - \theta_1 Be_t$

On simplification and introducing the seasonal parameters

$$X_t = X_t X_{t-1} + \phi_1 X_t X_{t-2} + \phi_{12} X_t X_{t-12} - \phi_{12} X_t X_{t-13} - \phi_1 X_t X_{t-14} - \phi_1 X_t X_{t-3} + \phi_1 X_t X_{t-15} + e_t - \theta_1 e_{t-1}$$

The above equation was solved for the parameters ϕ_1, ϕ_{12} and θ_1 which introduces the minimum mean square error (MSE), by the method of decomposition. Graphic plots of the results obtained are shown in figure 10, and the parameter values are as given below

$\phi_1 = - 0.025$

$\phi_{12} = + 0.750$

$\theta_1 = - 0.975$

obtained through further analysis (at shorter intervals of ϕ_1, ϕ_{12} and θ_1), in the region of minimum MSE.

Substituting these parameter values into the equation above,

$$X_t = X_t X_{t-1} - 0.025X_t X_{t-2} + 0.025X_t X_{t-3} + 0.750X_t X_{t-12} - 0.750X_t X_{t-13} + 0.025X_t X_{t-14} - 0.025X_t X_{t-15} + 0.975e_{t-1} + e_t \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (28)$$

This gives the time - series expression for estimation at t - period. However, it must be noted that the parameter values were obtained through the decomposition of the original expression, and are therefore very likely to be in error.

To obtain better parameter values, an approximate expression was derived to describe the system by substitution for X_t, X_{t-1}, X_{t-2}, X_{t-3}, X_{t-12}, X_{t-13}, X_{t-14}, and X_{t-15}, at the various time lags of 1 to 48. Initial value of zero was fixed for $e_{t-1} = e_o$ and all other values of e_t were computed. Results obtained shows that there is a general expression of the form

$$X_t = a_o + a_1{}^1 e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (29)$$

where, a_o and $a_1{}^1$ are parameters of the equation. Equation (29), is the equation for a straight line.

e_{t-1} are the computed errors at the various time lags allowing only one - hundredth of the errors computed in

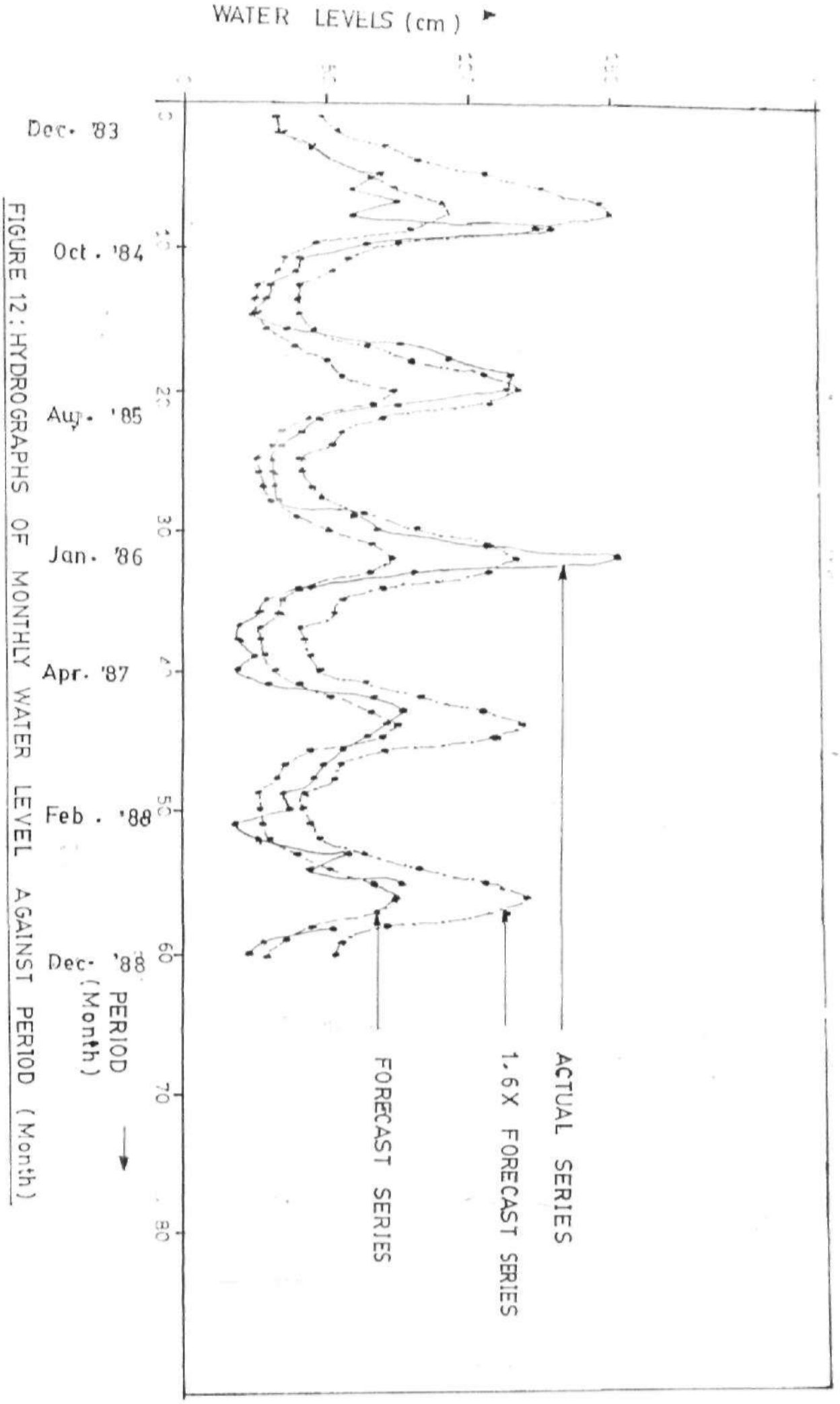


FIGURE 12: HYDROGRAPHS OF MONTHLY WATER LEVEL AGAINST PERIOD (Month)

the model above, we would have ammended Equation (29) to

$$X_t = a_o + a_1{}^1 \quad \frac{e^{t-1}}{100} \qquad \text{..... (30)}$$

or

$$X_t = a_o + a_1 e_{t-1} \qquad \text{..... (31)}$$

in which e_{t-1} is the independent variable and X_t, the dependent variable. The expression shows that the water levels in the river at any period is dependent only on the error in the immediate previous obtained estimated value. The parameter a_o and a_1 of this simple regression equation were obtained using the method of least squares, and the values are

$a_o = 0.6009$

$a_1 = -2.9865$

Therefore, the estimation model for the transformed data is given by

$$X_t = 0.6009 - 2.9865e_{t-1} \qquad - \qquad (32)$$

This was plotted for different values of X_t and e_{t-1}, and is shown in Fig. 11. The mean error in the forecast data was computed to be 0.00046.

The graph is slightly rotated to the left which shows some non randomness in the forecast data which must be removed by transformation.

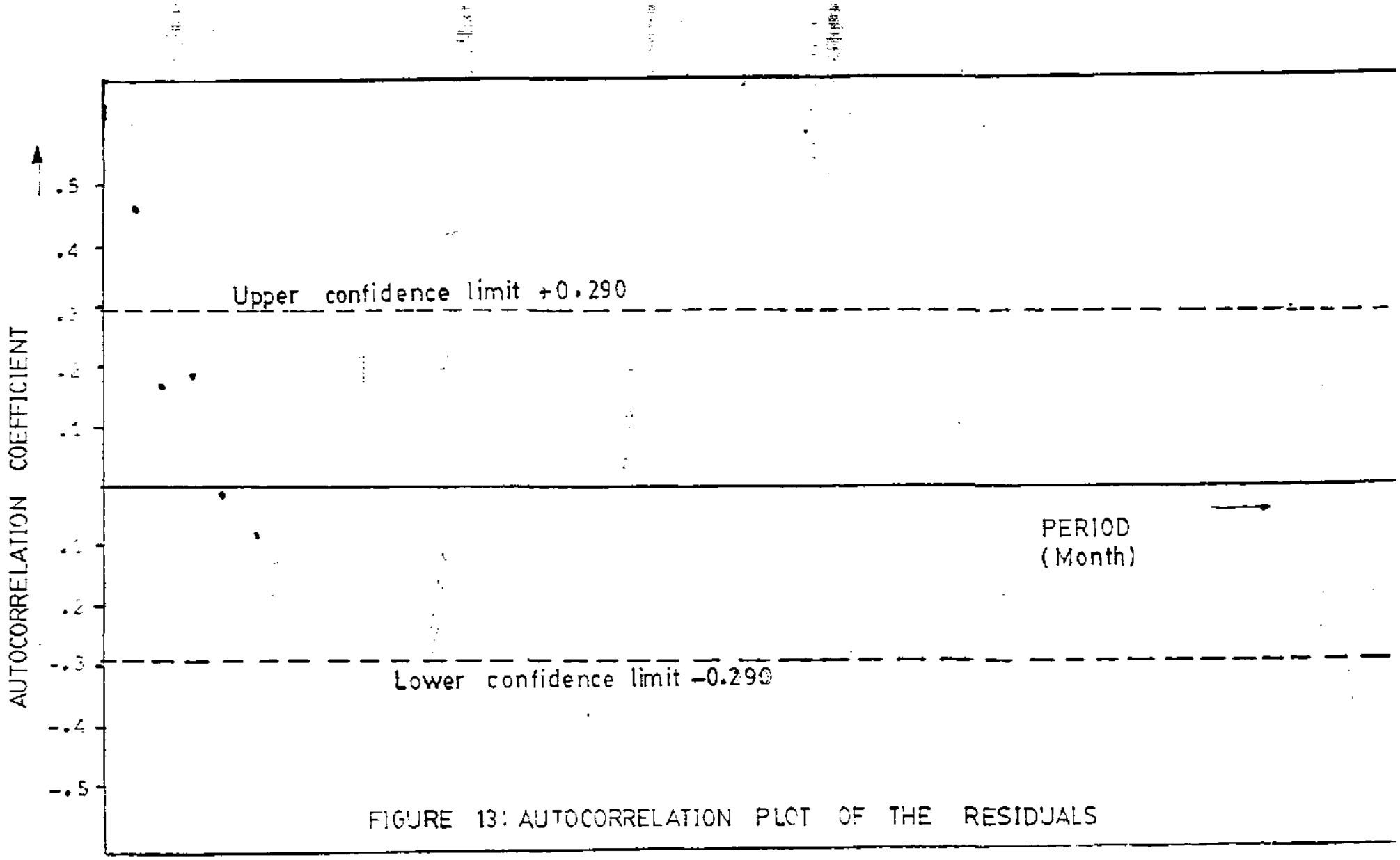


FIGURE 13: AUTOCORRELATION PLOT OF THE RESIDUALS

5.18 TRANSFORMATION OF THE ESTIMATED DATA.

The transformation was performed using the Box-Cox transformation, and the power transformation was obtained to be for

