RAINGAUGE NETWORK DESIGNS FOR UGANDA

BY

CHARLES P.K. BASALIRWA

A thesis submitted in fulfilment for the Degree of Doctor of Philosophy in Meteorology at the University of Nairobi.

August, 1991

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University.

Signed: ______________________ 30/11/90

CHARLES P.K. BASALIRWA

This thesis has been submitted for examination with our approval as University Supervisors.

Signed: ______________________ 2/12/90

PROF.L.J. OGALLO

Signed: ______________________

DR. F.M. MUTUA

Department of Meteorology
University of Nairobi
P.O. Box 30197
NAIROBI

TABLE OF CONTENTS

CONTENTS		PAGE
Title		(i)
Declaration		(ii)
Table of Contents		(iii)
Abstract		(ix)
List of Figures and their Captions		(xii)
List of Tables and their Captions		(xvii
CHAPTER 1		
1.0	Introduction	1
1.1	Literature Review	6
1.2	Objective of Study	8
1.3	Data Sources	9
1.4	The Climatology of Uganda	17
CHAPTER 2 METHODOLOGY		
2.0	Introduction	25
2.1	Estimation of Missing Records	25
2.1.1	The Isopleth Method	25
2.1.2	Correlation Methods of Estimation	26
2.2	Areal Averages	27
2.2.1	Arithmetic Mean	28
2.2.2	The Isohyetal Method	28
2.2.3	The Isopercentile and Thiessen Polygon Methods	29
2.2.4	Empirical Methods	31
2.2.5	The Use of Empirical Orthogonal Functions	33
2.2.6	Remote Sensing Methods	34

CONTENTS PAGE

2.2.6.1	The Use of Radar	34
2.2.6.2	Satellite Methods	36
2.2.6.2.1	Cloud Statistics	36
2.2.6.2.2	Vegetation Indices	38
2.2.6.2.3	Landsat Data	39
2.3	Data Quality Control	41
2.3.1	The Use of Mass and Double Mass Curves	42
2.3.2	Data Quality Control Using Correlation Methods	43
2.3.3	Trend Analyses	45
2.3.4	Maronna and Yohai Method	47
2.3.5	Wald-Wolfwitz One Sample Runs Test	48
2.3.6	Other Homogeneity Tests	49
2.4	Raingauge Network Design	50
2.4.1	Minimum Raingauge Network Design	51
2.4.1.1	Spatial Delineation of Climatic Zones	52
2.4.1.2	Cluster Analysis	53
2.4.1.3	Map Typing Method	55
2.4.1.4	Empirical Orthogonal Functions	57
2.4.1.4.1	Factor and Principal Component Analyses	58
2.4.1.4.2	Factor Analysis	59
2.4.1.4.3	Principal Component Analysis (PCA)	62
2.4.1.4.4	Number of Significant Principal Components (Eigenvectors)...............	64

CONTENTS PAGE

2.4.1.4.5	Kaiser Criterion	65
2.4.1.4.6	The Scree Method	65
2.4.1.4.7	The logarithm of the Eigenvalue (LEV) Method	66
2.4.1.4.8	Use of Sampling Errors of the Eigenvalues	66
2.4.1.4.9	Physical Reality (Stability) of the PCA patterns	68
2.4.1.4.10	Rotation of Principal Components	69
2.4.1.4.11	The Varimax Method	69
2.4.1.4.12	The Oblique Method	71
2.4.1.4.13	Other Varification Methods	72
2.4.1.4.14	The Use of Correlation Maps	72
2.4.1.4.15	Vector Plots and other Methods	73
2.4.1.5	The Ability of the Minimum Network in the Representation of the Regional Rainfall Characteristics	74
2.4.2	Optimum Raingauge Network Design	76
2.4.2.1	The Use of Principal Component Analysis (PCA) in Optimum Raingauge Network Design	77
2.4.2.2	Large Raingauge Density Methods	79
2.4.2.3	The Use of Running Costs of a Network	80
2.4.2.4	Physiographic Effects	81
2.4.2.5	Coefficient of variation Methods	82
2.4.2.6	Optimum Error of Interpolation	87
2.4.2.6.1	Graphical Methods for Determining The Correlation Distance Relationship	92

CONTENTS PAGE

2.4.2.6.2 The Linear Regression Model 92

2.4.2.6.3 Exponential Regression Model 93

CHAPTER 3

RESULTS AND DISCUSSION

3.0 Introduction 95

3.1 Quality of the Estimated Rainfall Records 95

3.2 The Results of Homogeneity Tests and the Adjustment of Heterogeneous Records 101

3.2.1 Double Mass Curves Results 102

3.2.2 Wald-Wolfowitz One Sample Runs Test Results 102

3.3 Areal Rainfall Estimates 108

3.4 Results From Raingauge Network Design 111

3.4.1 Results From Minimum Raingauge Net-Work Design 111

3.4.1.1 Results From Principal Component Analysis (PCA) 112

3.4.1.2 Significant Factors 112

3.4.1.3 The Delineated Homogeneous Zones of Uganda 118

3.4.1.4 Stability Results of the Delineated Zones 135

3.4.1.5 The Regional PCA Results 136

CONTENTS PAGE

3.4.1.6 The Use of Interstation Correlations 138

3.4.1.7 Results of Vector Space Mapping 141

3.4.1.8 Physical Factors Associated with the Homogeneous Regions 141

3.4.1.9 The Mean Seasonal Rainfall Characteristics Over the Individual Homogeneous Regions 151

3.4.1.9.1 Region A1 151

3.4.1.9.2 Region A2 153

3.4.1.9.3 Region B 154

3.4.1.9.4 Region C 155

3.4.1.9.5 Region D 157

3.4.1.9.6 Region E 157

3.4.1.9.7 Region F 158

3.4.1.9.8 Region G 158

3.4.1.9.9 Region H 160

3.4.1.9.10 Region I 161

3.4.1.9.11 Region J 162

3.4.1.9.12 Region K 162

3.4.1.9.13 Region L 164

3.4.1.9.14 Region M 165

3.4.1.10 Minimum Raingauge Network Design Derived From PCA 165

3.4.2 Optimum Raingauge Network Design Results 172

CONTENTS PAGE

3.4.2.1 PCA Regional Network Optimization Results 173

3.4.2.2 Optimum Raingauge Network Design Derived From Physiographic Effects Method 190

3.4.2.3 Optimum Network Designs Derived From Coefficient of Variation 194

3.4.2.4 Optimum Raingauge Network Design Derived From Error of Interpolation Method 199

3.4.2.4.1 Results From Graphical Methods 200

3.4.2.4.2 Regression Methods Results 203

CHAPTER 4

4.0 Summary and Conclusion 214

4.1 Summary of Results 214

4.2 Conclusions 216

4.3 Suggested Future Work 217

Acknowledgements 219

References 220

ABSTRACT

The effective use of rainfall informations for all water use activities can only be derived from well designed raingauge networks. The objective of this study was to design raingauge networks for Uganda. Two basic concepts were used, namely: the 'minimum, and 'optimum' network principles. Under the minimum network design, the minimum number of raingauges that must be operated in order to represent all the climatological divisions are determined. The optimum network design on the other hand, constitutes the least number of raingauges which must be operated in order to obtain optimum rainfall informations for the various water use activities.

Monthly rainfall records at 102 stations spanning within the years 1940-1985 were used in the study. Data quality control tests were used to test the quality of the few missing estimated records and all rainfall records before they were used in the study.

Principal component analysis (PCA) was used first to delineate Uganda into homogeneous rainfall subdivisions based on the spatial patterns of the dominant PCA modes. Areal rainfall values are usually better indicators of the synoptic systems controlling rainfall than point values at the individual locations. The areal rainfall estimates derived from the simple arithmetic mean and those obtained from PCA solutions were used to examine the skill of the developed network designs in the representation of the spatial and temporal rainfall characteristics at the PCA derived climatological divisions.

The principle of the minimum network design ensured that each of the delineated homogeneous rainfall subdivisions was represented by at least one raingauge. The methods which were used in the optimum network design included the PCA method of optimization, physiographic effects, coefficients of variations and optimum error of interpolation techniques.

Finally, the skill of the designed networks in effectively representing the rainfall characteristics over the individual rainfall subdivisions of Uganda were examined using analysis of variance techniques.

Results from data quality control tests indicated that all estimated and other rainfall records were homogeneous, indicating no significant shifts in the spatial and temporal rainfall characteristics over Uganda.

The results from PCA delineated Uganda into fourteen (14) homogeneous rainfall subdivisions. The minimum raingauge network design for Uganda would therefore consist of fourteen (14) stations with one station at each of the homogeneous rainfall subdivisions. Only two of the current synoptic stations in Uganda were included in the designed minimum network. Thus, the current synoptic network stations of Uganda do not give a good representation of the climatology of the country.

Results from the study indicated that the variances of the areal rainfall explained by the minimum network design varied from region to region and season to season. The lowest variances were explained during season 2 (March to May) and season 3 (June to August). This was attributed to the high temporal rainfall variability during these seasons. It was also noted that many large areas of the country were unrepresented by the minimum network design stations.

The generated optimum raingauge network designs gave 89 stations for the PCA method, 1517 for the physiographic effects method, 2483, 374 and 1477 stations for the three coefficient of variation methods and one twenty (120) stations for the optimum error of interpolation method. Thus, the largest raingauge networks were obtained with the physiographic effects, the three coefficient of variation and optimum error of interpolation method. Most stations in these large networks were located in the dry areas, near water bodies or were needed during the dry seasons. Such large networks would be uneconomical to run since rainfall dependent economic activities are least during dry seasons and at dry locations.

It may be concluded from the study that Uganda should have at least fourteen (14) synoptic stations based on one station at each of the homogeneous climatological subdivisions. The optimum number of raingauges for Uganda were however based on those found through the PCA method of at least 89. The use of orthogonal functions in representing the spatial characteristics under PCA ensures an optimum representation of the complex spatial rainfall characteristics over Uganda. This method of network optimization is being recommended and can be extended to any region of the world due to the flexibility in the characteristics of the empirical orthogonal functions.

LIST OF FIGURES

FIGURE		PAGE
1:	Number of stations in each half degree square with at least five years record in East Africa (Tomsett, 1969):	4
2:	The stations used in the study	11
3:	Relief map of Uganda	18
4:	Examples of the seasonal distribution of Rainfall in Uganda	23
5:	Examples of the estimated missing records from the Isopleth method	97
6:	Examples of the estimated rainfall records for some selected periods at station 153 using the isopleth method	100
7a:	Season 1 double mass curve for stations 71 and 73	103
7b:	Season 2 double mass curve for stations 71 and 73	103
7c:	Season 3 double mass curve for stations 105 and 107	104
7d:	Season 4 double mass curve for stations 105 and 107	104
7e:	Season 2 double mass curve for stations 137 and 73	105
7f:	Season 4 double mass curve for stations 137 and 73	105

FIGURE PAGE

7g: Season 1 adjusted cumulated rainfall data at station 71 versus that of station 73 106

7h: Season 2 adjusted cumulated rainfall data of station 71 versus that of station 73 106

8a: Example of the results from the scree test 115

8b: Example of the results from LEV test...... 115

9: The delineated climatological zones of Uganda from PCA results................ 119

10a: Map patterns of factor 1 loadings for climatological records................. 122

10b: Map patterns of factor 2 loadings for climatological records..................... 123

10c: Map patterns of factor 3 loadings for climatological records..................... 124

10d: Map patterns of factor 4 loadings for climatological records 125

11a: Map patterns for factor 1 loading during season 2................................ 126

11b: Map patterns for factor 2 loadings during season 2................................ 126

11c: Map patterns for factor 3 loadings during season 2......................... 127

FIGURE PAGE

11d: Map patterns for factor 4 loadings during season 2............................ 127

11e: Map patterns for factor 9 loadings during season 2 128

11f: Map patterns for factor 11 loadings during season 2 128

11g: Map patterns of factor 2 loadings during season 3 129

11h: Map patterns of factor 3 loadings during season 3 129

12a: The spatial patterns of interstation correlations with station 112............. 139

12b: The spatial patterns of interstation correlations with station 54 139

13: Example of a space vector plot for monthly rainfall climatological data...................................... 140

14: Mean annual rainfall map........................ 143

15a: Season 1 (December-February) mean rainfall map 147

15b: Season 2 (March-May) mean rainfall map.......... 147

15c: Season 3 (June-August) mean rainfall map 149

15d: Season 4 (September-November) mean rainfall map 149

16: Seasonal rainfall distribution at Kyanamukaka(station 119) in region Al. 152

FIGURE PAGE

17: Seasonal rainfall distribution at Buvuma (station 87) in region A2.................... 152

18: Seasonal rainfall distribution at Bumangi (station 129) in region A1................... 152

19: Seasonal rainfall distribution at Mutai Forest(station 93) in region B................ 152

20: Seasonal rainfall distribution at Mbarara (station 107) in region C................ 156

21: Seasonal rainfall distribution at Dabani (station 100) in region D........................ 156

22: Seasonal rainfall distribution at Serere (station 165) in region E........................ 156

23: Seasonal rainfall distribution at Mbale (station 5) in region F....................... 156

24: Seasonal rainfall distribution at Kotido (station 52) in region G..................... 159

25: Seasonal rainfall distribution at Kitgum (station 43) in region H..................... 159

26: Seasonal rainfall distribution at Gulu (station 54) in region I...................... 159

27: Seasonal rainfall distribution at Moyo (station 39) in region J................. 159

28: Seasonal rainfall distribution at Nebbi (station 53) in region K 163

FIGURE PAGE

29: Seasonal rainfall distribution at Kijura (station 19) in region L.................... 163

30: Seasonal rainfall distribution at Mubende (station 22) in region M 163

31: Map of the minimum raingauge network design generated from the study..................... 170

32: PCA derived optimum raingauge network design.... 189

33a: Examples of the relationship between interstation correlation and distance during season 1. 201

33b: Examples of the relationship between interstation correlation and distance during season 2 201

34a: Averaged relationship between interstation correlations and distances during season 1 202

34b: Averaged relationship between interstation correlations and distances during season 2 202

35: The standard deviations of season 2 209

36: Existing percentage errors interpolation for season 2 210

37: Intergauge distances based on γ = 25% during season 2. 212

LIST OF TABLES

Table		Page
1:	Details of the Stations used in the study......	12
2:	Percentage of Land Area of Uganda Receiving some selected values of Annual Rainfall Amounts	21
3A:	Examples of the Rainfall Estimates for Station 153 with Isopleth and Correlation Methods.......	96
3B:	Observed and Estimated Rainfall at Station No. 153	99
4:	The Chi-Square Values between Arithmetic Mean values and PCA Factor Weighting averages for various homogeneous regions	110
5:	Example of PCA Results with the Mean Climatological Records...........................	113
6:	Example of significance of sampling Errors of the Eigenvalues for the mean climatological Rainfall records	113
7:	Maximum number of significant factors with monthly and seasonal rainfall records..	116
8:	Loadings of the four major Factors for Regions C, H and J based on climatological Records	121

Table		Page
9:	Summary from Analysis of the Dominant PCA modes from Climatological and Seasonal Rainfall Records for Regions given in Figure 9.	133
10A:	The PCA results of a homogeneous group of stations (Region A1 with 9 stations)	136
10B:	The PCA results of a homogeneous group of stations (Region B with 19 stations)	137
11:	The Mean Annual Rainfall for the Wettest and driest stations in the delimited homogeneous regions	145
12A:	The correlation coefficients (r) and their Squares (r^2) between areal rainfall and representative station rainfall totals in the various homogeneous regions	168
12B:	Regression parameters for the linear regression model (equations 33)	169
13A:	The regression parameters for the PCA optimum network design (model equation 34) for season 1	174
13B:	The regression parameters for the PCA optimum network design (model equation 34) for season 2	175
13C:	The regression parameters for the PCA optimum network design (model equation 34) for season 3	176
13D:	The regression parameters for the PCA optimum network design (Model equation 34) for season 4.......	177
13E:	The regression parameters for the PCA optimum network design (model equation 34) for the mean annual values	178

Table Page

13F: Regional Areal Variances ($x10^2$) explained at each regression step for season 1 179

13G: Regional Areal Variances ($x10^2$) explained at each regression step for season 2 180

13H: Regional Areal Variances ($x10^2$) explained at each regression step for season 3 181

13I: Regional Areal Variances ($x10^2$) explained at each regression step for season 4 182

13J: Regional Areal Variances ($x10^2$) explained at each regression step for the mean annual values ... 183

14: The desired number of gauges (N) derived from the use of physiographic effects method (equations 36) 191

15A: The results from the use of coefficient of variation (model equation 37) 195

15B: The results from the use of coefficient of variation (model equation 38, P = 10%) 197

15C: The results from the use of coefficient of variation (model equation 39, D = 10%) 198

16: Parameter estimates for the linear and exponential regression models 204

17: Sample Estimates of εopt values at various distances (x) 206

18: Sample Estimates of the optimum probability weighting (p) at various distances (x) 207

CHAPTER I

1.0 INTRODUCTION

In the tropics, rainfall is the climatic element with the highest variability. The high spatial and temporal variability of rainfall is reflected in all rainfall dependent activities. Knowledge of the temporal and spatial characteristics of rainfall is, therefore, not only important in meteorology, but is also crucial in the planning and management of all rainfall dependent activities. These activities include, among many others:-

(i) Agriculture, which is the backbone of Uganda's economy. The agricultural activities which can be optimized with good rainfall informations are the decisions on suitability of crops, the introduction of new ones, planting dates, growth period, time of harvest, forecasting methods and many others.

(ii) Hydrological systems such as in irrigation schemes, industrial, domestic urban and rural water needs, wildlife management and biomass monitoring, hydro-electric power production and many others.

(iii) The transport industry, communication, public health safeguards etc.

(iv) Delineation of risk zones for any water use activities.

(v) Real time assessment and forecasting of available food and water resources together with expected anomalies and their corresponding impacts.

Examples of the dangers inherrent in ignoring the spatial and temporal rainfall variability at the planning stages of the projects include the 1961 Uganda Marine disaster (Daily Nation, 1988). In 1961, lake Victoria levels rose by about 10m. The rising lake levels submerged all port facilities throughout the country. This caused a huge economic loss to the Railways and Habours Corporation. It was evident from these impacts that extreme rainfall variations over the lake region were not considered at the planning stages of these ports.

Another example was in 1984 when the so called Kitgum Farming Company was formed with a major objective of growing coffee in the Kitgum area. The company operated on the pseudo assumption that the whole of Uganda has similar climatic characteristics. Coffee farming practices which were used in the best coffee growing areas of Southern and Eastern parts of Uganda were, therefore, extended to the Northern parts (Kitgum area) which have comparable total annual rainfall amounts but different temporal rainfall characteristics. The non consideration of the temporal rainfall variability at the planning stage resulted in a complete failure of the project and a huge ecomic loss to the investors.

In the early 1950's, investors ignorant of the spatial and temporal rainfall variability in Central Tanzania started a groundnut scheme in Dodoma. The scheme failed and caused investors a huge loss of capital.

These three examples, illustrate the need for optimum use of the knowledge of the spatial and temporal variability in rainfall in the planning of all rain dependent activities.

One good example in which the climatic characteristics were considered is the successful Kibimba rice scheme in Eastern Uganda. The temporal rainfall characteristics were obviated in the planning stages and alternative projects like poultry farming and pig husbandary were introduced during anomalous rainfall periods.

Optimum planning and management of all rainfall dependent activities, therefore, require good knowledge of the spatial and temporal characteristics of rainfall informations. These informations are often derived from the surface raingauge network.

Figure 1 shows the raingauge densities in East Africa in 1969 (Tomsett, 1969). It is clear from the figure that the raingauge networks in East Africa like in many other countries, have developed from ad hoc responses to particular problems (Rodda, 1969, Langbein, 1972). Most of the raingauges are generally

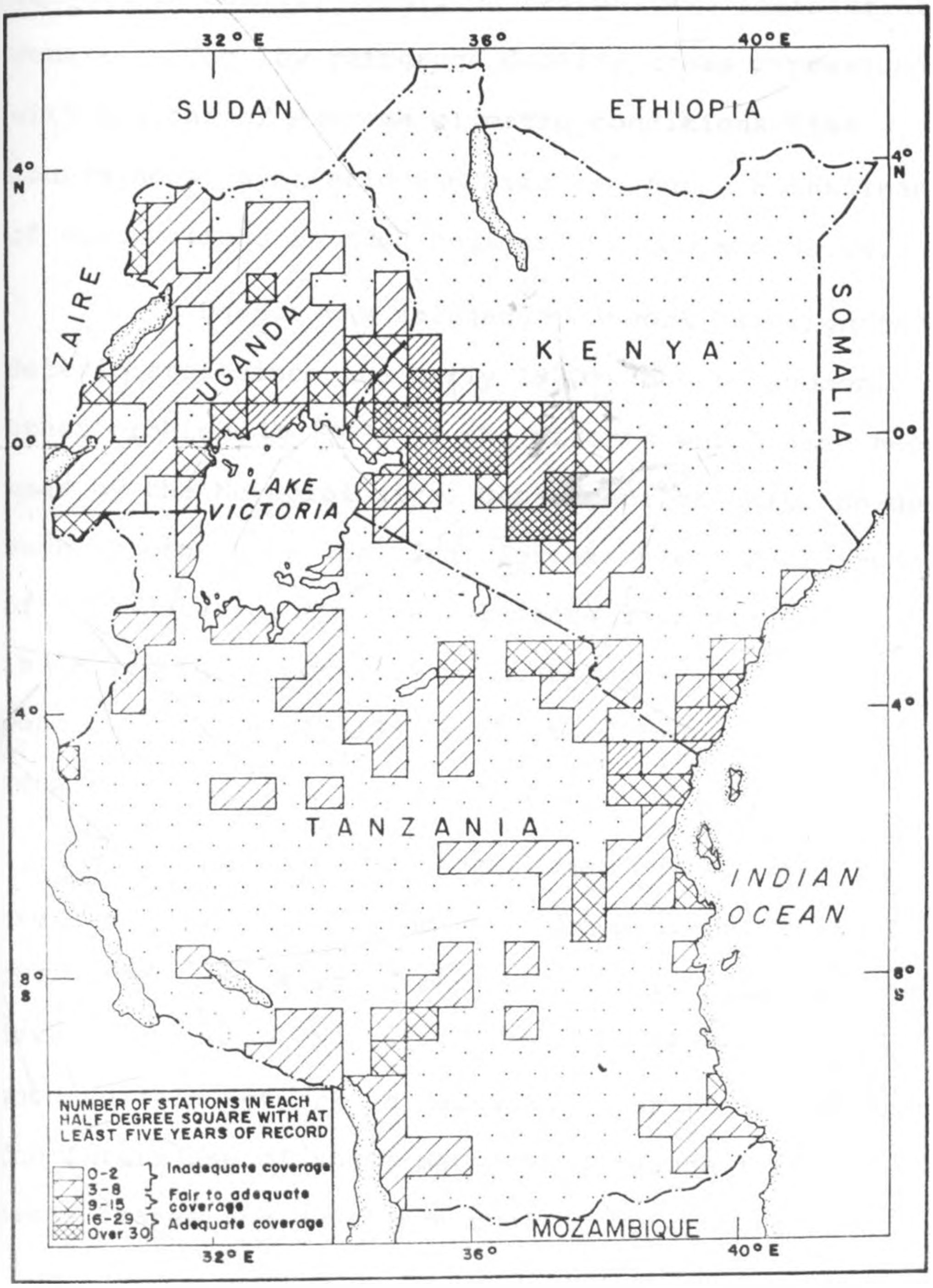


Fig. 1 Number of stations in each half degree square with at least five years record in East Africa (Tomsett, 1969)

located in heavily populated regions and administrative centres. The low raingauge density areas correspond with regions of extreme climatic conditions like mountainous, semi-arid and arid regions. Maintainance of observations in such regions is too expensive.

In Uganda the raingauge network has significantly deteriorated since the early 1970's due to economic and other problems in spite of the efforts which have been made by the Meteorological Department to halt the decline. Such efforts have included, for example, the introduction of an honorarium for the voluntary rainfall observers in 1977. The honorarium payments were, however, suspended in the early 1980's and this led to a further decline in the raingauge network.

Although remote sensing methods are now commonly used in many parts of the world to monitor rainfall characteritics, most of the methods are still at the development stage and all of these methods require a good surface network of raingauges for their calibrations. The future use of these methods in Uganda further necessitates the development of an optimum network.

The major objective of the study is to develop a minimum and an optimum raingauge networks for Uganda which can provide rainfall informations required by a variety of users.

A minimum raingauges network design ensures that each spatial rainfall pattern (climatological division) is represented by at least one raingauge. A minimum density raingauge network has, therefore, the advantages of ensuring the availability of sample rainfall data representative of all the climatological divisions and at minimum cost.

The impacts of the regional and local factors like topography, water bodies, land use activities and others induce large spatial and temporal variations in rainfall even within the individual homogeneous climatological zones. These factors can introduce large local and regional variations in the requirements of the rainfall informations which cannot be obtained from a minimum raingauge network design. This necessitated the development of an optimum raingauge network, which in spite of its higher costs can give optimum rainfall information to the various users.

The next section reviews some of the literature on areal rainfall estimation and raingauge network designs.

1.1 LITERATURE REVIEW

Rainfall is the major source of water for many water use activities. Many attempts have therefore been made to obtain reliable rainfall estimates at local, regional and national levels. A number of these attempts have sought to improve the accuracy of the areal rainfall estimates by using weighting functions of the individual raingauge catches. These studies have included those of Thiessen (1911), Chidley and Keys (1970), Akin (1971), Edwards (1972), Delfiner and Delhomme (1973), Bras and Rodriguez-Iturbe (1976) and many others. The derivation of weighting functions for a raingauge network may not, however, improve the accuracy of the areal rainfall estimate in areas with complex spatial and temporal rainfall variations.

Most of the raingauges are located at urban, administrative, scientific research institutions or agricultural farm sites where manpower for periodic rainfall observations can be obtained. A number of attempts have therefore been made to determine raingauge networks that provide reliable areal rainfall estimates. Such studies have included those of Wilm et al (1939), Ganguli et al (1951), Thom (1951) Kohler (1958), Rodda (1972), Dymond (1982), Ahuja (1959), Bleasdale (1965), Gandin (1972), Alaka (1970), Kagan (1972a, 1972b), Cislerova and Hutchnison (1974), Hutchnison (1974) and many others. These studies have ranged from the use of the variance of the mean areal rainfall to the determination of the relationship between the structural function and intergauge distance errors. Difficulties often arise in the practical application of many of these methods in the determination of the actual mean rainfall for the assessment of the variance or in attempts to model intergauge distance relationships in areas with complex space and time rainfall characteristics. Details of these methods, together with their associated difficulties will be discussed with methodology in Chapter 2.

In East Africa, no significant work has been done on the design of raingauge networks or the determination of the adequacy of existing networks. The general guidelines provided by the World Meteorological Organization (WMO) for example WMO (1962, 1970, 1974, 1981) and many other references can be used. These WMO guidelines, however, may be inadequate especially in regions with large spatial and temporal rainfall variations. For Uganda in particular, the raingauge network has declined since the early 1970's due to social and economic difficulties. The determination of the minimum and optimum raingauge work designs would provide a useful basis for decisions on priority gauge re-establishment and the development of a reliable raingauge network.

Details of the objectives of this study are provided in the next section.

1.2 OBJECTIVE OF STUDY

The primary objective of the current study is to design the minimum and optimum raingauge networks for Uganda. These networks are developed from the historical rainfall records.

Some of the problems which are inherrent with rainfall records include non-uniformity of record lengths and the missing of rainfall informations. The initial step in the study therefore involved the determination of the best period of continuous record length for the available stations. Even within the selected period, a few records were still missing at some locations. The first part of the study concentrated on the estimation of the missing records using some of the standard techniques. The second part of the study examined the quality of the estimated records. The homogeneity of the historical records were also examined before these data were subjected to any analyses.

Another problem with the use of rainfall informations is the high spatial variations of the rainfall characteristics. Uganda's rainfall is generated from complex local and well developed regional and synoptic features. Under such circumstances, areal derived rainfall statistics sometimes give more useful informations for regional planning purposes than point values. The third part of the study was devoted to comparison of point and areal rainfall values. These records formed the fundamental base for the development of the minimum and optimum network designs for Uganda as highlighted in Chapter 2.

The next two sections are devoted to the data used in the study and the climatology of Uganda.

1.3 DATA SOURCES

The rainfall data used in this study consisted of monthly rainfall totals. The bulk of the data were obtained from the Uganda Meteorological Department. Some data were also obtained from the Kenya Meteorological Department which has an archive of the old records collected before the collapse of the East African Community in June, 1977.