$$\beta = 0.8898$$

That is

$$X_t^T = X_t^F$$

where,

$$X_t^F = X_t = 0.6009 - 2.9865e_{t-1}$$

That was presented in figure 11. Therefore, the actual transformed estimation model for the system is

$$X_t = (0.6009 - 2.9865e_{t-1})^{0.8898}$$

Recall that

$$X_t = Y_t^{-0.13347}$$

where, Y_t is the observed data.

Hence,

$$Y^{-0.13347} = (0.6009 - 2.9865e_{t-1})^{0.8898}$$

Raising both sides of the equation to power $-\frac{1}{0.13347}$

$$Y_t = (0.6009 - 2.9865e_{t-1})^{-0.8898/0.13347}$$

or

$$Y_t = (0.6009 - 2.9865e_{t-1})^{-6.6667} \quad \ldots\ldots\ldots\ldots\ldots (34)$$

This expression means that to determine the water level in the river at time t, we must first of all compute the error,

between the measured and the estimated values of water level (that is, measured minus estimated) at time t-1, and substitute into the equation to obtain the water level at time t.

5.19 MODEL SENSITIVITY ANALYSIS

An attempt was made to estimate with the model

$$Y_t = (0.6009 - 2.9865e_{t-1})^{-6.6667}$$

the results obtained were found to be grossly not in agreement with the observed data, and in the situation where the expression in the parenthesis becomes negative the results were simply impossible. It was therefore thought that model modification would be necessary.

5.20 MODEL MODIFICATION

In order to modify the model the index - 6.6667 was replaced by β_t as it was noted to vary with time, such that

$$Y_t = (0.6009 - 2.7865e_{t-1})^{\beta_t} \quad \ldots\ldots\ldots\ldots\ldots\ldots (35)$$

In this equation Y_t is known for the first 48 time lag, e_{t-1} and β_t are not known. However, using the errors in the linear regression model at the different period (months), values were extrapolated for β_t for time lags 1 to 48. Table 2.

The problem now was to model into the future, and then use the rsults obtained to obtain actual model errors at the different periods. and then model the error terms. This

therefore means that at least three models would need to be combined to obtain the desired forecast model,

Viz i) model for β_t

ii) model for e_{t-1}

iii) model for Y_t

5.20.1 MODEL FOR β_t

In order to model β_t it was thought that it would be best to look at the disturbances in the system causing the observed instability of the model. From the concept of fluid mechanics, we assumed that there exist some external forces which is proportional to the displacement, in accordance with Newton's 2nd law of motion, such that we have,

$$\frac{d^2}{dt^2}\beta_t = a\,\beta_t \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (36)$$

where, a = constant

On simplification of equation (assuming a steady state) by transformation and modelling using the least square method, we have

$$\ln\ln|\beta_t|^{0.1} = C_p + b\ln|\beta_{t-1}|^{0.1} \qquad \ldots\ldots\ldots (37)$$

where,

b and C_p = constants

This over simplification was understood would result in errors, which is dealt in Appendix 1, page 147.

Fourty seven (47) data at various time step were generated using the earlier extrapolated values of β_t, to obtain values for

$$\ln\ln|\beta_t|^{0.1} \text{ and } \ln|\beta_{t-1}|^{0.1}$$

These values were used in the least square method to obtain values for the coefficients C_p and b, such that

$$b = 4.6569$$

and $C_p = -2.54168$

Therefore, the linear prediction is

$$\ln\ln|\beta_t|^{0.1} = -2.54168 + 4.6569 \ln|\beta_{t-1}|^{0.1}$$

Model error was estimated to be zero.

Rearranging the equation above, we have

$$\ln\ln|\beta_t|^{0.1} = -2.54168 + \ln|\beta_{t-1}|^{0.1} \times 4.6569$$

or

$$\ln\ln|\beta_t|^{0.1} = \ln 0.078734 + \ln|\beta_{t-1}|^{0.46569}$$

or

$$\ln\ln|\beta_t|^{0.1} = \ln(0.078734|\beta_{t-1}|^{0.46569}) \quad \ldots\ldots\ldots (38)$$

Taking exponential of both sides twice gives,

$$|\beta_t|^{0.1} = \exp\left[0.078734|\beta_{t-1}|^{0.46569}\right] \quad \ldots\ldots\ldots (39)$$

Raising both sides to the power of 10,

$$|\beta_t| = \left[\exp\left[0.078734|\beta_{t-1}|^{0.46569}\right]\right]^{10} \quad \ldots\ldots\ldots (40)$$

since β_t is negative, we write

$$\left[\left[\right.\right.$$

$$\beta^1_t = -\left|\exp\left|0.078734\right|\ \beta^1_{|t-1}\right|^{0.46569}\Big|\Big|^{10} \quad \ldots\ (41)$$

The model was tested for accuracy, and the mean error was found to be - 0.431141 and unstable.

The computed values of β^1_t, Table 2, was subtracted from the extrapolated values, and the results obtained E_t are shown in table 2. The results obtained termed the errors were modelled regressively as illustrated in Appendix 2, and the model obtained was

$$\hat{E}_t = E_t + 1.175E_{t-1} - 0.19875E_{t-2} - 0.125375E_{t-3} + 0.01345E_{t-4} - 0.00252E_{t-5} \quad \text{---------- (42)}$$

With $\hat{E}_t$ being the error in the model at time t, the mean of which after five iterations was found to be - 0.02181, and it was found to decrease as the number of interations increased, and at this level was only about 5 percent, that is

$$\frac{0.02181}{0.43114} \times 100$$

of the initial mean error of $|-0.43114|$ (or modulus of -0.43114).

The error in the model was now modelled using Fourier transform, considering the seasonal length of twelve (12) months in the analysis and the average for the four years on monthly basis therefore used, the following monthly adjustment constants, k, were obtained, Table 5.

TABLE 5:

MONTHLY ADJUSTMENT CONSTANTS $K(E_t)$

January	–	+0.45463
February	–	-0.07349
March	–	-0.02096
April	–	-0.04505
May	–	-0.36742
June	–	-0.32292
July	–	-0.32837
August	–	-0.17273
September	–	+0.16722
October	–	+0.64139
November	–	+0.16037
December	–	-0.10832

In equation (42) above the K value replaces ε_t obtain e_t from the equation, this e_t value is added to β^1_t from equation (41) to get the β_t required for the estimation.

5.20.2 MODEL FOR e_{t-1}

The model for e_{t-1} in the equation for Y_t, Eq. 35, was obtained by considering the first two terms in equation (42) for E_t and seeking for the situation whereby if the data generated for E_t, equation (43), below are used in a Fourier series, the constant term can be approximated to zero.

That is

$$e_t = \varepsilon_t - 1.175 \times 10^{-3} e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (43)$$

was found to be a suitable model in which the mean of ε_t was found to the equal to 0.0000154, and this was approximated to zero, such that

$$e_t = -1.175 \times 10^{-3} e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (44)$$

Since e_t and e_{t-1} by reason of the above equation (44) are both small of the order of 10^{-3} or less, it was decided to set

$$e_{t-1} = - 1.175 \times 10^{-6} e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (45)$$

The error introduced by this approximation and by dropping the contributions of e_t was accounted for in the model for Y_t below.

5.20.3 MODEL FOR Y_t

The model for the Y_t given earlier is

$$Y_t = (0.6009 - 2.9865e_{t-1})\beta_t$$

From given values of Y_t and forecast values of β_t, e_{t-1} values were computed and the average was found to be 0.00385, subtracting this from 0.6009 in the above equation gave

$$Y_t = (0.59705 - 2.9865e_{t-1})\ \beta_t \quad \ldots\ldots\ldots\ldots\ldots\ (46)$$

Substituting for e_{t-1}, the equation (45) above we obtain

$$Y_t = (0.59705 - 2.9865 \times -1.175 \times 10^{-6}e_{t-1})\ \beta_t$$

$$Y_t = (0.59705 + 3.5091 \times 10^{-6}e_{t-1})\ \beta_t \quad \ldots\ldots\ldots\ldots\ (47)$$

This model gives the mean level of water in the river at the gauging station at the different time steps. The results obtained are as tabulated in table 2, and plotted in figure 12.

The 95 percent confidence limits for 0.00385 was also computed and was found to lie within,

$0.0035308 < \bar{e} < 0.0041692$ The results for the lower bound and upper bound 95 percent confidence limits are also tabulated in table 2.

5.21 MODEL VALIDITY

Once the estimation model was obtained, it was decided to test its appropriateness in describing the data by examining the residuals (or errors), table 2. The first five auto correlation coefficients for the errors were computed and the results are as listed below and plotted in figure 13.