A scan through the rainfall records indicated that 102 stations had relatively few missing records within the period 1940-1975 inclusive.

The spatial distribution of the stations used in the study are given in Figure 2, while Table 1 gives their names, locations and altitudes. In the next section a brief review of the climatology of Uganda, the study area, is given.

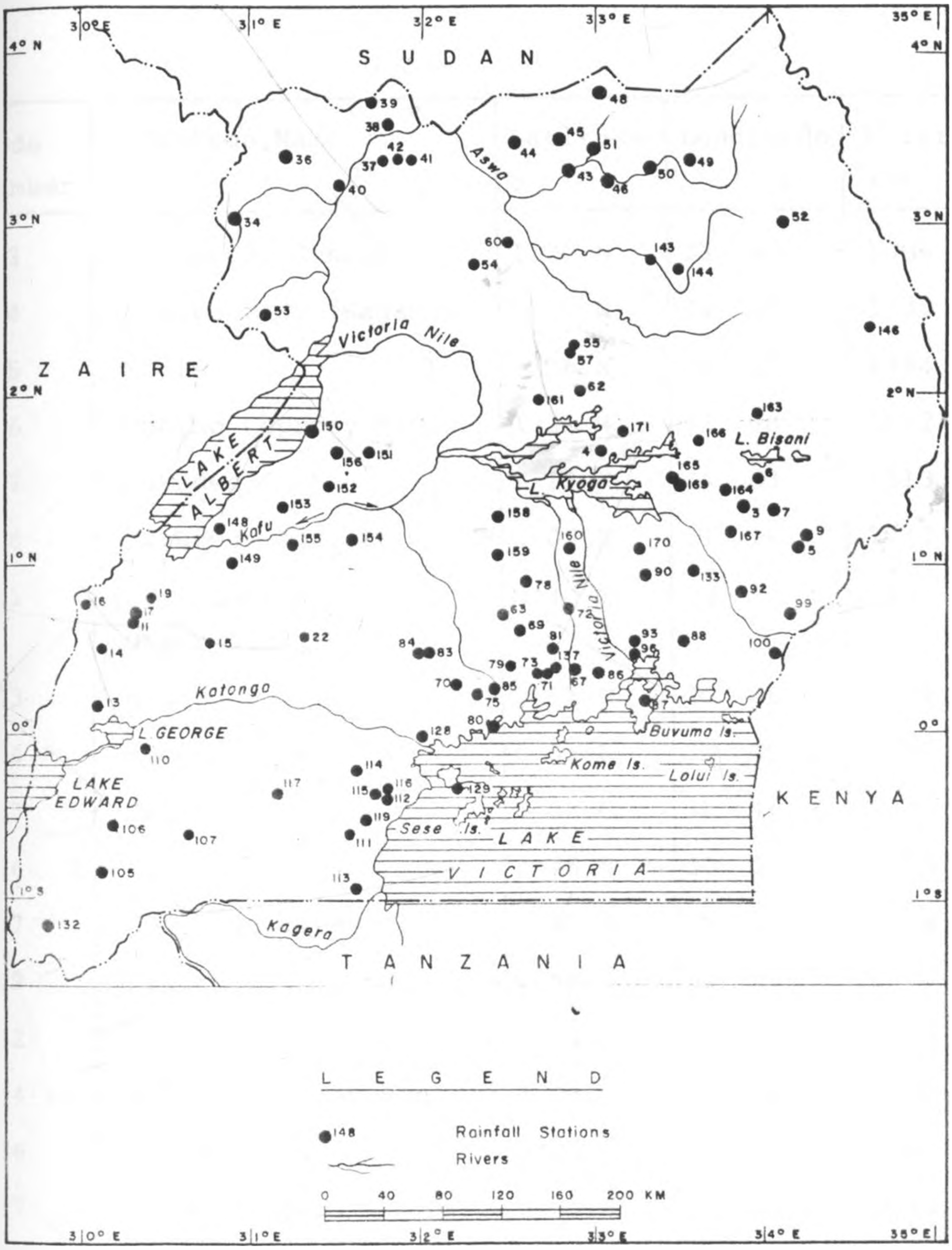


Fig. 2 The stations used in the study

Table 1: Details of the stations used in the study.

Code Number	Station Name	Latitude o '	Longitude o ' E	Altitude (m)
3	Mukongoro Ginnery	1 20 N	33 53	1006
4	Ochero P.S. (Kagaya)	1 40 N	33 01	1036
5	Mbale	1 06 N	34 11	1494
6	Ongino Leprosy Camp	1 31 N	33 58	1052
7	Bukedes	1 19 N	34 03	1113
8	Bugusege Agr.Stn.	1 09 N	34 16	1433
11	Fort Portal 2nd Order Stn.	0 40 N	30 17	1539
13	Bugoye	0 17 N	30 06	1829
14	Kisomoro	0 31 N	30 09	1585
15	Matiri	0 32 N	30 46	1341
16	Nyaruru	0 45 N	30 02	914
17	Kyembogo Farm	0 41 N	30 20	1524
19	Kijura Tea Factory	0 48 N	30 25	1524
22	Mubende 2nd Order Stn.	0 35 N	31 22	1553
34	Arua 2nd Order Stn.	3 03 N	30 56	1204
36	Yumbe Disp.	3 25N	31 14	1067
37	Ajumani Disp.	3 22 N	31 47	732
38	Laropi Disp.	3 34 N	31 50	610
39	Moyo Boma	3 41 N	31 44	1036
40	Obongi Disp.	3 14 N	31 33	610
41	Zaipi Disp.	3 22 N	31 58	732
42	Pakelle N.T.C.	3 22 N	31 52	762

Code No.	Station Name	Latitude o '	Longitude o ' E	Altitude (m)
43	Kitgum Cotton V.T.C.	3 17 N	32 53	937
44	Palabek Gomb	3 27 N	32 34	914
45	Padibe	3 29 N	32 49	914
46	Kitgum Matindi Gomb	3 17 N	33 03	914
48	Madi Opei Admin.	3 46 N	33 01	1372
49	Rom Gomb.	3 23 N	33 33	1219
50	Naam Okora	3 19N	33 19	1219
51	Mucwini Gomb.	3 26 N	33 01	914
52	Kotido	3 01 N	34 06	1219
53	Nebbi V.T.C.	2 29 N	31 05	1372
54	Gulu 2nd Order stn.	2 45 N	32 20	1106
55	Ngetta Farm-Lira	2 17 N	32 56	1095
57	Lira 2nd Order stn.	2 15 N	32 54	1085
60	Abera Forest stn.	2 53 N	32 25	1109
62	Bardyang	2 01 N	32 57	1097
63	Bukalasa Agr. stn.	0 43 N	32 31	1128
67	Moniko Estate	0 23 N	32 55	1250
69	Kalagala Agr. stn.	0 37 N	32 37	1097
70	Kabasanda	0 17 N	32 13	1158
71	Mukono Agr. stn.	0 21 N	32 45	1184
72	Ntenjeru Agr. stn.	0 44 N	32 53	1158
73	Namanve Forest stn.	0 21 N	32 41	1135
75	Mpigi Agr. stn.	0 13 N	32 20	1250
78	Wabusaana	0 53 N	32 39	1219

Code No.	Station Name	Latitude o '	Longitude o ' E	Altitude (m)
79	Kawanda Agr. stn.	O 25 N	32 32	1196
80	Entebbe Vet. Lab	O 02 N	32 27	1183
81	Nakifuma	O 32 N	32 47	1128
83	Bakijjula	O 28 N	32 03	1219
84	Senda	O 29 N	32 01	1219
85	Nsimbe Estate	O 16 N	32 26	1219
86	Bwavu Estate	O 21 N	33 01	1280
87	Buvuma Island	O 11 N	33 18	1158
88	Nawanzu	O 33 N	33 30	1189
90	Namwendwa Railway stn.	O 55 N	33 17	1105
92	Budumba Railway stn.	O 49 N	33 51	1070
93	Mutai Forest stn.	O 33 N	33 13	1146
96	Nakabango	O 29 N	33 14	1219
99	Tororo 2nd Order stn.	O 42 N	34 10	1226
100	Debani M.H.M	O 28 N	34 05	1219
105	Rwashamaire	O 50 S	30 08	1646
106	Bushenyi	O 33 S	30 12	1631
107	Mbarara 2nd Order stn.	O 37 S	30 39	1443
110	Kicheche	O 05 S	30 24	1372
111	Kiteredde Mission	O 37 S	31 34	1219
112	Buwunga	O 23 S	31 47	1250
113	Katera Sango Bay Estate	O 55 S	31 38	1189
114	Katigondo W.F.M.	O 13 S	31 44	1311
115	Masaka	O 20 S	31 44	1311

Code No.	Station Name	Latitude ° '	Longitude ° ' E	Altitude (m)
116	Kiwala Estate	0 20 S	31 48	1298
117	Lyangonde Disp.	0 20 S	31 09	1219
119	Kyanamukaka	0 30 S	31 41	1219
128	Nkozi	0 01 S	32 01	1189
129	Bumangi W.F.M.	0 19 S	32 14	1143
132	Kabale 2nd Order stn.	1 15 S	29 59	1871
133	Vukula	0 57 N	33 36	1097
137	Kivulu Estate	0 23 N	32 47	1228
143	Patonga	2 47 N	33 19	1067
144	Adilang	2 44 N	33 28	1067
146	Moroto 2nd Order	2 33 N	34 36	1524
148	Kyangwari	1 13 N	30 49	1067
149	Mugalike	1 01 N	30 53	1219
150	Butiaba 2nd Order stn.	2 33 N	34 36	621
151	Masindi 2nd order stn.	1 41 N	31 43	1146
152	Bulindi	1 28 N	31 28	1036
153	Kiziranfumbi	1 20 N	31 12	1067
154	Butemba	1 09 N	31 36	1158
155	Nalweyo	1 07 N	31 16	1219
156	Nyabyeya	1 40 N	31 32	1189
158	Nakasongola	1 19 N	32 28	1274

Code No.	Station Name	Latitude	Longitude	Altitude (m)
159	Kakoge	1 04 N	32 28	1189
160	Bale	1 06 N	32 53	1097
161	Aduku V.T.C.	1 59 N	32 43	1036
163	Katakwi Disp.	1 54 N	33 59	1158
164	Ngora C.M.S.	1 27 N	33 46	1128
165	Serere Agr. stn.	1 31 N	33 27	1139
166	Soroti 2nd Order stn.	1 43 N	33 37	1127
167	Kibale V.T.C.	1 12 N	33 47	1097
169	Kyere	1 29 N	33 37	1067
170	Bugaya	1 06 N	33 15	1067
171	Kaberamaidc Agr. stn.	1 47 N	33 10	1067

Note: The missing code numbers were excluded in the study.

V.T.C = Variety trial centre.
Agr. stn. = Agricultural station.
W.F.M. = White Fathers Mission
Disp. = Dispensary
Gomb = Gombolola (Sub-county headquarters).
N.T.C. = National Teachers College

1.4 THE CLIMATOLOGY OF UGANDA

Uganda lies within Latitudes 5°N - 1°S and Longitudes 29°E - 36°E (Figure 3). The major factors controlling the climate of Uganda include the Inter-Tropical Convergence Zone (ITCZ) a broad zonal low pressure belt in which the north east and south east trade winds of the two hemispheres meet. The ITCZ moves northwards and southwards lagging behind the overhead sun by about one month. The location of this zone in a particular region concides with the season of maximum precipitation in that area (Griffiths, 1972, 1969, Johnson and North, 1961a and 1961b, Johnson, 1962). The mean monthly position of the ITCZ in Africa has been given by many authors including Dhonneur (1974). The effectiveness and depth of the ITCZ is mainly controlled by the intensity of the subtropical anticyclones. These include the Arabian high to the north, the Azores high to the north-west of Africa, the Mascarene high to the south and the St. Helena high to the southwest. These anticyclones determine the characteritics of the monsoonal winds over East Africa.

During the months of July and August when the ITCZ is furthest north, there is a maximum influx of moist westerly winds from Congo/Zaire basin and the Atlantic Ocean (Trewartha, 1961, Thompson, 1957, Ogallo, 1988). These winds are locally known as "Congo air mass."

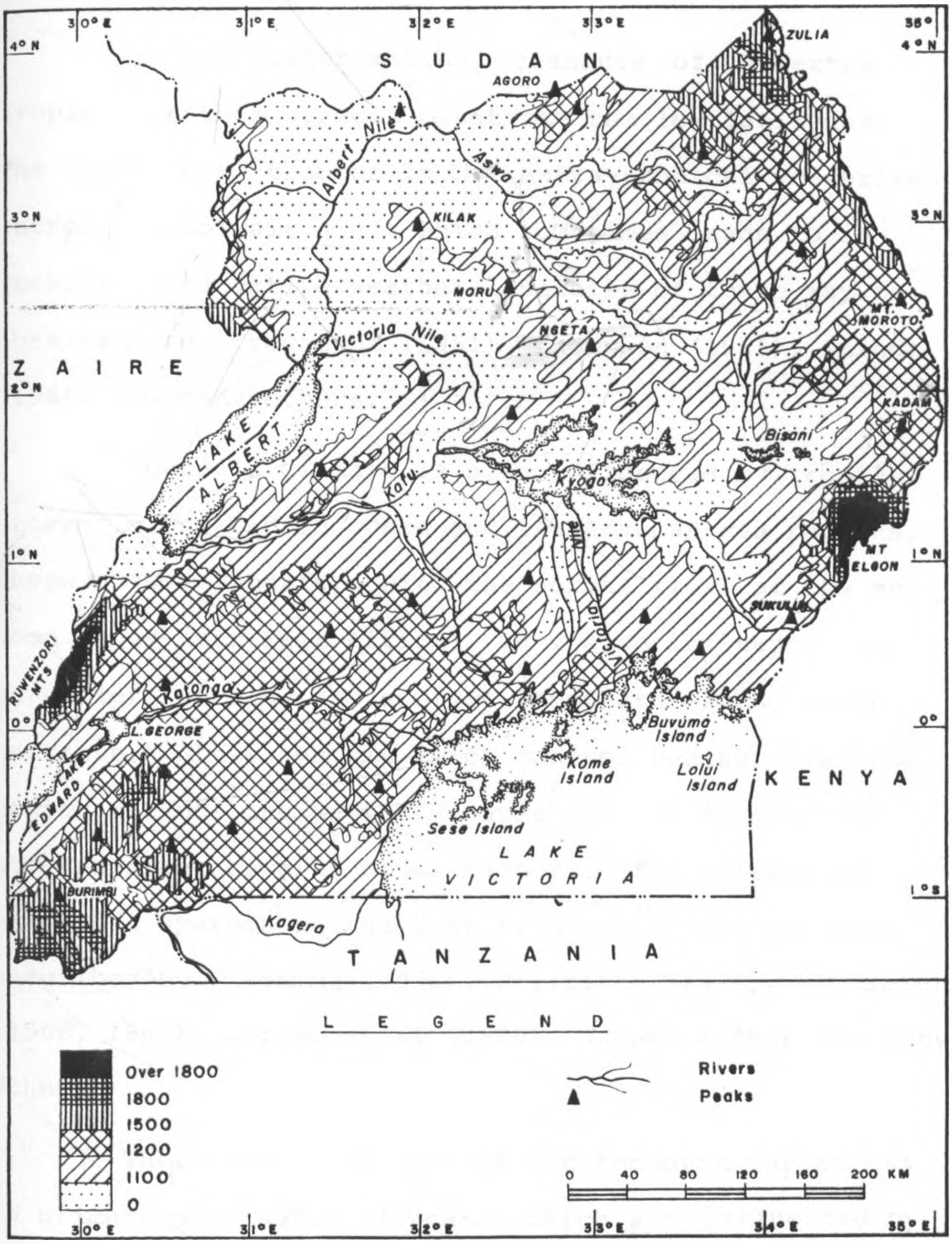


Fig. 3: Relief map of Uganda

The influence of these westerly winds are generally increased by the intensification of the St. Helena anticyclone.

In the winter months, remnants of the extra-tropical weather systems penetrate the low latitudes. The upper level troughs in the extra-tropical westerlies introduce cool air aloft which sometimes leads to instability and active weather. Details of these may be obtained from Trewartha (1961), Thompson (1965), Ininda (1985) and many others.

Uganda's climate is also influenced by tropical storms in the Indian ocean and the Arabian sea regions. These systems interfere with the normal flows of the monsoon winds over East Africa. They also influence ocean parameters such as sea surface temperatures and ocean currents. Other systems which control weather over the region include the Easterly waves and the African jet streams among several other factors. The effects of all these systems on the East African Climate has been described by Flemming (1970), Griffiths (1972), Findlater (1968, 1969), Ogallo et al (1988), Ininda (1985), and many others.

There are large spatial and temporal variations in climate over parts of Uganda which are influenced by some regional features. These features include the large water bodies and complex topographic features. The large water bodies include lakes Victoria to the south, Kyoga in the central, Albert in the northwest and lakes Edward and George in the southwest of Uganda. The major rivers include river Nile originating from lake Victoria and passing through lakes Kyoga and Albert as it flows northwards into the Sudan and beyond. Other rivers include Katonga in the South and Aswa in the north among many other streams and swamps.

The major mountains are Elgon to the East and Rwenzori to the West. Other highlands include the hills of Kigezi in the southwest, the craggy peaks of Moroto, Kadam and many others in the north-east as well as the northern central highlands of Kilak and Moru. Besides these mountains and highlands are valleys; notably the western shoulder of the great rift valley. This part of the great rift valley runs through lakes Edward and George, along the Uganda/ Zaire border, through lake Albert and forms part of the Albert Nile valley northward. Many of these physical features are shown in Figure 3.

These physical features introduce significant modifications in the general wind flow over the region. A good example of the impact of the regional systems is observed over lake Victoria. This is one of the largest fresh water lakes in the world with an areal extent of about 70,000 km^2. Lake Victoria has a vigorous circulation of its own. A trough of low level pressure extends over the lake and parts of central Uganda throughout the year (Ogallo, 1988). The trough moves towards the centre of the lake at night (Asnani et al, 1979). The intensity, location and orientation of this trough has a significant influence on the rainfall of the lake region (Anyamba, 1983, Ininda, 1985, Asnani et al, 1979).

Like in most of the tropical regions, rainfall is the weather parameter with the highest spatial and temporal variability. Table 2 gives some indications of the spatial distribution of the annual rainfall.

Table 2: Percentage of Land Area of Uganda Receiving some Selected values of annual Rainfall Amounts

Rainfall (mm)	Percentage Area
Less than 500	12
500 - 750	10
750 - 1250	72
more than 1250	6

(Griffiths, 1972)

The above distribution, however, does not take into account other temporal rainfall characteristics.

When a rainy month is considered to have mean precipitation of at least 50 mm (Griffiths, 1972, Nieuwolt, 1978) and a rainy season considered to constitute a sequence of such months, the spatial distribution of Uganda's rainfall can be grouped into three broad categories. Areas of rainfall throughout the year, bimodal and unimodal rainfall structures. Examples of these are shown in Figures 4a-e.

Areas which receive rainfall throughout the year are concentrated around large water bodies due to the lake/land breeze effects. An example of this is shown in Figure 4a.

Bimodal seasonal rainfall distributions generally dominate over the Southern, Central, Eastern and Western Uganda (Figure 4b).

Over parts of central and Nothern Uganda, a unimodal seasonal rainfall structure is observed between March and November (Figure 4c).

The seasonal rainfall peaks which are observed at the various locations can be grouped into three categories. The two seasonal peaks which are centred around March to May and October to November are associated with the double passage of the ITCZ. The third peak is related to Congo airmass dominant at some locations during July/August (Figure 4d).

The driest part of Uganda lies to the north east in the wind shaddow of the mountain chain along the Kenya/Uganda border. Figure 4e shows the rainfall

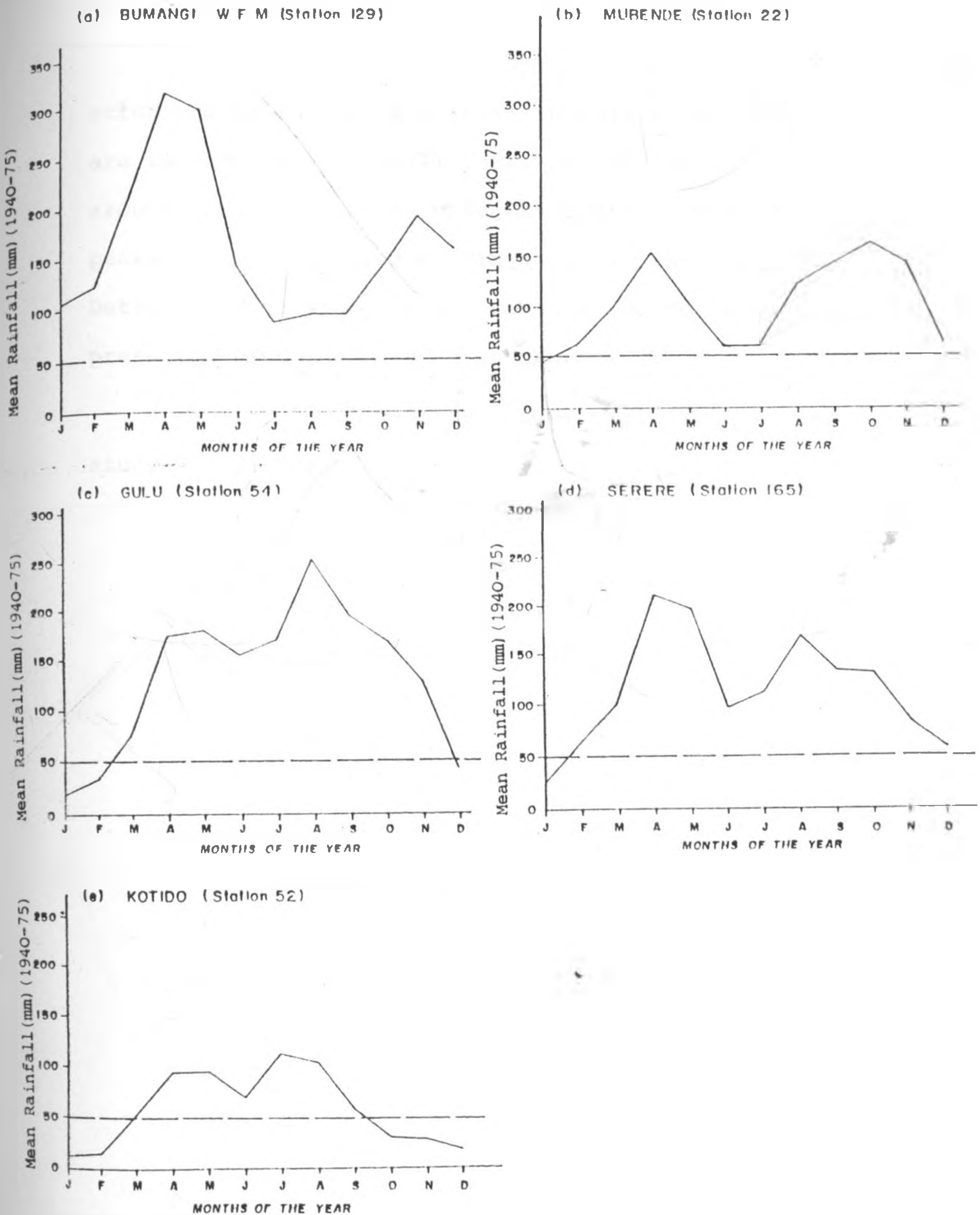


Fig 4 Examples of the seasonal rainfall distribution in Uganda

structure of Kotido, a station in this area. There are two seasonal rainfall peaks in this region centring around April/May and July/August months. The rainfall peaks are concentrated within a single rainy season. Detailed discussions of the seasonal patterns will be presented with the results in chapter 3.

In the next chapter, the methods used in the study are discussed.

CHAPTER 2

METHODOLOGY

2.0. INTRODUCTION

This chapter is divided into four major sections based on the various steps which were undertaken to fulfill the objectives of this study. The first section examines some of the methods that can be used in the estimation of point rainfall values, while sections two and three are devoted to areal estimation of the rainfall records and quality control of the data used.

In the fourth section, some of the methods which can be used in the design of the minimum and optimum raingauge networks are discussed. Lastly the ability of the generated minimum and optimum network designs in the representation of the regional rainfall informations are examined.

2.1 ESTIMATION OF MISSING RECORDS

Many methods have been used to estimate missing climatological records. Details of these can be obtained from U.N.(ECAFE) - WMO (1960), WMO (1962) and many standard references in climatology . The next two sections will review only the isopleth and correlation methods which were used in the study.

2.1.1 THE ISOPLETH METHOD

Under this method, missing data records are estimated using the informations from the neighbouring stations. This involves mapping all the available records for a period with missing data at some location. Lines of equal magnitude (isopleths) are then drawn. The missing values are extrapolated from the magnitudes of the two isopleths which enclose the location of the missing record.

The method, however, requires a dense network of stations in order to give good estimates of the isopleths. It is also very tedious especially when it involves the estimation of many missing records.

An advantage of this method is the possible inclusion of knowledge of physiographic relations, storm tracks, storm types and other factors in the drawing of the isopleths.

2.1.2 CORRELATION METHODS OF ESTIMATION

This method first determines whether a missing variable is related to some other parameter(s) at the locality or variables at other locations. This is generally done through correlation analysis. Regression methods are then used to derive the mathematical function(s) which can describe the relationship between the variables.

Simple linear relationships between the records with a variable with missing values (x_t) and a related parameter (y_t) at any given time t may be expressed as:-

$$x_t = \alpha y_t + \beta \quad .... \qquad [1]$$

where α and β are coefficients to be estimated from the available data. Equation (1) can then be used to estimate missing x_t values using the available y_t records. A more generalised approach has, however, been based on the ratios of the arithmetic averages of the form:-

$$x_t = \frac{\bar{x}_t}{\bar{y}_t} y_t \quad \ldots\ldots \qquad [2]$$

An advantage of the method is that the significance of the estimated records can be statistically tested using standard statistical methods.

A major disadvantage of the method is the development of the mathematical function(s) relating the variables, since some locations may have unique characteristics which cannot be accounted for by any simple function.

For variables which have large spatial and temporal variations like rainfall, areally averaged indices are better indicators of the synoptic and regional scale fluctuations than records from individual locations. Some of the methods which can be used in areal rainfall estimation are discussed in the next section.

2.2. AREAL AVERAGES

Methods which have been used to give areally averaged indices have ranged from the simple arithmetic mean, isohyetal, the Thiessen polygon, the isopercentile, empirical methods to remote sensing and other complex weighting methods like those derived from empirical orthogonal functions. Few of these methods are discussed briefly in the next sections.

2.2.1 ARITHMETIC MEAN

The arithmetic mean is one of the simplest objective methods of obtaining average rainfall over a region (WMO, 1974). Under the method, the rainfall totals (x_i) at all of the stations located within a homogeneous area are divided by the number of stations in the region in order to the derive an areal averaged value. The method is suitable for areas with large numbers of rain-gauges which are evenly spread within any homogeneous climatic zone. It cannot, however, be a good method over regions with large spatial variations in rainfall characteristics which may be induced by topography, large water bodies, and other meso and macro-scale features.

2.2.2 THE ISOHYETAL METHOD

To obtain areal rainfall using this method, a suitable map of the area is required. Mill (1908) defined scales of suitable maps for use in drawing of isohyets for the British isles. The station locations are drawn on this map and the point rainfall totals over the desired time period plotted at the station locations. Isohyets (lines of equal rainfall amounts) are then drawn on the map. Glasspoole (1922, 1928) gave detailed principles for the drawing of isohyets across ungauged spaces on the map.

To obtain mean areal rainfall, the area between pairs of isohyets is obtained. The mean rainfall between pairs of isohyets is multiplied by the area between them to form products. These products are summed for the whole region. The sum of these products divided by total area of the region of interest is an estimate of the mean areal rainfall.

The method is very accurate when skilfully used over areas with uniform topography (WMO, 1974). However, the method is tedious and time consuming when areal rainfall values over many regions are required. Conrad and Pollack (1962) have also observed that the rules of Glasspoole (1922, 1928) for drawing of the isohyets generally present many problems including thorough knowlege of the locations of the gauges. These data are used to provide prevailing wind directions, terrain and other aspects which can be used to modify the patterns of the isohyets (Reed and Kincer, 1917).