$r_1 = 0.462$ $r_2 = 0.171$ $r_3 = 0.187$

$r_4 = -0.008$ $r_5 = -0.082$

The confidence limits for the auto correlation is

$$-0.290 \leq r_k \leq + 0.290$$

which shows that only one auto correlation r_1, is outside the limits; And since theoretically auto correlation values decay exponentially it follows that no other auto correlation value is likely to fall outside the confidence limits.

The model can therefore be said to be able to appropriately describe the data, the location of r_1 outside the confidence limits is expected due to the nature of occurence of storm rainfall which are highly unpredictable, hence floods. The fact that rainfall and not floods are seasonal, the model is not able to estimate the peak values of water levels appropriately.

TABLE 6: RESULTS USING GENERATED DATA

Period t	Month	$\hat{e}_t$=K | E_t	β^1_t	θ_t | Y_t		Actual	Forecast, F^Y_t			Error	%
							Lower Bound	Mean	Upper Bound	e_t	Error
49	Jan. '88	.45463	.85824	-7.12080	-6.35256	36.1	26.4	26.5	26.6	9.6	26.6
50	Feb.	-.07349	.82126	-7.21152	-6.39026	37.8	26.9	27.0	27.1	10.8	28.6
51	Mar.	-.02096	.72961	-7.21218	-6.48257	19.2	28.2	28.3	28.4	-9.1	47.4
52	Apr.	-.04505	.54584	-7.21279	-6.66695	26.3	31.0	31.1	31.3	-4.8	18.3
53	May.	-.36742	.03654	-7.21336	-7.17682	58.1	40.4	40.5	40.7	17.6	30.3
54	Jun.	-.32292	-.47106	-7.21388	-7.68494	46.0	52.4	52.6	52.9	-6.6	14.3
55	Jul.	-.32837	-.94982	-7.21436	-8.16418	77.3	67.1	67.4	67.7	9.9	12.8
56	Aug.	-.17273	-1.19422	-7.21479	-8.40901	-	76.1	76.5	76.8	-	-
57	Sep.	.16722	-.98904	-7.21520	-8.20424	-	68.5	68.8	69.1	-	-
58	Oct.	.64139	-.17072	-7.21557	-7.38629	52.5	45.0	45.1	45.3	7.4	14.1
59	Nov.	.16037	.29448	-7.21591	-6.92143	29.0	35.4	35.5	35.6	-6.5	22.4
60	Dec.	-.10832	.38196	-7.21623	-6.83427	23.0	33.8	34.0	34.1	-11.0	47.8

Furthermore, the x^2 of the auto correlations were computed to be 13.65, while the corresponding value from the table of x^2 at 95 percent confidence level is 63.98. Since x^2 - computed is less than x^2 - theoretical, one can conclude that the model given by equation (47) for Y_t is an approriate one for estimating future values of the time series of water levels in Lokko river.

From table 2, it will be observed that the error in estimating about 29 percent of the data is less than 5cm while the error in estimating about 56 percent of the data is less than 10cm. Similarly, the error in estimating about 73 percent of the data is less than 15cm, and the error in estimating about 81 percent of the data is less than 20cm.

Therefore, depending on the nature of the project to be executed on the river and the degree of accuracy required, a selection can be made from the above listing to help in deciding on the factor of safety to be applied, that is

for 5cm error,

$$\text{factor of safety} > \frac{100}{29} = 3.4$$

for 10cm error,

$$\text{factor of safety} > \frac{100}{56} = 1.8$$

for 15cm error,

$$\text{factor of safety} > \frac{100}{73} = 1.4$$

for 20cm error,

$$\text{factor of safety} > \frac{100}{} = 1.2$$

Above statement is corroborated by the fact that at period 32 (that is at the 32nd month), table 2, the historical level recorded was 152.7cm but only 74.0cm was estimated by the model, resulting in an error of 78.3cm, which means that a dam of height 74.0cm above zero of gauge (94.980m) will be overtopped several times; But if a factor of not less than 152.7/74 = 2.06 is applied, the safety of the dam, buildings and other infrastructures near the dam, farm and farm settlements, human lives as well as those of livestock and other farm produce will be greatly enhanced.

Therefore it is recommended that for estimating peak water levels in the project area which normally occurs between the months of July and September, the factor of safety of

$$\text{F.S} \frac{\text{for 5cm error + F.S for 10cm error + F.S. for 15cm error}}{3}$$

$$= \frac{3.4 + 1.8 + 1.4}{3} = 2.20$$

(F.S being Factor of Safety) should be applied to ensure error restriction to a mean of 10cm.

Historically, we are told that there was flood only in one year of the four years study period, and the year referred to is 1986, with periods 1 to 12 referring to January to December, 1984, periods 13 to 24 refers to January to December, 1985, periods 25 to 36 refers to January to

December, 1986 and period 37 to 48 refers to January to December 1987. The next highest water level to that recorded at period 32 of 152.7cm
is 129.8cm corresponding to the period, 9 (that is, September, 1984), however there were no flooding that year. It follows that for up to 130cm water level floods may not occur. This water level of 130cm is therefore selected as the base water level which if the estimated values with the factor of safety applied is exceeded, is a strong indication of the possibility of flooding.

5.22 USE OF THE MODEL FOR ESTIMATION.

To use the model for estimation, let us consider for the purpose of illustration the estimation for the month of February, 1988. This correspond to the period of 50 in table 2.

The model has been designed such that past information must be preserved (as in table 2) and added to continually, however, only the previous five (5) time steps recorded are needed to estimate (see table 2). The procedure for the estiamtion are enumerated below:

5.22.1 Select the time step (or month) and enter it into column 1. For the purpose of this illustration we choose a time step of 50.

5.22.2 Noting that the time step of 50 correspond to the month o f February, we select the corresponding monthly adjustment constant, E_t = - 0.07349 from the listing on table 5.

5.22.3 Using the Eq (42), given below we obtain

$$\hat{E}_t = E_t + 1.175E_{t-1} - 0.19875E_{t-2} - 0.125375E_{t-3} + 0.01345E_{t-4} - 0.00252E_{t-5}$$

Substituting from table 2,

$\hat{E}_t$ = - 0.07349

E_{t-1} = 0.85824

E_{t-2} = 0.38499

E_{t-3} = 0.29843

E_{t-4} = - 0.16586

E_{t-5} = - 0.98344

into the equation, we have

$E_t = E_{50}$ = - 0.07349 +1.175 x 0.85824
- 0.19875 x 0.38499
- 0.125375 x 0.29843
+ 0.01345 x - 0.16586
- 0.00252 x - 0.98344

E_{50} = - 0.07349 + 1.008432 - 0.076517 - 0.037416
- 0.002231 + 0.002478
= 0.821256 = 0.82126

5.22.4 Using equation (41),also given below, we obtain

$$\beta^1_t = -\left[\exp\left(0.078734 \left| \beta^1_{t-1} \right|^{0.46569}\right)\right]^{10}$$

for t = 50

$$\beta^1_{50} = -\left[\exp\left(0.078734 \left| \beta^1_{49} \right|^{0.46569}\right)\right]^{10}$$

from table 2, $\left| \beta^1_{49} \right| = 7.2108$

Therefore,

$$\beta^1_{50} = -\left[\exp\left(0.078734 \times 7.2108^{0.46569}\right)\right]^{10}$$

$$\beta^1_{50} = 7.21152$$

5.22.5 From $\beta_t = E_t + \beta^1_t$

Therefore,

$$\beta_{50} = E_{50} + \beta^1_{50}$$

Substituting earlier obtained values of E_{50} and β^1_{50} into above equation gives

$$\beta_{50} = 0.82126 - 7.21152$$

$$= -6.39026$$

5.22.6 From equation (47), given below

$$Y_t = (0.59705 + 3.5091 \times 10^{-6} e_{t-1})\ \beta^t$$

where, e_{t-1}, is the measured difference between the computed and the actual water level and the previous time step. Above can be determined in case of estimation for only one time

step ahead, where forecast is to be done for two or more time steps, e_{t-1} is neglected and we use

$$Y_t = (0.59705)\ \beta^t \qquad \text{..... (48)}$$

values so obtained can be adjusted at one time step ahead in the future using previous equation (47). The difference in the results will be of the order of 10^{-4} or less.

Using equation (48) above

$$Y_{50} = 0.59705\ \beta^{50}$$

$$= 0.59705^{-6.39026}$$

Therefore,

$$Y_{50} = 27.0\text{cm}$$

Estimation values obtained for periods 49 to 60 are shown in table 6, and plotted in figure 12.

CHAPTER 6: DISCUSSION OF RESULTS

6.1 GENERAL

River flow level estimation is an art involving the use of the theory of hydrology or the statistical theory to estimate values for future events. In this study the statistical theory was adopted. The study was aimed at providing a model which will serve as an advance warning system of flood occurence in the Lokko river basin;

A consideration of the topographical heights of the area showed that when the flow level has risen up to 96.23m above the ordinance datum (or 1.30m on the reference scale) and it is still rising it follows that flooding is likely to occur. This sort of forecasting according to Hurst (1956) hardly explains more than 70 percent of the behaviour of the hydrological process because of the complex nature of the hydrological phenomena, however, it is very important as it helps in reducing the potentially disastrous effects of floods as well as providing needed information for water resources planning. The methodology discussed in the study, is the well documented method of Box and Jenkins (1970) for univariate models, discussed earlier; which according to Jamieson and Wilkinson (1972) cannot account for over 45 percent of the series, the main limitation being that they have no parameters representing real conditions and,

therefore cannot account well for incidences of convective storms which results in high flow .pm level or flash flood.

4.2 ON RANDOMNESS OF DATA

The autocorrelation analysis of the historical data shows that the data generating the autocorrelations are non-random; this was confirmed with the χ^2-test which gave values of the computed χ^2- test to be greater than the theoretical χ^2 -test values. A good number of autocorrelation values were outside the random limits, indicating that there is some pattern in the historical data, a sinusoidal repetition in the pattern was observed after every 12 periods (months), suggesting the presence of yearly pattern accounting for the non-randomness of the data.