2.2.3 THE ISOPERCENTILE AND THIESSEN POLYGON METHODS

The isopercentile method attempts to reduce the necessity of obtaining isohyetal maps for each time period by utilizing the mean annual rainfall of the points (rainfall stations) and the annual general rainfall of the area obtained once through the use of the isohyetal method. The average areal rainfall has been given by the U.K. Meteorology Office (1975) as:-

$$\text{Average Areal Rainfall} = \frac{AAGR}{N} = \sum_{n=1}^{N} \frac{T_n}{AAR_n} \quad \text{......} \quad [3]$$

where:-

AAGR is the annual average general rainfall of the area determined through the use of the isohyetal method (Section 2.2.2) for a period of normally 30 years,

AAR_n is the average annual rainfall at the n^{th} station taken over the same period as above,

T_n the total rainfall of the n^{th} station, and

n the total number of stations.

Under the Thiessen polygon method, the areas enclosed by each station are represented by the Thiessen polygon weights (Thiessen, 1911). Lines are drawn on a suitable map connecting adjacent stations. The perpendicular bisectors of these lines form a pattern of polygons, each around a station. The areas of those polygons, found through planemetry or other techniques as fractions of the area of the region of interest form weights for the rainfall totals for their corresponding stations. The sum of the products of these weights and the corresponding rainfall totals at the stations give an estimate of the average areal rainfall.

These two methods use simple weights and good results may be obtained for the areal rainfall estimate.

Basalirwa (1979), for example, reported obtaining results similar to those of the isohyetal method using the iso-percentile techniques for some catchments in the upper Tana river. However, the isopercentile method still requires suitable base maps to compute the annual average general rainfall (AAGR) using the isohyetal method (section 2.2.2) which may not generally be available. In the case of the Thiessen polygon method, apart from the maps, new weights for the stations have to be recomputed each time a station is added or removed from the network or has a missing record.

2.2.4 EMPIRICAL METHODS

A number of studies have used empirical observations between rainfall, altitude, latitude, longitude, season and many other surface parameters in the estimation of point and areal rainfall values. Under these methods the rainfall records from existing raingauges are used to derive empirical models which relate rainfall as a function of the various spatial and temporal parameters. The equations developed can then be used to estimate rainfall values without further use of the raingauge network.

Examples of these models include the quintic trend surface models developed by Kagenda (1976) for the 37 successive 10 day periods over the year for Uganda using latitudes and longitudes of the areas alone. Kagenda (1976) did not, however, verify his models over an independent period of time. It is also difficult to believe that a complex variable, like rainfall, can be expressed in terms of latitudes and longitudes alone without considering other physical factors which control it.

Gregory (1965) used linear regression relationships to express annual total rainfall as a function of the altitude, latitude, longitude and distance inland of a station from the coast in Sierra Leone. It should be noted that rainfall generating processes along coastal areas are significantly influenced by mesoscale circulation systems like the land and sea breezes, where topography and distance inland from the coast may play some significant roles. Gregory (1965), however, concluded that these four orthogonal factors accouted for only 75% of the total observed annual rainfall amounts.

Other studies have sought to optimize the weighting functions of the various rainfall influencing factors to derive the rainfall models. These techniques known as kriging have usually been advocated in circumstances where the mean rainfall is known to be stationary, which may not usually be the case. Such studies have included the works of Chua and Bras (1982), Delfiner and Delhomme (1973), Bras and Rodriguez (1976) and many others. Other methods using least squares optimisation techniques, fitting surfaces to observed rainfall values or using orthogonal functions include those of Edwards (1972), Childley and Keys (1970) and Akin (1971).

A disadvantage of these methods is the need for a good base raingauge network for their calibrations. Such networks are often lacking in many regions. Another problem of their use is the need to specify the time scales (periods) over which particular models can be used. This may necessitate the development of several models for different time scales. These models are also often difficult to develop.

One empirical technique, however, which has been used widely in recent years is the use of the time coefficients derived from empirical orthogonal functions. Some details of this method are described in the next section.

2.2.5 THE USE OF EMPIRICAL ORTHOGONAL FUNCTIONS

Empirical orthogonal functions use the time coefficients to represent areal indices. Details of these areal indices can be obtained from Johnson (1980) and Ogallo (1986). Ogallo (1986), for example, used the uniqueness of the weights on the eigenvectors for each station to represent the degree of correlation between the rainfall at any station and that particular eigenvector. He used these weights to describe mean areal rainfall conditions in a homogeneous rainfall region with a single dominant eigenvector to form a composite index of areal relative wetness/dryness at any time $t(F_t)$ given by:-

$$F_t = \frac{1}{m} \sum_{j=1}^{m} a_j z_{tj} \quad \ldots\ldots \quad [4]$$

where: z_{tj} represents standardized annual rainfall at station j in year t,

a_j a regression weight on the first eigenvector at station j, and

m the number of stations.

A similar principle can be used by representing each homogeneous regional area, with a single dominant eigenvector, as the sum of the dominant eigenvector re- regression weights. The magnitudes of the unique regression weights can, therefore be regarded as fractions of the total area of the homogeneous region influenced by each station.

The next section considers some of the more recent remote sensing methods of determining areal rainfall.

2.2.6 REMOTE SENSING METHODS

Rainfall informations have been derived through the application of remote sensing technology. These methods have included the use of the weather radars and satellites. These are briefly discussed in the next subsections.

2.2.6.1 THE USE OF RADAR

A weather radar, can after suitable calibration be used to obtain rainfall estimates over an area.

The rainfall rate from a precipitating cloud is related to the median rain-drops diameter in the Z - R relation which has been expressed by Battan (1976) as:-

$$Z = AR^{b} \qquad \ldots\ldots \qquad [5]$$

where R = The rainfall rate in $mmhr^{-1}$

Z = The radar reflectivity factor, and A and b constants.

The value of A and b are obtained from empirical observations from the existing ground raingauge network and by measuring the rain-drop spectra on the ground. From these measurements, it is possible to determine the rainfall total over a given time period.

Weather radars, however, suffer from many drawbacks Battan (1976) pointed out some of these which include:

(i) Erroneous values of R (the rainfall rate) mainly because of errors in the measurement of vertical air velocities on which Z depends.

(ii) In mountaneous areas, orographic background echos have to be eliminated to avoid erroneous results.

(iii) Thunderstorm movement across several regions may cause griding problems when areal rainfall values over only one particular region are required rather than the total dissipation of the cloud.

(iv) Z depends on the rain-drop diameter, in light falls, therefore, the radar may fail to detect the rainfall.

(v) Continuous areal monitoring is required. This may be expensive.

(vi) The initial and running costs of the radar are very high and the radar requires specialised training. The radar, therefore, is only suitable in the study of particular storms rather than as a means of obtaining climatological rainfall records.

2.2.6.2 SATELLITE METHODS

Areal and point rainfall data have been estimated from many satellite derived indices including cloud statistics, vegetation indices, outgoing long wave radiation and many others. Some of these are briefly reviewed in the next subsections.

2.2.6.2.1 CLOUD STATISTICS

Rainfall estimates can be derived from satellite data based on cloud statistics. The cloud statistics which have been used include Cold Cloud Durations (CCD), cloud top temperatures, the BIAS method, Outoing long-wave radiation among many others.

In the cold cloud duration method, a threshold temperature of a convective cloud below which all clouds will precipitate is used. Cold cloud durations (CCD's) indicate the time taken by cold clouds whose temperatures are equal or less than the threshold temperature in staying over a specified area. Relationships are then established between these CCD's and point or areal rainfall values from a suitable network of observations in the area (Flitcroft et al, 1986, Dugdale and Millford, 1987, Chadwick, et al, 1986, Ouma, 1988). These relations are then used in obtaining estimates of precipitation amounts.

The BIAS method entails establishing relationships between the area of clouds covering the basin and the cloud types to surface rainfall observations (Barret, 1970, Barret et al, 1986). The BIAS method expresses accumulated rainfall (R) as a function of other parameters in an equation of the form:-

$$R = f(C_a, C_t, S_w, M_c) \quad \ldots\ldots \qquad [6]$$

where: C_a is the cloud area

C_t the cloud type

C_w a synoptic weight, and

M_c a Morphoclimatic weight.

Other cloud statistics which have been used include cloud top temperatures which are related to cloud brightness. The methods attempt to establish relationships between the video brightness of convective clouds and their cloud temperatures to the observed surface rainfall over the area (Radok, 1966, Lethbridge, 1967, Arkin and Meisner, 1987). Another cloud statistic used is outgoing long wave radiation (OLR) which is related to the fourth power of the atmospheric temperature according to Stefan-Boltzmann's radiation law (Krishnarmurti and Ogallo, 1989, Janowiak et al, 1985). The tropical clouds are generally produced from convective systems . Low OLR values are therefore associated with high clouds and convective regions. These relationships can be used to estimate surface rainfall values. Details of the use of these cloud statistics can be obtained from Ouma (1988).

2.2.6.2.2 VEGETATION INDICES

Some climatological rainfall informations have been derived from very high resolution scanning radiometers data by assuming that the distribution of vegetation is a function of environmental factors of which rainfall is an important factor in the tropics. The relationship between rainfall and vegetation has been used by Whitmore et al (1961). Through physical observations of plant communities on the ground, Whitmore et al (1961) were able to obtain useful estimates of rainfall in remote areas.

In recent times, some attempts have been made to map rainfall climatology in terms of the spatial patterns of the observed vegetation species. The satellite data used are normalised difference vegetation indices (NDVIs) derived from the Advanced Very

High Resolution Radiometers (AVHRR) [NOAA, et al 1986, Tucker, 1986, Justice et al, 1986, Hielkema et al, 1986 and many others]. The AVHRRs use the principle that in the red part of the spectrum of in-coming radiation (0.55-0.68µm) chlorophyll in the vegetation causes absorption of incoming radiation. In contrast, in the infred portion of the spectrum, reflectance from spongy mesophyll leaf structures is high. The ratio of these two channels, known as the normalized difference vegetation index (NDVI) which is given by:- NDVI = (IR-red)/(IR + red), has been shown to be highly correlated with vegetation parameters such as green-leaf biomass and green-leaf area, and absorbed photo-synthetically active radiation and hence can be used for vegetation discrimination (Justice et al, 1986). Values of NDVI have then been correlated with rainfall values from a suitable ground raingauge network. These relations are then used to monitor spatially everaged climatology of the precipitation characteristics since the data (NDVIs) are often provided at grid points.

2.2.6.2.3 LANDSAT DATA

Landsat data is mainly used for monitoring earth resources. The data can, however, be used to give some details of the spatial distribution of the landscape and surface features. These informations can be used to derive some spatial climatological characteristics of rainfall in terms of the impacts of some of natural resources on climate. Rainfall is the climatic parameter with the highest space and time variations over the tropical regions.

Satellite data are useful over remote areas and over the open seas where they may serve as the only means of obtaining precipitation estimates. In settled areas, however, satellite methods are only supplementary to existing raingauge networks. One disadvantage of their use is that a particular satellite does not remain in orbit indefinately. This necessitates periodic changes of receiver sets, which could be different depending on the new owners of the satellite. This is very expensive for developing countries. Other problems of satellite technology are associated with the calibrations of satellite observations.

In this study, arithmetic averages and time coefficients derived from the dominant principal component modes will be used to give areal rainfall estimates for the delineated homogeneous zones. The methods of determining the homogeneous rainfall regions are discussed later in the text (section 2.4). The areal rainfall estimates will be used in testing the ability of the generated network in the representation of the regional rainfall characteristics.

The quality of the rainfall records were, however, examined first using some of the methods which are discussed in the next section.

2.3 DATA QUALITY CONTROL

The quality and quantity of data are the fundamental base of any research. Civilisation and technological advancement can affect the quality of meteorological records. These effects can be due to the use of newer and sophisticated equipments, changes in observation routines and the need to shift stations to other sites due to new equipments and requirements. It can, therefore, hardly be expected that meteorological records will remain strictly comparable over a long period of time at all stations. Some records before and after the changes are introduced will not be comparable. Climatological records must, therefore, be declared homogeneous (a sample from the same statistical distribution) before being applied in any studies.

Some attempts have been made by Ogallo (1981) to examine the homogeneity of East African Rainfall Records. These studies included 15 Uganda stations. The data used also consisted of annual rainfall totals. Even with such a small sample and accumulation of rainfall to an annual time scale, heterogeneity was indicated in the rainfall records of Tororo and Masindi stations.

In the next subsections, some of the standard methods which have been used to test the data quality are discussed. These include the mass curves, correlation methods, trend analyses and several other non-parametric approaches.

2.3.1 THE USE OF MASS AND DOUBLE MASS CURVES

Under these methods, cummulative records (mass curve) or the deviations from the mean (residual mass curve) are plotted versus time. A straight line graph will be obtained from the homogeneous records. Heterogeneity can be deduced from the significant deviations of some points from the straight line. These methods can also be used to adjust heterogeneous records.

Under double mass curves, one or more stations situated within the same climatic zone and known to have homogeneous records are chosen. Temporal variations in these records and those of the test station are then compared. A common practice is to plot corresponding cummuluative values to obtain a double mass curve. A straight line will be obtained for homogeneous test records. The heterogeneous records can be adjusted using the slopes of the graphs of these curves. Details of mass curves can be obtained from Ogallo (1981, 1987), WMO (1970, 1986) among many others.

2.3.2 DATA QUALITY CONTROL USING CORRELATION METHODS

If a variable being tested for homogeneity is known to be correlated with some other climatic variable, the functional relationships between these variables are used under this test.

Let y be the variables to be tested and x the other variable. The correlation coefficient r between the variables x and Y has been given by Gregory (1963), Ebdon (1977), Yamane (1973) and in many other statistics text-books as:-

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2 \right]^{\frac{1}{2}}} \quad \ldots\ldots [7]$$

where r = the correlation coefficient

x_i = the ith value of variable x

y_i = the ith value of variable y

$\bar{x}$ = the mean of variable x,

$\bar{y}$ = the mean of variable y and

n = the number of observations

Equation (7) is tested statistically for significance. If x and y are significantly correlated, the functional equation can then be developed. A simple relationship may take the form:

$$y_i = a + bx_i \qquad \ldots\ldots \qquad [8]$$

with x_i and y_i as in equation (7) a a constant and b a coefficient. The values of a and b can be estimated by least squares as:-

$$b = \frac{\sum_{i=1}^{n} (x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n} (x_i-\bar{x})^2} \qquad \ldots\ldots \qquad [9]$$

and $a = \bar{y}-b\bar{x}$ ...... [10]

Estimated values $\hat{y}_i$ can be generated through the use of equations (8), (9) and (10) as:-

$$\hat{y}_i = a + bx_i$$

The residuals $(\hat{y}_i-y_i)$ are than studied for any unsual trend to determine homogeneity.

A disadvantage of this method, however, is that high correlations may not necessarily lead to meaningful results. These could arise as a consequence of seasonal effects affecting both variables x and y which once removed leads to non-significant correlations between the variables (WMO, 1986).

2.3.3 TREND ANALYSES

Under these methods, temporal changes in the time series are investigated in order to determine any significant shifts in the fluctuations. Trend analyses can be divided into two major divisions, namely, graphical and statistical methods.

The graphical methods incude graphical representation of time averaged values (running means). An example of this type of analysis is given by (WMO, 1986). A variable X_t, $t = 1,2,\ldots,n$, can be averaged over a time interval $m (m \ll n)$, either as overlapping means ($t_1 = 1,2,\ldots,m$; $t_2 = 2,3,\ldots,m+1$, $t_3 = 3,4,\ldots m+2,\ldots$) or as non-overlapping means ($t_1 = 1,2,\ldots,m$; $t_2 = m+1, \ldots, 2m; \ldots$) and the resulting smoothed plot examined for any significant departures.

The statistical methods on the other hand include both parametric and non-parametric methods. An example by Buishand (1982) uses accumulated sums of the deviations from the mean. In the equation

$$S_t = \sum_{i=1}^{t} (x_i - \bar{x}), \quad i = 1,2,\ldots,t \quad \ldots\ldots \quad [11]$$

S_t is the sum of deviations of variable x from the mean up to time interval t.

x_i is the ith observation of variable x, and

$\bar{x}$ the mean of x over all observations n.

If the mean of x is stationary, than the sums S_t should fluctuate about zero. However, if heterogeneity changes the mean from μ, to $\mu + \Delta$ ($\Delta > 0$) at the point t = m say, then before this point all x_i will be mostly less than $\bar{x}$ and S_t will be increasingly negative. After t = m, x_i will be mostly greater than $\bar{x}$ and S_t will become increasingly positive, returning to zero at t = n. That is $S_o = S_n = 0$.

The statistics used to measure the bias in the mean is the greatest range (R) of S_t given by:-

$$R = Max(S_t) - Min(S_t) \quad 0 \leq t \leq n \quad ...... \quad [12]$$

OR The standardized range (R_s) given by:-

$$R_s = R/\hat{\sigma}_x \quad ...... \quad [13]$$

where $\hat{\sigma}_x$ is the standard deviation of x.

The statistics of R and R_s in equations [12] and [13] are not normally provided by the statistical tables but percentage points of R_s for values of sample size n can be obtained from Buishand (1982).

More details of trend analysis methods and their associated problems can be obtained from Kendall (1976), WMO (1986), Siegel (1956) and several other standard references.

2.3.4. MARONNA AND YOHAI METHOD

This method is an extension of double mass curve analysis discussed in section (2.3.1). In this method, however, a quantitative measure of the limits in which the data are declared homogeneous is included.

The method requires that each series of data be distributed normally, be serially independent and be stationary (except for a shift in the mean of the series under test).

For any two random variable x and y, new series of cummulative sums are formed as:

$$x_i = \frac{1}{i} \sum_{j=1}^{i} x_j \quad \text{and} \quad y_i = \frac{1}{i} \sum_{j=1}^{i} y_i \quad \ldots\ldots \quad [14]$$

Then $\overline{x} = x_n$ and $\overline{y} = \overline{y}_n$ when n is the number of observations.

The variances (S_x and S_y) and covariances (S_{xy}) are then computed as given by WMO (1986) as:-

$$S_x = \sum_{j=1}^{n} (x_j - \overline{x})^2, \quad S_y = \sum_{j=1}^{n} (y_i - y)^2 \quad \ldots\ldots \quad [15]$$

and $$S_{xy} = \sum_{j=1}^{n} (x_j - \overline{x})(y_j - \overline{y})$$

For each value in the series i, the following are computed:-

$$F_i = S_x - (x_i-\bar{x})^2 ni/(n-i); \quad i < n$$

$$D_i=(S_x(\bar{y}-y_i)-S_{xy}(\bar{x}-x_i))n/((n-i)F_i), \; i < n. \quad \ldots\ldots [16]$$

and $$T_i = [i(n-i)D^2F_i]/(S_xS_y-S^2_{xy})$$

The value of i when T_i is maximum for the series T_i, i=1,2,....,n-1 indicates the year before the mean change. The value of D_i at this point is an estimate of the amount of change. Critical values of T_i have been produced by Moronna and Yohai (1978).

A disadvantage of the method is that the data has to be normally distributed which may not always be the case with tropical rainfall data.

2.3.5 WALD-WOLFOWITZ ONE SAMPLE RUNS TEST

This nonparametric test is useful in testing for homogeneity of a data record against an alternative hypothesis of trend in the data, bias, drift or oscillation or a combination of two or more of these (Siegel, 1956, WMO, 1979, 1986 and several others). The test has been applied in many climatological studies (Ogallo, 1981, WMO, 1979).

In order to determine the homogeneity of a record, the rainfall process is assumed to be random. The distribution of the number of runs, r, above and below the median has been given by Siegel (1956), WMO (1979, 1986) as being approximately normal with:-

$$\text{Mean, } \mu = \frac{2n_1n_2}{N} + 1 \quad \ldots\ldots \quad [17]$$

$$\text{and variance, } \sigma_r^2 = \frac{2n_1n_2(2n_1n_2-N)}{N^2(N-1)}$$

where: N = record length, N > 40

n_1= number of runs below the median

n_2= number of runs above the median

The normal variate, Z, can be calculated (after correction for the approximation of a discrete to a continuous variable by subtracting 0.5) as:

$$Z = \frac{|r-\mu_r|-0.5}{\sigma_r} \quad \ldots\ldots \quad [18]$$

Z is normally distributed with zero mean and unit variance. The significance of the observed number of runs, r, can be obtained by using the normal, N(0,1), tables. For $N \leq 40$, approximate tables can be found in Siegel (1956).

2.3.6 OTHER HOMOGENEITY TESTS

Other tests for detecting heterogeneity in climatic data include Von Neumann's ratio, Mann-Whitney, Kolmogrov-Sminov, Akaike's Information Criteria (AIC) and many other tests. Details of both parametric and non-parametric tests can be obtained from WMO (1986), Siegel (1956) and many other standard references.

The methods which were used to test the quality of the rainfall records in this study included the

Wald-Wolfwitz runs test and mass curves. Any heterogeneous records were adjusted using double mass cuve method.

The homogeneous records were used to determine the optimum and minimum raingauge network designs for Uganda discussed in the next section.

2.4 RAINGAUGE NETWORK DESIGN

The basic principle of a raingauge network is to meet the needs for precipitation information as efficiently and as economically as possible. A good raingague network must, therefore, be able to provide optimum rainfall information to the various users of the rainfall records.

Three major criteria which can be considered in the design of raingauge network are:-

(i) A basic network for providing precipitation information for wide ranging national and regional data needs such as the desertification or flood problems of the invidual regions and to serve as a basis for climatological data collection to be operated indefinately.

(ii) A secondary network to serve in specific national and regional projects when increased densities are required, for example, in agricultural hydrological and land use observations.

(iii) Specific raingauge networks for research programmes (WMO, 1981, 1989, Rodriquez-Iturbe et al, 1974).

WMO (1981) gives recommendations for the rain-gauge densities for various locations and terrains. These densities, however, are often inadequate, in some cases, for national or regional planning and the management of rainfall activities given the spatial and temporal complexity of tropical rainfall characteristics.

A minimum raingauge network design can best meet criterion (i) above in that it is meant to represent each climatological division of an area in the raingauge network. The rainfall informations obtained from a minimum network may not, however, satisfy the needs specific in criteria (ii) and (iii) above. An optimum raingauge network design is necessary to achieve these criteria. The methods used for a minimum and optimum raingauge network designs are, however, described under separate sections.

2.4.1 MINIMUM RAINGAUGE NETWORK DESIGN

Under a minimum raingauge network design, an attempt is made to represent each climatological division of every area with unique spatial and temporal rainfall characteristics with at least one raingauge in the regional or national raingauge network. This would ensure the availability of rainfall samples regionally or nationally representing all the different distinct spatial and temporal rainfall characteristics. The delineations of the different regional climatological divisions can be derived from the spatial mapping and identification of the dominant rainfall patterns over the area of interest. The methods which have been used to identify the dominant rainfall patterns over the regions include the spatial delineation of climatic zones using graphical methods, cluster analyses, map typing and complex methods involving empirical orthogonal functions. These methods are discussed in the next sections.

2.4.1.1 SPATIAL DELINEATION OF CLIMATIC ZONES

The graphical techniques formed the basis of most of the early attempts of climate zoning. Some of these graphical techniques were based on subjective visual grouping of the graphical results. Among these is the work of Gregory (1954) who defined the general regional climatic patterns of exceptionally high and exceptionally low rainfall values. Composites of both results were used in the final delineation of Southern England into climatic zones.

Glasspoole (1922) on the other hand used isopleths of correlation coefficients derived from annual rainfall to delimit Britain into climatic zones. Kraus (1955) used long term rainfall means and interannual variability in his climatic zoning while Singh (1986) used the relationship between the standard deviation of annual precipitation totals and the average annual rainfall totals to produce probability zoning of the annual rainfall patterns over India.

In East Africa, Johnson (1962) used daily rainfall totals to delimit areas with similar daily rainfall characteristics while probability maps have been used by Glover et al (1954), Griffiths (1958) and Hanna (1975) to delimit uniform climatological zones.

The complex nature of the East African rainfall was reflected from the results of harmonic analyses and the general mean monthly rainfall patterns (Potts, 1971, Griffiths, 1969, 1972). Potts (1971) and Ogallo (1983) have observed unique regional patterns from these harmonic analyses although the boundary divisions were not quite distinct. Griffiths (1969, 1972), however, obtained 52 distinct regional patterns from the spatial patterns of mean monthly rainfall graphs of East Africa.

Other methods of spatial climate zoning which have developed with the advancement of computer technology are highlighted in the next sections. These include cluster analyses, map typing methods and empirical orthogonal functions.

2.4.1.2 CLUSTER ANALYSIS

This method often involves the use of running means, mean monthly or annual totals, or the correlation matrix between stations (variables). Under this method, variables (stations) with similar charcteristics are clustered together. An example of the correlation method of clustering is the linkage analysis which was used by Gregory (1965) to group rainfall stations of Sierra Leone.

The linkage analysis uses the interstation correlation matrix to determine the stations which are closely related. The method examines the highest correlation coefficients for each column which is then underlined. The highest intercorrelated pair in the matrix is then selected as a nucleus of the first group. The rows of nucleus stations in the matrix are then examined to identify all those variables which have their highest degree of correlation with them. These variables together with the nucleus variables form the first group. The next highest intercorrelated pair is then selected and the procedure repeated to obtain a second group. The procedure is repeated until all the variables (stations) have been allocated to the most appropriate group. These patterns of clusters can be used in the design of rainfall networks.

The method works quite well when small numbers of variables are being considered. Sorting out the inter relationships between variables becomes complex as the number of the variables increases.

In general cluster analysis techniques are not based on sound probability models and the results are poorly evaluated and unstable when evaluated, Hartigan (1975). Some of these drawbacks however, are being overcome with the statistical significance tests being proposed (Sneath, 1979).

2.4.1.3 MAP TYPING METHOD

The aim of map typing techniques is to identify a subset of a data patterns which occur most frequently, thus summarizing large sets of data into meaningful patterns which can be identified. The linear correlation map typing technique was developed by Lund (1963). Lund's work was based on the identification of recurring map configurations of sea level pressures which appeared more frequently than expected by chance during the winter months of December to March (1949-53) for a small area of North Eastern United States. A correlation matrix of mean sea level pressure observations was first obtained. A threshold value of the significance of a correlation (r), of $r > 0.7$, was adopted. A map pattern which had the largest number of observations correlated to it at that level of significance or more was removed from the correlation matrix and designated map type A. The same procedure was then applied to the remaining correlation matrix and designated map type B and so on until 10 map types had been extracted. The pressure patterns of these maps A to J were then examined and classified to determine the final classification of the regional climatic patterns.

Map typing patterns with modified correlation methods have also been used by Blassing and Lofgren (1980), Paegle and Kierulff (1974), Nicholson (1986) and many others. Nicholson (1986), for example, defined an index of correlation representative of rainfall departures from the mean as:-

$$C_{jk} = \frac{\Sigma R_j R_k}{\left[\Sigma (R_j)^2 \Sigma (R_k)^2\right]^{\frac{1}{2}}} \quad \ldots\ldots \quad (19)$$

where R_j and R_k represent regional rainfall departures in the years j and k respectively.

C_{jk} the index of correlation.

and summation is made over all rainfall regions on the map. The proceedure to obtain map types has been explained in six steps as:

(1) Correlate the pattern for each year i, with that of every other year.

(2) From the matrix obtained in (1), a year which has the most correlations above an arbitrary set value is selected.

(3) A composite pattern is derived from the average with the two best correlated patterns with it and is designated anomaly type 1.

(4) All the years are now correlated with type 1, and those correlated above the chosen significance level are labelled type 1 years and are removed from the data set.

(5) Steps 2 and 4 are repeated to obtain types 2,3,.....etc. and

(6) All types obtained are now correlated with all the years, and years reclassified where appropriate when their correlations correlate better with other types than the ones where they were originally designated.

Some of the problems associated with the map typing methods includes the fact that they are time consuming, tedious and the assumption of orthogonality (independence) between the variables (stations). It is, however, well known that weather characteristics at neighbouring stations have some similarities, especially, where synoptic scale effects are dominant.

2.4.1.4 EMPIRICAL ORTHOGONAL FUNCTIONS

More recent research work of data reduction and regionalization of climate has employed Empirical Orthogonal Functions (EOF). The advantages of EOF's include:

(i) Ability to reduce the dimensionality of a complex problem by replacing the several complex variables with fewer orthogonal ones.

(ii) Reduction of the amount of data processed.

(iii) Ability to represent complex variables in terms of recognizable physical processes controlling the variables with the patterns of each orthogonal function representing each of those processes.

(iv) They do not require even (equal distant) points of observations, unlike orthogonal functions.

(v) Their time coefficients can be used as composite indices of the temporal anomalies.

The EOF's will therefore be used in this study because of their above characteristics and advantages in data analyses.

The basic principles of EOF's are derived from the concepts of variance. In general, the first step of EOF solutions involves the computations of some measure of association between the set of variables. This is followed by the construction of a set of orthogonal functions which are representative of the complex variables. The two methods which have been commonly used to describe the orthogonal solutions are the Factor and Principal Component Analyses. These are briefly discussed in the next sections.