In order to confirm that there is a seasonal pattern of 12 periods (months), it was necessary to ensure that the data were stationary; the autocorrelations, Figure 7, some what fluctuate about the theoretical mean for a random autocorrelated series of zero mean, it dropped to zero after the second value, and it does not have a noticeable trend, hence it was concluded that the historical data are stationary and spurious autocorrelations are not expected. The twelve period or twelve months seasonal length was therefore confirmed.

The power transformation of the Box and Cox (1964), yielded a transformation of $X_t = Y_t^{-0.13347}$ with a minimum skewness of - 0.012. The data generated from the above transformation gave auto correlation coefficients, some of which were found to lie outside the confidence limits for random data, in addition, the computed χ^2 test values were found to be greater than the theoretical χ^2 - test values, indicating that the data were still non-random after the transformation.

The twelve months pattern persisted with the exponential decay of the auto correlations with increasing period. A new series was therefore generated from the product of the period to - period differencing and season-to-season differencing. Auto correlation analysis of this new series shows that all the auto correlation coefficients lie within the random limits, except r_1, which means that the new series are random, this was confirmed by the computed value of the χ^2 test which was found to be lower than the theoretical χ^2 - test value. The sinusoidal and exponential decay pattern in the original data was eliminated by the differencing, and the auto correlation functions r_8, r_{20} and r_{32} corresponding to the month of August shows the incidence of high water levels in this month and also confirms the twelve months seasonal pattern.

4.3 ON STATIONARITY OF DATA

The auto correlation analysis of the historical data, figure 7, shows that the auto correlation coefficients somewhat fluctuate about the theoretical mean for a random auto correlated series of zero mean, it drops to zero after the second value, and it does not have a noticeable trend; Therefore, it was concluded that the historical data are stationary.

After the Box and Cox (1964) transformation, it was observed that the data were still stationary, Figure 9; However, the fluctuations of the auto correlation about the mean was not perfect and, it decays exponentially, as with the case of the historical data. This was believed to have introduced spurious auto correlation in the data.

To eliminate this one short term and one long-term differencing were performed to remove period-to-period and season-to-season non-stationarity that could introduce spurious auto correlations in the data. Figure 10, shows that the differenced-transformed data are stationary as the auto correlations dropped to zero after the third period (month).

4.4 ON IDENTIFICATION OF TENTATIVE ARMA(p,q) MODEL.

Some partial auto correlations were calculated until values obtained were no longer significantly different from

zero, the order of the auto regressive process (AR) corresponding to the number of partial auto correlations that are significantly different from zero. Figure 11, shows that only one partial auto correlation ϕ_1, is significantly different from zero, therefore, the order of the estimation model was identified as AR(1).

The trailing off which is usually sought for from the partial auto correlation plot in order to confirm the presence of a moving average (MA) process is rather difficult to notice, Figure 11; However, a general exponential decay could be noticed, and therefore it was assumed that an MA process was also present. The order of which was sought for from the auto correlation plot, Figure 10, by noting the number of auto correlations which are significantly different from zero. Since only one auto correlation value is significantly different from zero, it was concluded that the order of the estimation model is also MA(1).

By combining the two models identified above according to Box-Jenkins (1970), an appropriate model to describe the process was found to be ARMA(1,1), where $p = 1$, and $q = 1$.

6.5 ON ARMA(1,1) MODEL PARAMETER ESTIMATES.

The model expression was obtained to be

$$X''_t = \phi_1 X''_{t-1} + e_t - \theta_1 e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots (12)$$

where

$X''_{t-1} = (Y_{t-1} - Y_{t-2})(Y_{t-1} - Y_{t-17})$

$X''_{t} = (Y_{t} - Y_{t-1})(Y_{t} - Y_{t-12})$

ϕ_1 = AR Parameter

θ_1 = MA Parameter

e_t = error at t period

e_{t-1} = error at t-1 period

By applying the backshift operator, and accounting for the seasonal parameter, ϕ_{12}; The general model expression become

$$X_t = X_t X_{t-1} + \phi_1 X_t X_{t-2} - \phi_1 Y_t Y_{t-3}$$
$$+ \phi_{12} X_t X_{t-12} - \phi_{12} Y_t Y_{t-13} - \phi_1 Y_t X_{t-11}$$
$$+ \phi_1 Y_t Y_{t-15} + e_t - \theta_1 e_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (50)$$

The above expression was solved for the values of ϕ_1, ϕ_{12} and θ_1 which induces the minimum mean square error, by the method of decomposition, and the results obtained were

$\phi_1 = -0.025$, $\phi_{12} = 0.750$ and $\theta_1 = -0.975$.

These values were substituted into the general expression, equation (50), above and the following equation was obtained.

$$Y_t = X_t X_{t-1} - 0.025X_t Y_{t-2} + 0.025X_t Y_{t-3}$$
$$+ 0.750X_t X_{t-12} - 0.750X_t Y_{t-13}$$
$$+ 0.025X_t Y_{t-14} - 0.025X_t Y_{t-15}$$
$$+ 0.975e_{t-1} + e_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (51)$$

Values of X's were substituted into the equation (51), for

the various period setting an initial value of zero for e_o

Results obtained shows that there is a general equation of the form,.is

$$X_t = a_o + a_1{}^1 e_{t-1}$$

or

$$X_t = a_o + a_1 e_{t-1} \quad \text{..... (52)}$$

by setting $a_1{}^1 = a^1$

100 ..... (53)

This equation shows that the water level in the river at any time is dependent on the immediate past recorded error between the estimated value and the measured value.

The model, equation (52), parameter values a_o and a_1 were obtained using the method of least squares, and the values obtained were

$a_o = 0.6009$ and $a_1 = -2.9865$.

The estimated mean error was computed to be 0.00046.

The model was therefore given as

$$X_t = (0.6009 - 2.9865e_{t-1})^{0.8998} \quad \text{.....(54)}$$

after Box - Cox transformation.

Since,

$$X_t = Y_t{}^{-0.13347} \quad \text{..... (55)}$$

by appropriate rearrangement

$$Y_t = (0.6009 - 2.9865e_{t-1})^{-6.6667} \quad \ldots\ldots\ldots\ldots (56)$$

was obtained. This model was found to be grossly inadequate to describe the system as instability was obtained in the system, when sensitivity analysis was performed on the system, by varying e_{t-1}.

It was therefore recognized that model modification was necessary.

6.6 ON THE MODIFICATION OF THE MODEL

In view of the inadequacies of the model in estimating future events due to instability, a general model of the form,

$$Y_t = (0.6009 - 2.9865e_{t-1})^{\beta t} \quad \ldots\ldots\ldots\ldots\ldots\ldots (57)$$

was adopted in which Y_t, e_{t-1} and β were modelled separately.

Model result obtained for β_t was

$$\beta_t = \beta^1_t + E_t \quad \ldots\ldots\ldots\ldots\ldots\ldots (58)$$

where,

$$\beta^1_t = -\left|\exp\left|0.07073\right| \beta^1_{t-1}\right|^{0.45569}\Big|\Big|^{10} \quad \ldots\ldots\ldots\ldots (59)$$

$$E_t = E_t + 1.175\, E_{t-1} - 0.19875\, E_{t-2} - 0.125375\, E_{t-3} + 0.01345\, E_{t-4} - 0.00252\, E_{t-5} \quad \ldots\ldots\ldots\ldots\ldots (60)$$

values of E_t for the various months of the year were determined using Fourier transform and are given in table 6, and are referred to as the monthly adjustment constants.

The error term in the expression for Y_t, equation (57), was modelled by considering only the first two terms of the expression, equation (60) above, and replacing E_t by e_t, and the expression $e_t = \varepsilon_t - 1.175 \times 10^{-3} e_{t-1}$ ..... (61) was used to model e_t. This expression is different from equation (60) which only gave an idea of its possible form. The expression was found suitable to describe e_t as the mean error ε_t was found to be equal to 0.0000154 and was equated to zero, hence the equation (61) was approximated to $e_t = 1.175 \times 10^{-3} e_{t-1}$ ..... (62)

noting that both e_t and e_{t-1} are small of the order of 10^{-3}, e_{t-1} was set as $e_{t-1} = -1.175 \times 10^{-6} e_{t-1}$ ..... (63) such that e_{t-1} can be approximated to zero, the error incurred is accounted for in the model for Y_t below..

In order to finally model Y_t, e_{t-1} values were computed from known values of Y_t and β_t estimate, the average value of e_{t-1}, termed error term, ε_{t-1} was found to be 0.00705, this value was subtracted from 0.6009 in the equation for Y_t to obtain.

$$Y_t = (0.59705 - 2.9865 e_{t-1}) \beta t$$

and

$$e_{t-1} = e_{t-1} + 0.00129 \quad \text{..... (64)}$$

and

$$0.6009 = 2.9865(e_{t-1} + 0.00129)$$

or $0.6009 - 2.9865e_{t-1} = 0.00385$ ..... (65)

or $0.59705 = 2.9865\, e_{t-1}$

in which case the error term is considerably strapped to close to zero, except for positive deviations which would add up to 0.00129 to give higher values.

Further strapping was achieved by substituting equation (63) for e_{t-1} to obtain

$$Y_t = (0.59705 + 3.5091 \times 10^{-6} e_{t-1}) \quad \text{..... (66)}$$

This model gives the mean level of water in Lokko River at the gauging station for various periods (months), but at only one step ahead.

This model was further simplified by ignoring e_{t-1} as its contribution is insignificant and simply writing

$$Y_t = (0.59705)^{\beta t} \quad \text{..... (67)}$$

This simplified expression apart from its compactness have an added advantage of extreme importance in that it can be used to estimate for any number of time steps ahead, whereas, the former equation (66) is only suitable for one step ahead. Estimation results obtained are listed in table 3, and plotted in figure 12.