2.4.1.4.1 FACTOR AND PRINCIPAL COMPONENT ANALYSES

Factor and Principal Components analyses constitute the two methods most commonly used to describe solutions of empirical orthogonal functions. Their central role is the orderly simplication of diverse numbers of interrelated measures into few understable ones. Therefore both methods have been extensively applied in many research studies including meteorology. Details of the theory on which they are based can be found in Harman (1976), Child (1970) and many advanced statistics textbooks dealing with data analysis and simplification.

In this section, therefore, only some of the more important details of these methods are provided.

2.4.1.4.2 FACTOR ANALYSIS

The classical factor analysis model designed to maximally reproduce the correlations r_{jk} between variables j and k over individual observations $i(i = 1,2,....,N)$ assumes that the observed correlations in the data are the results of an underlying regularity. The observed variable is influenced by various determinants, some of which are shared by other variables in the set, while others are not shared by any other variable. That part of a variable that is influenced by the shared determinants is called common, and the remaining part is called unique. The unique part of a variable, therefore, does not contribute to the relationship between the variables.

It is assumed also, that the common determinants will not only account for all the observed relations in the data, but will also be smaller in number than the variables, n. The basic factor model may therefore be put in the form given by Harman (1976) as:-

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + \ldots a_{jm}F_m + U_jB_j \quad (j=1,2,\ldots,n)$$

OR

$$Z_j = \sum_{k=1}^{m} a_{jk}F_k + U_jB_j \quad \ldots\ldots \quad [20]$$

where Z_j = variable j in standard form

F_k = hypothetical factor k.

a_{jk} = standardized multi-regression coefficient of variable j on factor k (factor loading).

U_j = Standardized regression coefficient of variable j on unique factor j.

B_j = Unique factor for variable j.

In model equation [20] it is assumed that the correlations

$$r(F_k, B_j) = 0 \quad (k = 1,2,\ldots,m;\ j=1,2,\ldots,n) \ldots [21].$$

$$r(B_j, B_k) = 0 \text{ when } j = k \quad \ldots\ldots \quad [22]$$

That is, the unique factor B_j, is orthogonal to all common factors and unique factors associated with other variables in the set. This means that the unique portion of a variable is not related to any other variable or that part of itself due to the common factor.

The model given in equation [20] for j = 1,2,....,n forms the factor pattern. To obtain a_{jk} of the factor pattern , the data, in this case the monthly or seasonal rainfall totals are normalized to have zero mean and unit variance. A matrix of interstation correlation coefficients is then derived.

Any correlation between the variables (station records) j and k is assumed to be orthogonal. If there are m common factors in the data of n variables (stations) usually (m < n) the fundamental factor theorem as given by Nie et al (1970) can be utilized to obtain:

$$r_{jk} = r_jF_1r_kF_1 + r_jF_2r_kF_2 + \ldots + r_jF_mr_kF_m$$

$$= a_{j1}a_{k1} + a_{j2}\, a_{k2} + \ldots + a_{jm}a_{km}$$

$$= \sum_{i=1}^{m} a_{ji}a_{ki} \qquad \ldots \qquad [23]$$

Thus the correlations between variables j and k is the sum of the cross products of the correlations of j and k with their respective common factors.

If only a single common factor exists in equation [23], then it reduces to:

$$r_{jk} = r_jF_1r_kF_1 \qquad \ldots\ldots \qquad [24]$$

Meaning that the correlation between variables j and k is due sorely to factor F_1. If F_1 is controlled, then the partial correlation between j and k would be zero.

This approach has been used by Kagenda (1975) to delimit Uganda into climatic zones. The records used were 10 day rainfall totals. Ogallo (1988) observed, however, that the use of 10 day rainfall totals would be expected to induce errors in the covariance or correlation matrix during the dry months when many zero values exist. These errors would be propagated into the factor analysis solutions. The method has also been used by Barring (1987) for daily Kenyan rainfall and in many other researches.

The major problem with factor analyses is the determination of the unique component, U_jB_j. Many studies have neglected the unique component of the variance. This reduces the factor analysis model to Principal Component Analysis (PCA) model of the next section.

2.4.1.4.3 PRINCIPAL COMPONENT ANALYSIS (PCA)

In meteorological research, the difficulties in estimating the uniqueness component in equation [20] necessitates a principal component approach in factoring. Under PCA, the uniqueness is ignored. That is, in the correlation or covariance matrix used in factoring, the correlation of a variable with itself, r_{ii}, is given by

$$r_{ii} = 1 \quad \text{all } i \qquad \ldots\ldots \qquad [25]$$

Thus no account is taken of observational or instrument errors. Equation [20], therefore, becomes:

$$Z_j = \sum_{k=1}^{m} a_{jk}F_k \quad \ldots\ldots \quad [26]$$

Furthermore, the common variance of Z_j (equation [26]) can be deduced to be given by:

$$S_j^2 = \sum_{k=1}^{m} a_{jk}^2 \quad \ldots\ldots\ldots \quad [27]$$

Equation [27] signifies that the common factors m represent the variance of the variable. Thus all the common variables are equally determined by the common factors. These assumptions although not justified in real phenomena, have provided useful information especially in the representation of complex variables.

Examples of the application of PCA include the work of Gregory (1975) on the delimitation of patterns of climatic fluctuations in Britain. Rinne and Karhila (1979) on the studies of 500mb heights in the Northern hemisphere; Walsh and Mostek (1980) on quantitative description of meteorological anomaly patterns over the United States between 1900-1977; Cohen (1983) in the classification of 500mb anomaly patterns; Peak at al (1986) on cyclone forecasting, Barnston and Livezey (1987) on the classification, seasonality and persistence of low frequency atmospheric patterns; Ogallo (1980) on the classification of East African rainfall stations into homogeneous groups; Bärring (1987) on the determination of spatial patterns of daily rainfall in central Kenya; Ogallo (1983) on the regionalization of spatial and temporal patterns of East African stations and many others.

Ogallo (1980, 1989) included only 15 Uganda rainfall stations in deriving 3 and 6 major climatic zones of Uganda respectively. This climatic classification of Uganda is not realistic over some areas especially regions with high spatial and temporal rainfall variability.

Some of the major problems in using PCA are the determination of:

(i) The number of significant principal components (eigenvectors) and

(ii) The physical reality (stability) of the delineated zones. Details of these problems are discussed in the next two sections.

2.4.1.4.4 NUMBER OF SIGNIFICANT PRINCIPAL COMPONENTS (EIGENVECTORS)

Both principal component and factor analyses deal with falliable data. Individual measurements and correlations among variables are subject to the vicissitudes of sampling. As a consequence, there is sampling variation present in the results of the analyses. The judgement concerning the statistical significance of the number of common factors, m, to be extracted during the factoring should be based on their contribution to the reproduced correlations as related to the actual sampling variations of these correlations. Many researchers, therefore, have proposed various methods of determining the numbers of significant principal components (factors). In the next four subsections, some of these methods are discussed.

2.4.1.4.5 KAISER CRITERION

One of the simplest methods of determining the significant principal components was developed by Kaiser (1959). The method assumes that all principal components with eigenvalues greater or equal to one are significant.

The Kaiser criterion has the advantage of ease of application to computer factoring and has been incorporated in many computer subroutines (Nie et al., 1970).

2.4.1.4.6 THE SCREE METHOD

Another approach to finding the appropriate number of orthogonal vectors of practical significance in the factoring is the scree method. The scree method assumes that the component variance approaches some constant value when components are measuring random errors. When the number of components is graphically plotted against the proportion of variance extracted in the unrotated components result, an exponential curve is obtained. The scree test determines the number of significant principal components from the point where the graph starts to be almost linear (Harman, 1976, Cohen, 1983).

2.4.1.4.7 THE LOGARITHM OF THE EIGENVALUE (LEV) METHOD

Another method of finding the significant components to be retained in the factoring is based on the criterion of the logarithm of the eigenvalues (LEV) developed by Craddock and Flood (1969). It has been used by Rinne and Järvenoja (1979) among others. The method assumes that meteorological errors are in logarithmic progressions. It therefore modifies the scree method by plotting the logarithms of the eigenvalues versus the component number. The components of significance are determined from the LEV graph where the graph can be started by approximating it with a straight line.

2.4.1.4.8 USE OF SAMPLING ERRORS OF THE EIGENVALUES

A method proposed by North et al (1982) of finding the significant components to be retained in the factoring uses the sampling errors of their associated eigenvalues. In this proceedure, the objective is to determine whether a sample component is a faithful representation of the true eigenvalue.

In their derivation, North et al (1982), suggested that the sampling error is of the order of $(2/N)^{\frac{1}{2}}$, where N is the total number of observations (cases). A first order shift in the component is shown to be strongly dependent upon the spacing of the eigenvalues, whereas the shift in the eigenvalues does not depend on their spacing.

North et al (1982), therefore, proposed that if the sampling error of a particular eigenvalue, λ_i, given by: $[\delta\lambda_i = \lambda_i (2/N)^{\frac{1}{2}}]$, is comparable to or larger than the spacing between λ_i and λ_{i-1}, a neighbouring eigenvalue, than the component associated with λ_i will be comparable to the size of the neighbouring component associated with λ_{i-1}.

The interpretation of the North et al (1982) rule is that if a group of true eigenvalues lie within one or two $\delta\lambda$ of each other, then they form an "effective degenerate multiplet" and the sample components are a random mixture of the true components. The difficulty in choosing the proper linear combination within the multiplet would lead to enhanced sampling errors, therefore, it is better to cut off factoring at eigenvalue λ_{i-1}.

A difficulty with the application of this method lies in the necessity of the inclusion of a separate subroutine in the factor programme to compute sampling errors and the spacing between λ_i and λ_{i-1}, to determine the significance of the component associated with λ_i. This complication cannot be easily incorporated in routine factor programmes. Alternatively a halt in the factor programme at the unrotated components has to be made to find out the sampling errors $[\lambda_{i-1}-\lambda_i]$ to determine the significant components to be retained in the rotation before restarting the programme to obtain rotated components. This process would be very costly in computer time.

Harman (1976) notes that practising statisticians admit that tests of significance are in many cases superfluous and may or may not indicate meaningful factors. Never the less, the Kaiser criterion, the scree test, LEV method and sampling errors test have been included in this study.

The next section is devoted to the stability of the number of Principal Components which are declared significant by the above methods.

2.4.1.4.9 PHYSICAL REALITY (STABILITY) OF THE PCA PATTERNS.

PCA assumes that the records are orthogonal (independent). Since meteorological parameters like rainfall are influenced by synoptic or regional factors, some similarities should be expected between rainfall patterns of neighbouring stations. Such similarities are not taken into account in PCA solutions, making some of the derived patterns physically unrealistic

(Richman, 1981). This problem is sometimes reduced through the rotation of the principal components (Child, 1970) and other verification techniques.

2.4.1.4.10 ROTATION OF PRINCIPAL COMPONENTS

Under the rotations, the frames of reference of the principal components are turned about the origin until some alternative position has been reached. Such changes have been noted to reduce some ambiguities associated with the unrotated component solutions (Harman, 1976, Richman, 1981, Barnston and Livezey, 1986 and many others).

The most common methods of rotation are the varimax and oblique rotations. In the next two subsections these two methods of rotation are discussed.

2.4.1.4.11 THE VARIMAX METHOD

The varimax rotation approach owes its development to Kaiser in 1958 (Harman, 1976). The method defines a simple factor as one with ones and zeros in the column. That is to say, if equation [26] is written in the expanded form for each standard variable Z_j, $j = 1,2,\ldots,n$ as:

$$\left.\begin{aligned} Z_1 &= a_{11}F_1+a_{12}F_2 + \ldots\ldots\ldots+a_{1m}F_m \\ Z_2 &= a_{21}F_1+a_{22}F_2 + \ldots\ldots\ldots+a_{2m}F_m \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Z_n &= a_{m1}F_1 + a_{n2}F_2 + \ldots\ldots+a_{nm}F_m \end{aligned}\right\} \quad \ldots\ldots[28]$$

If factor 1, F_1, is a significant factor on variables Z_1, Z_3, Z_4, Z_6 and Z_n only and is not significant on variables Z_j, j = 2,5,7,8,....,n-1, then the coefficients (loadings), a_{11}, a_{31}, a_{41}, a_{61} and a_{n1} are approximately equal to one; and all other coefficients in column one, that is, $a_{21}, a_{51}, a_{71}, a_{81}, \ldots a_{(n-1)1}$ are approximately equal to zero.

A simplification to ones and zeros in a column is equivalent to maximizing the variance of the squared loadings in each column. The computational model has been given by Nie et al (1970) as:

$$V = n \sum_{k=1}^{m} \sum_{j=1}^{n} (a_{jk}/h_j)^4 - \sum_{k=1}^{m} \left(\sum_{j=1}^{n} a_{jk}^2/h_j^2\right)^2 \longrightarrow \text{Maximum} \qquad \ldots\ldots \quad [29]$$

where V = the variance

n = the number of variables

m = the number of common factors

a_{jk} = the loading of variable j on factor k

h_j = the communality of variable j; given by:-

$$h_j^2 = \sum_{k=1}^{m} a_{jk}^2; \quad (j=1,2,\ldots,n) \qquad \ldots\ldots \quad [30]$$

with a_{jk}, m and n defined as in equation [29].

In general, all the significant principal components are rotated through 90°. This method of rotation has been the most widely used in meteorological research (Ogallo, 1980, 1986, 1988; Bärring, 1987, Oludhe, 1987 and many others).

2.4.1.4.12 THE OBLIQUE METHOD

Oblique rotation involves a similar type of simplying principal to varimax. The requirement of orthogonality between the components, however, is relaxed. In the simplification, the components (factors) are allowed to be correlated if such correlations exist in the data. The initial factor axes are allowed to rotate to best summarize any clustering variables. Such rotation, however, can only be achieved with some visual graphical aid and the discerning eye of the researcher. Sometimes when the angle of rotation is too narrow, there may be no substantial solution.

The approach would be to define the so called reference axes and then simplifying them. That is, to minimize the cross products of the factor loadings. The introduction, however, of reference axes unnecessrily complicates the presentation and the theoretical pay off is negligble (Nie et al, 1970).