6.7 ON THE VALIDITY OF THE MODEL

Two standard validity tests were carried out on the model to determine its appropriateness in describing the data, in order to carry out the tests the estimation errors (or residuals) were first computed by comparing estimated values obtained from equation (67) with the historical (or known) data, as shown in table 2.

Some auto correlation values were computed and

6.7(a): Plotted in figure 13 shows that only one auto correlation, r_1 is outside the confidence limits of $-0.290 \leq r_k \leq +0.290$, and since theoretically auto correlation values decay exponentially, it follows that no other auto correlation value will fall outside the confidence limits. This model was therefore considered suitable to appropriately describe the data.

6.7(b): The Box-Pierce (1970) x^2 - test was performed in order to confirm the appropriateness of the model to describe the data as stated in 6.1.6(a) above; The computed x^2 - test value was 17.65 while the theoretical x^2 - test value at 95 percent confidence level was 67.93. Since x^2 computed is less than x^2 theoretical it follows that the model is an appropriate one for describing the

historical data and for estimating future events.

6.0 ON TABLE OF RESULTS

In table 2 are given listings of raw information, and estimation results, column 1, gives the listing of the periods, while column 2 gives the corresponding months and year, with period 1 corresponding to January, 1 1984 and period 60 corresponding to December, 1988. The table have been prepared in such a way that continuation can be achieved and stored in this comport form in office file where a computer facility is available a simple program can be developed to do the computations from model equations discussed earlier and to tabulate them as presented.

Column 3, is the adjustment constant, E_t and it repeats itself after every 12 months, column 4, gives the error ε_t in the index of the power transformation, $\beta_{,t}$ as computed from the linear approximation, β^1_t column 5; β^1_t has a gentle negative slope, while E_t oscillates sinusoidally from positive values to negative values to correct the error in β_t. Column 6 gives the computed values of β (power index) in which the starting value of - 5.64766 was selected in order, to eliminate error from $\beta_{,t}^1$ such that E_t is equal to zero.

Column 7, are the actual (historical) data, while column 8b gives the mean of the estimated data, columns 8a

and 8c gives the lower and upper bounds of the estimated data. Column 9, gives the magnitude of the estimation error, meansured in contimetres, same as the un't of the actual data, while column 10, gives the percent estimation error.

It was observe that the percent error computed from the equation

$$\% \text{ Error} = \left|\frac{\Delta_t}{Y_t}\right| \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (48)$$

varies from 3.4 to 77.6, this is expected as the complex nature of hydrologic phenomena according to Hurst (1956) makes it difficult for such models to account for more than 70 percent of the processes of hydrology, furthermore, acording to Jamieson and Wilkinson (1970), such a stationary model can only explain 45 percent of the rainfall time series; since rainfall are primary causes of the runoff processes, it follows that the model may not account for more than 0.70 x 0.45 x 100 percent, that is 31.5 percent, therefore errors up to 68.5 percent are expected during storm rainfall.

In view of the errors anticipated in the estimated values it was thought of to apply some safety factors to the values obtained before their use, in addition to this consideration is the error allowable in the estimate.

In consideration of the error allowable for the purposes of engineering design works, a 5cm error was determined to attract a factor of safety of 3.4, while a 10cm error would attract the application of 1.8 factor of safety, similarly, 15cm error would attract 1.4. In order to remove any ambiguity arising from the use of the factors of safety, the error margin was limited to 10cm and the general factor of safety fixed at a minimum of 2.2.

The factor of safety, 2.2, is believed would when applied to a dam project the dam would impound the maximum flood, for example, the maximum flood level recorded during the period under investigation is 152.7cm, while the estimated value is 74.0cm, applying a factor of safety of 2.2 to this would provide a level of 74.0 x 2.2 cm = 162.8cm. At this level the dam is not likely to be overflooded. However, for it to serve as an early warning system this factor of safety is too high, rather, a continuous observation of flow level changes is believed would serve the much needed purpose, for this reason, a height of 130cm is fixed as the upper most level to which the water could rise, but if it continues to rise all necessary alarm should be raised to avert disaster.

That not withstanding an advance information can be obtained from the estimated values, this is possible by multiplying the estimated value by a factor of safety of

1.6, figure 14, if the value obtained is greater than 130.0cm, all necessary emergency preparations should be made to combat the possible flood or to avert disaster.

(d) A minimum of four (4) years daily records should be used for any further study in the project area, since the frequency of flood occurence is three (3) years.

(e) The intervention analysis, as proposed by Box and Tiao (1975) should be applied to account for changes in the model response due to unexpected events of rainstorms.

(f) A procedure be incorporated in the model development for identifying when error are no longer random, such as, the Trigg (1964) monitoring system (or Tracking signal) in order to minimize errors in the model.

7.2 CONCLUSIONS

From the methodology discussed and the analytical solution to the problems proposed in this study as well as the explanations to the discrepancies observed and their analyses, the following conclusions can be drawn:

(a) An approximate methodology has been discussed which is applicable to other river basins, as results obtained shows close agreement with observed data and theoretical considerations. The error in estimating about 29 percent of the data is less than 5.0cm, while the error in estimating about 56 percent of the data is less than 10cm; this is in line with the theoretical consideration that such stationary

models can only explain 71.5 percent of the data.

(b) The model can approximately describe the data, as the autocorrelation analysis of the residuals shows that they are random and this was confirmed by the use of the χ^2-test.

(c) For purposes of engineering works a factor of safety of 2.20 should be applied to the maximum recorded (estimate) water level, and for the purposes of early warning a factor of safety of 1.6 should be applied to the corresponding estimated water level.

R E F E R E N C E S

Abel, P.(1979): Programming assembler language. Reston publishing Company, Inc. Reston. Virginia.

Aitchison, P.J., Bawden, M.G., Carrol D. M., Glover, P.E., Klinken berg; K., Leeuw, P.N. De., and Tuley, P. (1972): The Land Resources of North East Nigeria: The Environment. Vol. 1 No 9.

Aitchison, P.J., and Glover, P.E. (1972) The Land Resources of North East Nigeria: Tsetse and Trypanosomiasis. Vol. 2 Land Resources Study No. 9.

Akaike, J. (1974): Stochastic theory of minimal realization. IEEE Tran. Autom. control19(6):667-674

Albrecht, R.L., Finkel, L., and Brown, J.R.(1978) : Basic. 2nd edition John Wiley and Sons, Inc. New York.

Anderson, R.L (1942) : Distribution of serial correlation coefficents. Annals of mathematical Statistirs, Vol. 13, pp 1-13.

Bajpai, A.C., Calus, I.M., Fairley, J.A., and Walker, D.(1973): Mathematics for engineers and Scientists. Volume 2. John Wiley and Sons. London.

Bartlett, M.S.(1946): On the theoretical specification of sampling properties of auto correlated time series. Journal of the Royal Statistical Society, series B, Vol. 8, P. 27.

Bartlett, M.S. (1966): An introduction to stochastic processes. Combridge University press.London. Bawden, M.G., Carroll, D.M., and Tuley, P.(1972): The Land Resources of North East Nigeria. The land System. Vol. 3 Land Resources Study No.9.

Bawden, M.G., Carroll, D.M., and Tuley, P. (1969): Preliminary descriptions of the Land System of North East Nigeria. Unpubl. prelim. Rep.Land Resources Division Dir. Overseas Surv. No. PR/6/68.

Betson,R.P., Turker, R.L., and Haller, F.M. (1969): Using analytical method to develop a surface run off model. Water Resources. Res.; Vol. 5, PP.103 - 111.

Bigelstone, H.J. (1958): Harmattan haze at Kano. Tech. Note Br. W. Afr. Met. Serv. No. 10.
Biswas, A.K.(1976): Systems approach to water management. McGraw Hill Book Company. New York.

Blackman, R.B. and Tukey, J.W.(1959): The measurement of power spectra. Dover, New York.

Bolch, B.W., and Huang, C.J.(1974): Multivariate Statistical methods for Business and Economics. Prentice-Hall International series in management. New Jersey.Box, G.E.P., and Cox, D.R. (1964): An analysis of Transformations.Journal of the Royal Statistical Society, Vol. 26, series B, pp. 211-53.

Box, G.E.P., and Jenkins, G.M. (1970): Time series analysis, forecasting and control. San Francisco. Holden-Day.

Box, G.E.P., and Jenkins, G.M. (1976): Time series analysis, forecasting and control. Revised Edition. San Francisco. Holden-Day.

Box, G.E.P., and Pierce, D.A. (1970); Distribution of residual autocorrelations in Autoregressive integrated moving average time series models. Journal of the American Statistical Association Vol. 65, pp. 1509-26.

Box, G.E.P., and Tiao, G.C.(1975): Intervention analysis with applications to economics and environmental problems. Journal of the America Statistical Association Vol. 70,No.349, pp.70-79.

Bras, R.L.(1976): Short - term forecasting of rainfall and run off. IIASA proceeding series. pp. 125 - 137.

Bras, R.L., and Rodrigues-Iturbe, I. (1975): Rainfall run off as spatial stochastic process: Data collection and Synthesis. Hydrodynamics Lab. report No. 194, M.I.T. Dept. of Civil Engineering Combridge, Massachusetts. U.S.A.

Bras, R.L. and Rodrigues - Iturbe, I. (1976): Rainfall generation:Anonstationarytimevarying multidimensional model. Water Resources. 12(3).

Burges, S.J. (1970): Use of Stochastic hydrology to determine storage requirements for reservioirs - A critical analysis, Engineering - Economic Planning. Dept. No. 34, Dept. of Civil Engineering Standford University, Standford. U.S.A.

Carter, J.D., Barber, W., Tait, E.A., and Jones, G.P. (1963): The geology of parts of Adamawa, Bauchi and Borno provinces in North - Eastern Nigeria. Explanation of 1: 250,000 sheets Nos. 45, 46 and 47, Bulletin geological Survey, No. 30.

Chow, V.T. (1956): Hydrologic studies of floods in the United States. International Association Scientific Hydrology Publication 42, 134-70

Ciriani, T.A., Maione, U. and Wallis, J.R. (Edited 1974) Mathematical models for surface water hydrology. Proc. of the workshop held at IBM Scientific Centre, Pisa Italy. John-Wiley and Sons. London.

Clark, C.O. (1945): Storage and Unit hydrograph. Transaction American Society Civil Engineering, Vol. 110, pp. 1419-1488.

Cocheme, J. and Franquin, P. (1967): An agroclimatology Survey of a semi - arid area in Africa South of the Sahara.Technical Note World Meterological Orgasation. No. 86.

Cook, H.L. (1946): The infiltration approach to the calculation of surface run off. Transaction American Geophysical Union, Vol.27, pp. 726 - 747.

Consullint (1975): Water survey of North-east State of Nigeria.Engineering project report, Federal Republic of Nigeria, Upper Benue River Basin Development Authority.

Daniel, P.M., and Thompson, M.(1973): Mathematical models and applications, Prentice-Hall, Inc. Eaglewood cliffs, New Jersey.

Dar al Handasah (1980): Feasibility study for dam irrigation project on the river Song, Federal Government of Nigeria, Upper Benue River Basin Development Authority..

Darcy, H (1856) : Flow properties of water through layers of soil, Journal on soil studiespp.570, 590 - 594, V.Dalmont,Paris.

Diddy, R.L., and Page, R.L. (1974): Fortran for Humans, West publishing Co. New York.

Dorrell, A.T. (1947): A Severe Squall at Kano, Northern Nigeria. Metcorological Magazine. 76, 206-207.

Eckhouse, R.H. (Jr.), and Morris, L.R. (1979): Minicomputer systems Organization, programming and application (PDP-11). 2nd ed. Prentice - Hall Inc. New Jersey.

Fuller, W.E. (1914) : Flood flows. Transaction American Society Civil Engineering, 77, 564 - 617.

Grayson, W.M. and Eagleson, P.S. (1971): Evaluation of radar and raingauge systems for flood forecasting. Report No. 138. Ralp, M. Parsons, Laboratory, Massachusetts Institute of Technology.

Hannan, E.J. (1960): Time series analysis. Mathuen. London.

Hannan, E.J. (1970): Multiple time series. Wiley. New York.

Hashidu, A. Alh (1990): Hydro-Niger Project. Daily times, Nigeria, Nov. 22.

Hewitt, K., and Burton, I.(1970): The hazardousness of a place. A regional ecology of damaging events. University of Toronto, Dept. of Geology Resource Publication, 5, 154.

Hewlett, J.D., and Nutter, W.L. (1969): An outline of forest hydrology. University of Georgia Bulletin Athens.

Holmes, A.(1944): Principles of Physical Geography. T. Nelson and Sons.London

Horton, R.E. (1945): Erosional development of streams and theirdrainagebasins;hydrophysical approach to quantitative morphology.Geological Society American Bulletin, Vol. 56, pp.275 - 370.

Hoyt, W.G. (1942): An outline of the run off cycle.Pennsylvania State College Tech. Bull. 27, pp. 57 - 67.

Hughes, C.E., (Pfleeger, C.P.)and Rose, L.L. (1978): Advanced programming techniques; A second course in programming using Fortran. John Wiley and Sons, Inc. New York.

Hurst, H.E. (1951): Long-term storage capacity of reserviors. Transaction American Society Civil Engeering.; 116,770 - 799.

Hurst,H.E. (1956):Methods for using long-term storage in reserviors. Proceedings of the Institute of Civil Engineers 1, 519-543..

International Association for Scientific Hydrology (1974): Flash floods. Proceedings of the Paris Symposium. IASH/UNESCO/WMO.

Jacobson, R.R.E., Snelling, N.J., and Truswell, J.F. (1963): Age determination in the geology of Nigeria with special reference to the older and younger granites. Overseas Geology Ministry Resources. 9.

Jamieson, D.G., and Wilkinson, J.C. (1972): River Dee research program 3; A short term control strategy for multi- purpose reservior systems. Water Resources Research 8(4): 911-920.

Jenkins, G.M., and Watts, D.G. (1968): Spectral analysis and its applications. Holden-Day, San Francisco. U.S.A.

Johanson, R.C. (1971): Precipitation network requirements for stream flow estimation. Technical Report Dept. of Civil Engineers, No. 147 Stanford University.

Kashyap, R.L., and Rao, A.R.(1976): Dynamic Stochastic Models from empirical data. Academic press. New York.

Kendal, M.G. (1963): The advanced theory of statistics. Vol.1. 2nd ed. Butter and Tanner Ltd. London.

Kendal, M.G.(1967): The advanced theory of statistics. Vol. 2. 2nd ed. Butter and Tanner Ltd.London.

Kendal, M.G. (1968): The advanced theory of Statistics. Vol. 3. 2nd ed. Butter and Tanner Ltd. London.

Kirkby, M.J., and Chorley, R.J.(1967): Through flow, overland flow and erosion.Bulletin of the International Association Scientific Hydrology, 12, 5-21.

Klinkenberg, K., Tomlinson, P.R., Higgins, G.M., and Leeuw P.N. de (1963): The soils of the middle Gongola area: Samaru soil survey Bulletin No. 21.

Kohler, M.A. and Linsley, R.K. (1941): Predicting run off from storm rainfall. U.S.A. weather Bureau Research paper 34.

Kraeger, B. (1971): Stochastic monthly streaeflow by multi-station daily rainfall generator. Tech. Rept. 152., Dept. of Civil Engineering, Standford University.

Kresge, R.F. and Nordenson, T.J. (1955) : Flood frequencies derived from river forecasting In: Proceedings of the American Sococity Civil Engineerings, separate no. 630.

Lambert, W.G. and Millard, A.R. (1969): Atra-hasis: The Babylonian story of the flood. Oxford University press London.

Land Resources Division Commonwealth Overseas Development Administration (1972): Land resources of North-Eastern Nigeria. Land resources study No.2 Vols.1-5. Federal government of Nigeria, Upper Benue River Basin Development Authority.

Leeuw, P.N.de., Lesslie, A., and Tuley, P.(1972): The Land Resources of North East Nigeria; Present and potential land use. Vol. 4. Land Resources study No. 9.

Lehman, R.S.(1977): Computer simulation and modelling; An introduction. Laurence erlbaum associates, Hills-dale, New Jersey.

Leopold, L.B., Wolman, M.G., and Miller, J.P. (1964): Fluvial processes in geomorphology. Freeman, San Francisco.

Linsley, R.K., and Ackermann, W.C.(1942): A method for predicting the run off from rainfall. Transion American Society Civil Engineering vol. 107, pp. 825-842.

Linsley, R.K. and Crawford, N.H. (1960): Computation of synthetic streamflow record on a digital computer. International Association Science Hydrology, Publication. No. 51, pp.526- 538.

Linsley, R.K. Kohler, M.A., and Paulhus, J.L.H. (1958): Hydrology for Engineers McGraw Hill Book company, New York.

Long, A.J. (1975): An extended version of the adaptive filtering forecasting model.Unpublished working paper.

Makridakis, S., and Wheelwright, S.C., (1978):Forecasting: Methods and Applications John Wiley and Sons, New York.

Mandelbrot, B.B., and Wallis, J.R. (1969a): Computer experiments with fractional gaussian noise. Part 1 - Averages and variances. Water Resources Research, 5,1, 228 - 241.

Mandelbrot, B.B., and Wallis, J.R. (1969b): Computer experiments with fractional gaussian noise. Part 2 - Rescaled ranges and spectra. Water Resources Research, 5.1, 242 - 259.

Mandelbrot, B.B., and Wallis, J.R. (1969c): Computer experiments with fractional gaussian noise. Part 3 - Mathematical appendix. Water Resource Resources., 5,2, 321-340.

Mandelbrot, B.B., and Willis, J.R. (1967d): Some long - run properties of geophysical records. Water Resources Research, 5,2, 321 - 340.

Mandelbrot, B.B., and Wallis, J.R. (1969e): Robustness of the rescaled range R/S in the measurement of non-cyclic long-run statistical dependence. Water Resources Research, 5,5, 967 - 988.

Mandelbrot,B.B., and Wallis, J.R. (1971): A fast fractional gaussian noise generator. Water Resources Research, 7,3, 543 - 553.

Mann, A.B., and Wald, A. (1943): On the statistical treatment of linear stochastic difference equations. Econometrica; 173 - 220.

Matalas, N.C., and Huzzen, C.S. (1967): A property of the range of partial sums.Proceeding International Hydrology Sympossium. Fort Colins, Colorado state University. Vol.i, pp 252 - 257.

Matalas, N.C., and Wallis, J.R. (1971): Statistical properties of Multivariate fractional noise processes. Water Resources Research, 7,6, 1460 - 1468.

May, K., and Hanson, A.R.(1978):Fundamentals of Computing Science. Prentice-Hall Inc., New Jersey.

McCuen, R.H. (1975): The anatomy of the modelling process. In: Proceedings of the International conference Process. on modelling held at University. Southampton, England, Sept. 8 - 12.

Mckewn, H.P.J. (1958): Dust concentration in the Harattan. Journal of the International Research of the Meteorological society 84, 280 - 282.

Mejia, T.M., Rodriguez - Iturbe, I., and Dawdy, D.R. (1972): Simulation 2; The broken line process as a potential model for hydrologic simulation. Water Resources Research, 8,4, 931 - 941..

Meyer, R.E. (1915): Estimation of storm runoff from Water balance calculation. Poceedings of the Institute of Civel Engineers, Vol. 137

Miljukov, P.I. (1972):State and trends in hydrologial forecasts; status and trends of research in hydrology 1965 - 1974. UNESCO, Paris, pp. 85 - 100.

Morgan, C.W. (1966): Thunderstorms Texas Water Development Board Report 33, 31 - 44.

Mulvaney, T.J. (1850 -51): On the use of self registering rain gauge and flood gauges in making observations of the relations of rainfall and flood discharges in a given catchment. Proceedings of the Institute Civil Engineers vol. 4, pp.18-31.

Nash, J.E., and Sutcliffe, J.V. (1970) River flow forecasting through conceptual models. Part 1 - A discussion of pinciples. Journal of Hydrology 282 - 90.

Nayfeh, A.H., and Mook, D.T. (1979): Non-linear Oscillations. John Wiley and Sons Inc. New York.

Naylor, T.H., Balintfy, J.L., Burdick, D.S., and Chu, K. (1969): Computer Simulation Techniques. John Wiley and Sons, Inc. New York.

Nemec, J.(1964): Engineering hydrology Mc. Graw Hill London.

Nigeria, Geological Survey Department (1964): Geological map of Nigeria. Federal Surveys, Lagos.

National Oceanic and Atmospheric Administration Revised (1977): Floods, Flash floods and warnings. U.S. Dept. Commerce; NOAA/PA 71009.

O'connell, P.E. (1971): A simple stochastic modelling of Hurst's law. Proc. Int. symposium on mathematical models in hydrology; Warsaw. I.A.H.S.

O'connell, P.E. (1974): Stochastic modelling of long-term persistence in stream flow sequences. Ph.D thesis, Imperial college, University of London, pp. 284.

Ogden, G.M. (1970): Politics of water resources development in social and ecological aspects of irrigation and drainage. American Society of Civil Engineers, New York, pp. 353 - 375.

Oranugo, J.O. (1990): State of preparedness against natural disasters, Floods in urban and rural areas. Unpublished proc. of the 3rd National Workshop/exhibition on water resources and environmental pollution: Durbar Hotel Kaduna,

Ott, R. (1971): Streamflow frequency using stochastically generated hourly rainfall.Technical Report No. 151, Dept.of Civil Engineering. Stanford University.

Overton, D.E., and Meadows, M.E.(1976): Stormwater modelling. Academic press New York.

Oyewoye, M.O. (1964): Geology of the Nigeria basement complex. Journal Ministry of Geology. 1, 87 - 102.

Parzen, E. (1957): On consistent estimates of the spectrum of a stationary time series. Annals of Mathematical Statistics. 28: 329 - 48.

P.F.P. (Nig) Kennard and Lapworth, and Pell, Frischman and partners (1980): Kilange river basin pre feasibility study. Vol.1. Report., Federal Government of Nigeria, Upper Benue River Basin Development Authority.

Quenouille, M.H. (1949): The joint distribution of serial correlation coefficients. Annals of mathematical statistics, vol. 20. pp. 561 - 71.

Revised Standard Version:Genesis 7. The Holy Bible.

Richardson, C.W. (1977): A model of stochastic structure of daily precipitation over an area. Ph.d. Thesis, Colorado State University, Fort Collins, Colorado. U.S.A.

Rockwood, D. M. (1969): Application of Streamflow synthesisandreservoirregulation-"SSARR"-Program to the lower Mekong river. The use of analog and digital computers in hydrology, IASH - UNESCO, vol.1, pp 329-44.

Rosso, R., Roth, G., and Siccardi, F. (1983): Possible hydrological changes of the lower shabelle river (East Africa) produced by upstream land reclamation projects. Proceedings of the XX congress AHR - AIRH Moscow, September 5 - 7.

Rostvedt, J.O., et al(1968): Summary of floods in the United States during 1963. U.S. geological Survey, Water supply paper., 1870 -B.

Scheidegger, A.E. (1970): Theoretical geomorphology, 2nd rev. ed. Springer-verlag. Berlin.

Schulz, E.F., Koelzer, V.A., and Mahmood, K.(ed.1972): Floods and Droughts. Proceedings of the second International symposium in hydrology. Fort Collins, Colorado. U.S.A.

Sheaffer, J.R. (ed. White, G.F.1961): Flood - to -peak intervals. Papers on flood problems University of Chicago, Department of Geology Research paper 70, 95 - 113.

Sherman, L.K. (1932): Stream flow from rainfall by the unit graph method. Eng. New - Record. Vol. 108, pp. 501 - 505.

SOTESA (1990): Rural economic sample survey. Federal Government of Nigeria, Upper Benue River Basin Development Authority.

Sugawara, M.(1961): On the analysis of run off structure about several Japanese rivers. Japanese Journal Geophysical, 2.

Snyder, F.F.(1938): Synthetic unit hygrographs. Transaction American Geophysical Union, vol. 19(1), pp. 447 - 454.

Strahler, A.N. (1957): Quantitative analysis of water shed geomorphology. Transaction American Geophysical Union 38, 279 - 300.

Thomas, H.A., and Fiering, M.B. (1962): Mathematical synthesis of stream flow sequences for the analysis of river basin by simulation. In A. Mass.et al., Design of water reservior systems. Havard University press. Cambridge Mass.

Trigg, D.W. (1964): Monitoring a forecasting system. Operational research quarterly. volume 15, pp. 271 - 74.

Tuley, P. (1972): Climate. The Land Resources of North East Nigeria. The Environment. vol. Land Resources study No. 9.

Wheelwright, S.C., and Makridakis, S. (1977): _Forecasting methods for management._ 2nd edition. John Wiley and Sons, New York.

White, P.(1951): _Wedderburn bonding in time series analysis._ Almquist and Wiksell, Uppsala, Sweden.

White, P.(1952): Tests of fit in time series. _Biometrika_ 79, 209 - 213.

White, P.(1952a): The simulation estimation of a time series harmonic and covariance structure. _Trab. Estad_ 3: 12 - 87.

White, P.(1954): Statistical investigation of sunspot observations. _Astronomical physics Journal_ 120: 251 -260.

Wheelwright, S.C., and Makridakis, S. (1977): _Forecasting methods for management._ 2nd edition. John Wiley and Sons, New York.

White, P.(1951): _Wedderburn bonding in time series analysis._ Almquist and Wiksell, Uppsala, Sweden.

White, P.(1952): Tests of fit in time series. _Biometrika_ 79, 209 - 213.

White, P.(1952a): The simulation estimation of a time series harmonic and covariance structure. _Trab. Estad_ 3: 12 - 87.

White, P.(1954): Statistical investigation of sunspot observations. _Astronomical physics Journal_ 120: 251 -260.

Wheelwright, S.C., and Makridakis, S. (1977): _Forecasting methods for management._ 2nd edition. John Wiley and Sons, New York.

White, P.(1951): _Wedderburn bonding in time series analysis._ Almquist and Wiksell, Uppsala, Sweden.

White, P.(1952): Tests of fit in time series. _Biometrika_ 79, 209 - 213.

White, P.(1952a): The simulation estimation of a time series harmonic and covariance structure. _Trab. Estad_ 3: 12 - 87.

White, P.(1954): Statistical investigation of sunspot observations. _Astronomical physics Journal_ 120: 251 -260.

Tuley, F(Ed. 1972): _The Land Resources of North East Asia_ Vol. 5. Land Resource study No. 9.

UNESCO/IASH/WMO(1967): Floods and their computation. _Proceedings of the Leningrad Symposium_ Unesco/IASH, Belgium. Volume 2.

Ward, R. (1978):_Flood, A geographical perspective._ McMillan Press Ltd., London.

Weatherburn, C.E. (1962): _A first course in Mathematical Statistics._ University Press Cambridge.

Wallstead, P.E. (1979): _Introduction to physical system modelling._ Academic Press, Inc. London.

Wheelwright, S., and Makridakis, S.(1977): Ac examination of the use of adaptive filtering in forecasting. _Operational research quarterly_ vol 24, No. 1, pp. 55 - 64.

Whittle, P.(1954a): On stationary processes in plane. Biometrika 41: 434 - 49.

Whittle, P.(1954b): The statistical analysis of a seiche record. Journal Marine Research 13: 76-100.

Widrow, B.(1966): Adaptive filters I; Fundamentals. System theory laboratory, Stanford University Technical Report No. 67. pp. 64 - 66

Wiesner, C.J. (1970): Hydrometeorology. Chapman and Hall Ltd. London.

Wilde, D.J., and Beighter, C.(1964): Foundations of optimization. Eaglewood cliffs, N.J. Prentice - Hall. pp. 271 - 332.

Wilson, C.B. (1776): Effects of spacially - distributed rainfall on storm run off from a small catchment M.S. Thesis, M.I.T. Dept. of Civil Engineering.

World Meteorological Organisation (1970): Guide to hydro-meteorological practices, 2nd edition World Meteorological Organisation No. 168, Geneva.

Wold, H. (1938): The analysis of stationary time series. Almquist and Wicksell, Uppsala, Sweden.

Wood, E.F., and Szollosi- Nagy, A.(1976): Real time forecasting control of water resources systems. Selected papers from an International Institute of scientific Hydrology

Yule, G.U. (1927): On the method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot number. Philos. Trans. R.Soc. London A226; 267.

APPENDIX

1. Derivation of a linear model for β_t

Let us assume that there is some disturbance in the systems which is causing the instability and that this force is some external force which is proportional to the displacement; in addition to this are some decay and restoring forces, caused by sediment transport,floating debris and wave surges, such that the general equation and law of motion is represented by:

$$\frac{d^2\beta_t}{dt^2} = a\,\beta_t + b_1 e^{\lambda_1 t} + b_2 e^{\lambda_2 t} + b_{40} e^{\lambda_{40} t} \quad \ldots\ldots\ldots (69)$$

Neglecting terms in $e^{\lambda_i t}$, $i = 1, 40$

we have

$$\frac{d^2\beta_t}{dt^2} = a\,\beta_t \quad \ldots\ldots\ldots (70)$$

where,

a = constant

Rearrangement,

$$\frac{1}{\beta_t}\frac{d^2\beta_t}{dt^2} = a \quad \ldots\ldots\ldots (71)$$

Integrating twice yields

$$\beta_t(\ln|\beta_t| - 1) + c_1\beta_t = \frac{at^2}{2} + c_2 t + c_3 \quad \ldots\ldots (72)$$

where,

C's = constants

hence,

$$\beta_t \ln|\beta_t| - \beta_t + c_1\beta_t = a\underline{t^2} + c_2t + c_3 \quad \ldots\ldots (73)$$

but

$$-\beta_t + c_1\beta_t = (c_1 - 1)\beta_t \quad \ldots\ldots\ldots\ldots\ldots\ldots (74)$$

and

$$a\underline{t^2} + c_2t = fct(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots (75)$$

?

where,

fct = function of

Neglecting all functions of t we are left with

$$\beta_t \ln|\beta_t| + (c_1 - 1)\beta_t = c_3 \quad \ldots\ldots\ldots\ldots (76)$$

or

$$\beta_t \ln|\beta_t| = -(c_1 - 1)\beta_t + c_3$$

Taking the natural logarithm of both sides,

$$\ln(|\beta_t|\ln|\beta_t| = -(c_1 - 1)\ln|\beta_t| + \ln c_3 \quad \ldots(77)$$

expanding,

$$\ln(|\beta_t| + \ln\ln|\beta_t| = -c_1\ln|\beta_t| + \ln|\beta_t| + \ln c_3 \quad \ldots\ldots\ldots\ldots (78)$$

Simplifying, and taking

$$-c_1\ln|\beta_t| = b\ln|\beta_{t-1}| \quad \ldots\ldots\ldots (79)$$

and $\ln \quad c_z \quad = \varepsilon$ .....(80)

Then we would have

$$\ln \ln|\beta_t| = c + b \ln|\beta_{t-1}| \quad \ldots\ldots\ldots (81)$$

However, in order to induce simultaneously normality, constancy of variance and linearity in the model, β_t was transformed using the Box and Cox (1964) transformation which would first of all induce above properties in β_t; the final result obtained was

$$\ln \ln|\beta_t|^{0.1} = C_p + b\ln|\beta_{t-1}|^{0.1} \quad \ldots\ldots (82)$$

* Note that, $\ln|\beta_t|$ is used throughout instead of $\ln \beta_t$ this is because β_t is negative and $\ln \beta_t$ is not possible mathematically.

2. REGRESSION MODEL FOR THE ERROR TERM IN β_t

The regression model was developed based on the consideration that the error at time t is dependent on the error in the previous time step, t-1, and some other error term. This statement is represented by the expression

$$e_t = \beta'_t - \phi_1 e_{t-1} \quad \text{..... (03)}$$

where,

e_t = error in time step, t

e_{t-1} = error in time step, t-1

β_t = error in the model

ϕ_1 = Multiplying factor to e_{t-1}

ϕ_1 minimizes the error in the model called system response. Rearranging above expression, equation (03), we get

$$\beta'_t = e_t + \phi_1 e_{t-1} \quad \text{..... (04)}$$

The value of ϕ_1 which gave the minimum mean square error (MSE) in the model was found to be $\phi_1 = 0.8$, and the mean of the error was found to be - 0.42595. This was considered too high in view of the sensitive nature of β_t in the model. It was therefore decided to reduce the error in the model to about 5 percent or less of - 0.42595.

To do this some data were generated using the relation

$$\beta'_t = e_t - 0.8e_{t-1} \quad \text{..... (05)}$$

from equation (84) above.

These new data were modelled using equation (93) above, such that

$$\hat{e}''_t = \hat{e}^1_t + \phi_2 \quad \hat{e}^1_{t-1} \qquad \text{..... (86)}$$

The system response was found to be - 0.4.

To express $\hat{e}''_t$ in terms of the initial error terms, we recall equation (84),

$$\hat{e}^1_t = e_t + \phi_1 \, e_{t-1}$$

Similarly,

$$\hat{e}^1_{t-1} = e_{t-1} + \phi_1 \quad e_{t-2} \qquad \text{.....(87)}$$

Substituting equations (84) and (87) for $\hat{e}^1$ and $\hat{e}^1_{t-1}$ in equation (87), we have

$$\hat{e}''_t = e_t + \phi_1 e_{t-1} + \phi_2 (e_{t-1} + \phi_1 e_{t-2})$$

hence,

$$\hat{e}''_t = e_t + (\phi_1 + \phi_2)e_{t-1} + \phi_1 \quad \phi_2 \quad e_{t-2} \qquad \text{.....(88)}$$

The mean error due to $\hat{e}''_t$ was found to be - 0.06919 (that is about 15 percent of - 0.42595).

By a similar argument as before, a new set of data were generated using $\hat{e}''_t$, such that

$$\hat{e}'''_t = \hat{e}''_t + \phi_3 \quad \hat{e}''_{t-1} \qquad \text{..... (89)}$$

where ϕ_3 was found to be + 0.35. The mean error in the model was found to be + 0.03034 (that is about 7 percent of - 0.42595).

Continuing we had for

$$\hat{e}''''_t = \hat{e}'''_t + \phi_4 \hat{e}'''_{t-1} \qquad \text{..... (70)}$$

where ϕ_4 was found to be - 0.225. The mean error was found to go up to + 0.048796 (that is about 11.5 percent of - 0.42595) meaning that the model was not yet stable and that a further reduction in error was required.

This was done using the expression

$$\hat{e}''''''_t = \hat{e}''''_t + \phi_5 \hat{e}''''_{t-1} \qquad \text{..... (71)}$$

where ϕ_5 was found to be - 0.100. The mean error was found to be - 0.02181 (that is about 5 percent of - 0.42595). This was considered reasonable and using the expression for $\hat{e}''''''_t$ with respect to the original error data obtained recursively, we have

$$\begin{aligned}
\hat{e}''''''_t = {} & e_t + (\phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5)e_{t-1} \\
& + (\phi_1\phi_2 + \phi_1\phi_3 + \phi_1\phi_4 + \phi_2\phi_3 + \phi_2\phi_4 \\
& + \phi_3\phi_4 + \phi_1\phi_5 + \phi_2\phi_5 + \phi_3\phi_5 + \phi_4\phi_5)e_{t-2} \\
& + (\phi_1\phi_2\phi_3 + \phi_1\phi_2\phi_4 + \phi_1\phi_3\phi_4 + \phi_2\phi_3\phi_4 \\
& + \phi_1\phi_2\phi_5 + \phi_1\phi_3\phi_5 + \phi_2\phi_3\phi_5 + \phi_1\phi_4\phi_5 \\
& + \phi_2\phi_4\phi_5 + \phi_3\phi_4\phi_5)e_{t-3} \\
& + (\phi_1\phi_2\phi_3\phi_5 + \phi_1\phi_2\phi_4\phi_5 + \phi_1\phi_3\phi_4\phi_5 \\
& + \phi_2\phi_3\phi_4\phi_5)e_{t-4} \\
& + \phi_1\phi_2\phi_3\phi_4\phi_5 e_{t-5} \qquad \text{..... (72)}
\end{aligned}$$

Substituting for

$$\phi_1 = -0.800$$

$$\phi_2 = -0.400$$

$$\phi_3 = +0.350$$

$$\phi_4 = -0.225$$

$$\phi_5 = -0.100$$

$$\varepsilon''''_t = e_t - 1.175e_{t-1} + 0.19875e_{t-2} + 0.125375e_{t-3} - 0.01345e_{t-4} + 0.00252e_{t-5} \quad \ldots\ldots\ldots\ldots\ldots (93)$$

replacing ε''''_t by simply ε_t, and rearranging we have that

$$e_t = \varepsilon_t + 1.175e_{t-1} - 0.19875e_{t-2} - 0.125375e_{t-3} + 0.01345e_{t-4} - 0.00252e_{t-5} \quad \ldots\ldots\ldots\ldots\ldots (94)$$

The error ε_t was now modelled using Fourier transform as discussed in appendix 3.

3. FOURIER SERIES MODEL FOR ERROR θ_t AND THE DERIVATION OF THE MONTHLY ADJUSTMENT CONSTANTS, K.

A numerical method was adopted because of the complex nature of the error pattern, that it could not be defined easily by any known formula. This numerical method is less accurate especially at higher harmonics, therefore only the first (K/2 -1) harmonics were included in the series. K is the number of intervals, in this case, it is equal to twelve in order to conform with the cyclic nature of the system.

Hence there are only (12/2 - 1) =5 harmonics considered, such that the Fourier series was given as

$$\theta_t = -0.01352 + 0.28376\cos\theta - 0.11932\cos 2\theta - 0.0952\cos 3\theta - 0.00278\cos 4\theta - 0.00126\cos 5\theta + 0.10767\sin\theta - 0.06681\sin 2\theta + 0.00833\sin 3\theta + 0.17642\sin 4\theta + 0.06182\sin 5\theta \quad \ldots\ldots\ldots(95)$$

In which θ are given in degrees, with the whole cycle of 360° divided into 12 equal part of 30°, representing the seasonal length and the unit of 30° representing months.

Starting from 30°, the month of January was defined, 60° February, 90° March, and so on to 360°.

function, summing them up and dividing by 4, for example

$$0.28376 = \frac{\Sigma m \cos \theta}{4} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (94)$$

Substituting for the monthly units of 30°, 60° and so on in equation (95), the monthly adjustment constants, E, were obtained. Table 5.