A more beneficial approach suggested by Nie et al (1970) is to simplifying the expression:-

$$\sum_{k<\ell=1}^{m} \left(\sum_{j=1}^{n} a_{jk}^2 a_{j\ell}^2 - \frac{\delta}{n} \sum_{j=1}^{n} a_{jk}^2 \sum_{j=1}^{n} a_{j\ell}^2 \right)$$

The a's are the factor pattern loadings, $k<\ell$ and each refers to common factors and δ is an arbitrary value by means of the obliqueness of the solution is maintained.

The problem of quantifying complex multiple correlations between variables has limited the application of this method. Good results have, however, been reported by Richman (1981), Bärring (1988), Cohen (1983) and many others.

2.4.1.4.13 OTHER VERIFICATION METHODS

Other methods which have been used to verify the physical reality (stability) of PCA derived climatic patterns include the use of correlation maps, vector plots and the use of the relief map of the region. Some of these methods are discussed in the next two subsections.

2.4.1.4.14 USE OF CORRELATION MAPS

In this approach, the stations with the highest communality (equation [30]) in each cluster are chosen as focal points; since the stations with the highest communality correlate best with other stations in the cluster. These stations are used as reference centres while plotting on the maps the interstation correlations. The correlation between the focal point station and itself is assumed to be a positive unit and plotted as one at the location. The correlation between this station and all the others in the data set are then plotted at the various locations.

The spatial patterns of the correlations are then compared with the patterns derived from PCA clustering.

2.4.1.4.15 VECTOR PLOTS AND OTHER METHODS

Under vector plots, the coefficients of the significant principal components are used to map the locations of the individual variables (stations) on the vector spaces of pairs of significant principal components. The physical reality of the cluster patterns can then be quickly assessed from these graphs.

Finally, any physical characteristics of the region, for example, existing climatological informations, topography, large water bodies, land use patterns, remote sensing informations etc. can be used to confirm the physical reality of the dominant PCA modes.

The spatial distribution of the significant PCA modes were used to delineate Uganda into homogeneous climatic zones using the principle that stations with similar temporal characteristics cluster on to similar factors. The significance of any PCA mode at a given location was determined using the standard error (E_a) of a loading from a relation given by Child (1970) as:-

$$E_a = \text{standard error of a correlation} \quad \left(\frac{n}{n+1-r}\right)^{\frac{1}{2}} \quad \ldots\ldots \quad [31]$$

where: n is the number of variables in the factor analysis and

r is the factor number .

A PCA mode is considered significant at any location when its regression coefficient at the location is greater than E_a.

The concept of a minimum raingauge network design under this method therefore assumes that each delineated homogenous region must be represented by at least one raingauge station. Communality values derived from PCA can be used to determine the representative station for a given homogeneous zone. This entails the inclusion of the highest communality station in each homogeneous region in the minimum network design. The correlation coefficient (r) and r^2 were used here to determine the representativeness of the chosen stations for the various homogeneous zones.

2.4.1.5 THE ABILITY OF THE MINIMUM NETWORK DESIGN IN THE REPRESENTATION OF THE REGIONAL RAINFALL CHARACTERISTICS.

This section attempts to determine the ability of the generated minimum raingague network design in representing areal rainfall characteristics within the individual regions. The correlation coefficient (r) between the rainfall totals ($x_j(t)$) at the representative station (j) over time t(t=1,2,--,n) and the areal rainfall estimate from the individual homogeneous groups ($X_g(t)$) were first computed. The significance of the computed correlation coefficient (r) was determined by computing a t statistic in the equation given by Yevjevich (1972) as:-

$$t = r(n-2)^{\frac{1}{2}}/(1-r^2)^{\frac{1}{2}} \qquad \ldots\ldots \qquad [32]$$

where: r = the correlation coefficient between data at any two points and

n = the number of time intervals compared.

The correlation coefficient between any two data sets was significantly different from zero at any desired level of significance when the computed t in equation (32) was more than the tabulated value in the student t(n-2) statistical tables.

Haan (1977) has shown that the square of the correlation coefficient (r^2) between pairs of observations determines the variance of the dependent variable accounted for by the independent variable in a regression relationship. Thus r^2 gave the variance of the regional areal rainfall explained by the representative station rainfall totals.

An alternative method of assessing the representativeness of the chosen station is to use a regresssion model which expresses areal rainfall values as a function of the representative station values. An example of a linear model is given in equation (33) as:-

$$X_g(t) = a + b_i x_j(t) \quad \ldots\ldots \quad [33]$$

where: $X_g(t)$ is the areal rainfall estimate for group g over time t,

$x_j(t)$ is the rainfall total at station j over time t

a is a constant and

b is a regression coefficient.

Analysis of variance (ANOVA) techniques can be used to determine the proportion of variance accounted for by the representative station rainfall totals based on the linear model. Details of the analysis of variance techniques can be obtained from Haan (1977), Harnett and Murphy (1980) and many other standard statistics references.

The rainfall informations derived from the minimum network, however, may not satisfy the requirements of the various users. They may also not sufficiently represent the unique spatial or local rainfall characteristics such as those associated with topography, soil type, land and sea breezes and other factors that affect the nature of rainfall which are common at many locations. These factors necessitate optimization of the minimum raingauge network in order to include the unique characteristics of the individual regions. Some of the methods of optimum raingauge nework design are discussed in the next sections.

2.4.2 OPTIMUM RAINGAUGE NETWORK DESIGN

The objectives of optimum raingauge network designs are to provide optimum rainfall informations to a variety of users by including considerations of the unique spatial or local rainfall characteristics in their designs. It should be noted, however, that economic constraints often limit the operations of even the minimum raingauge network design in many countries.

The methods which have been used to determine the optimum raingauge networks include optimization techniques of the minimum raingauge network derived from PCA proposed in this study, empirical methods developed from consideration of physiographic effects, coefficient of variation and optimum error of interpolation techniques. Some of these methods are discussed in the next sections.

2.4.2.1 THE USE OF PRINCIPAL COMPONENT ANALYSIS (PCA) IN OPTIMUM RAINGAUGE NETWORK DESIGN

Under the method, attempts are made to optimize the minimum raingauge network design which was derived from PCA as discussed under section 2.4.1.

In the optimization scheme, the hypothesis of adequate representation of the homogeneous region by the representative station was accepted only if the representative station values accounted for at least 90% of the total areal variance (section

2.4.1.5). Otherwise, the representative station was declared inadequate to represent the given homogeneous region. More stations were then added in the network design using stepwise scheme as highlighted below.

Let the optimum number of stations in any homogeneous region be m. The equation (33) of the minimum network design now takes the form:-

$$X_g(t) = a + b_1x_1(t) + b_2x_2(t) + \text{---} + b_jx_j(t) + \text{---} + b_mx_m(t) \qquad \ldots\ldots \qquad [34]$$

where: $X_g(t)$ = the average regional areal rainfall for a given homogeneous region (g) at time t

$x_j(t)$ = the rainfall total at the representative station (j) in region g at the same time t.

a = a constant.

$b_1, b_2, \ldots b_j, \ldots, b_m$ = regression coefficients to be determined and

m = total number of stations included in the stepwise regression scheme.

The variables (stations) in the above regression scheme (equation (34)) were increased stepwise, testing at each stepwise stage whether the optimum number of stations (m) in the region has been attained. This was achieved using analysis of variance techniques which determined the proportion of the areal rainfall variance explained at each regression step.

2.4.2.2. LARGE RAINGAUGE DENSITY METHODS

Where funds are not limiting, large densities of raingauges can be established for a period of time using regional climatological knowledge and the physical features of the area. An assumption is then made that the average catch of the dense raingauge network is the "true" mean rainfall for comparative analyses (Kohler, 1958). Various numbers of evenly spaced raingauges are then used to find new averages. The standard errors of these sample means compared to the "true" mean are obtained over various time intervals. A student t test at the desired levels of accuracy is then used to test the representativeness of the sample gauges. An optimum spread of raingauges can then be determined depending on the desired accuracy.

Thom (1951) showed that changes in correlation coefficients with raingauge network densities were small and the error in the sample mean from the "true" mean varied approximately as the inverse of the square root of n, the number of raingauges in the network.

Other empirical studies have shown that the reliability of an n-station areal rainfall mean varies as the reciprocal of n to about the 0.6-0.7 power, in cases of thunderstorms and annual rainfall totals over even terrain.

Large raingauge density studies to find optimum networks include those of Nicks (1965), Balek and Holecek (1965) among many others. However, economic constraints often limit the use of these methods and other methods of optimising raingauge networks described in the next sections have been used.

2.4.2.3 THE USE OF RUNNING COSTS OF A NETWORK

Under this method, both accuracy and minimization of running costs of a network are considered. Bras et al (1976) proposed a method of network design incorporating the spatial uncertainty of the rainfall process, errors in rainfall measurements and their correlations and the non-homogeneous costs of running the network. Bras et al (1976) were able to obtain an optimal raingauge network in terms of numbers of raingauges and their locations, together with the resulting costs and mean error in the rainfall estimates derived from the network.

The assessment of running costs is, however, quite difficult to make and its inclusion may bias the raingauge siting and reduce the accuracy of the rainfall estimates (Gandin, 1970). Bras et al (1976) recommended a final decision on siting to be taken by the planners of the network after consideration of factors like suitability of a site and the availability of observers. It should also be noted that cost benefit factors based only on current activities may prove costly for some future activities.

2.4.2.4. PHYSIOGRAPHIC EFFECTS

Scientific advancement and understanding of the causative factors of rainfall influenced studies into planned raingauge networks especially in uneven terrain (mountaneous areas).

Noting the role of altitude and aspect of the landscape influences on rainfall intensity, Wilm et al (1939) proposed raingauge distributon in mountaneous terrain considering the physiographic features. Under this method, a contour map of the area is used to distribute the raingauges, taking into account altitudinal intervals and the elevation of the surface. This ensures obtaining rainfall samples of different characteristics and hence improved mean rainfall estimates.

Wilm et al (1939) used statistical methods to determine the number of gauges N, which may be needed to sample gross (areal) rainfall within stated limits of accuracy. This number was expressed as:

$$N = S^2/(S_{\bar{x}})^2 \qquad \ldots\ldots \qquad [35]$$

where: S^2 = the variance of the sample mean.

$(S_{\bar{x}})^2$ = the desired variance of the mean rainfall.

Equation [35] cannot, however, give any indications of the probability associated with the specified accuracy. Stein (1945), suggested a modification of equation [35] to be:-

$$N = t^2 s^2 / d^2 \quad \ldots\ldots \quad [36]$$

where: t = the value of the student t statistic from the t tables for desired confidence level and degrees of freedom.

d = the desired difference between population and sample mean.

Equation [36] was used by Helvey and Patrick (1965) in their interception of rainfall by vegetation studies.

2.4.2.5 COEFFICIENT OF VARIATION METHODS

Studies in raingauge network design generally seek to reduce the standard error of the sample estimates. The methods in this section were developed from empirical observations and therefore only their formulations are provided.

Ganguli et al (1951) proposed that in terms of existing numbers of raingauges, n, the new number of raingauges, N could be determined from:

$$N = n\left(\frac{C_{VE}}{C_{VR}}\right)^2 \quad \ldots\ldots \quad [37]$$

where C_{VE} = Existing coefficient of variation

C_{VR} = Required coefficient of variation.

Ahuja (1959) suggested the use of coefficient of variation when the areal rainfall is determined by the simple arithmetic average (arithmetic mean).
In the equation:-

$$N = (C_V/P)^2 \qquad \ldots\ldots \qquad [38]$$

C_V is the coefficient of variation,
P the desired degree of percentage error in the estimates of the areal mean and
N the number of raingauges.
Ahuja (1959) used equation [38] for the raingauge network planning in India.

In Britain, Bleasdale (1965) suggested the modification of the coefficient of variation equation to take the form:-

$$\text{Number of Raingauges,} \quad N = \left\{\frac{2C_V}{\text{Required Precision}}\right\}^2 \qquad \ldots\ldots \qquad [39]$$

where: C_V is the coefficient of variation. It should be noted, however, that equation [39] was derived from empirical obervations in a small catchment.

Rodda (1972) suggested the use of cummulated coefficients of variation of the mean, recommended by Stephen in 1967. Under this method, the required numbers of raingauges over a given area is calculated from:-

$$N = n\left(\frac{C'}{10}\right)^2 \quad \ldots\ldots \quad [40]$$

where: N = required number of raingauges

n = existing number of raingauges and

C' = is calculated as follows:-

Using monthly rainfall totals, the coefficient of variation of the mean, C_{Vm} is derived. Values of C_{Vm} are calculated using monthly rainfall totals expressed as percentages of the average annual rainfall. This calculation is performed over many months (at least 50) to enable construction of a cummulative frequency curve of C_{Vm} values to be made. From the C_{Vm} values, the value of C_{Vm} (C') exceeded on 5% of the occasions is determined. If C' is 10 or less, then the number of raingauges in the network is considered adequate for the region of interest. If C' is more than 10, then the required number of raingauges can be calculated from equation [40].

The choice of C' equal to 10 is arbitrary. Rodda (1972) recommended the application of equation [40] only in small catchments. However, the method is quite tedious especially when large numbers of raingauges are involved over a fairly long period of time.

Another method of finding an optimum number of raingauges to be retained in the network based on the concept of coefficient of variation is that proposed by Dymond (1982) for raingauge network reduction. The method derives an expression for the mean square error of areal rainfall which can then be used to determine the optimum number of raingauges to be left in the network. The relationship is calculated from:-

$$< e^2 >/< R >^2 = \frac{1}{3}C_v^2(1-\rho_n)N^{-\frac{1}{2}} \quad \ldots\ldots \qquad [41]$$

where: e is the difference between the average catch of the raingauges in a network and the average average that would be obtained if the number of gauges is increased to such a large number that further increases cannot change the average.

ρ_n is the average interstation correlation coefficient.

N the number of raingauges

< R > the expected rainfall at any point (areal mean) and

$C_v = (< R^2 > - < R >^2)^{\frac{1}{2}}/< R >$; $< R^2 >$ being the expected squared rainfall and is considered constant over the area.

Dymond (1982) showed that in the case of a linear decay of the correlation coefficient with distance between stations, the root mean square (r.m.s.) error of areal rainfall, expressed as a percentage of < R >, denoted by σ_e can be expressed as:

$\sigma_e = 100(< e^2 > / < R >^2)^{\frac{1}{2}}$; which when equation (41) is used becomes:

$$\sigma_e = 100C_v(\tfrac{1}{3}(1-\rho_n)N^{-\frac{1}{2}})^{\frac{1}{2}} \quad \ldots\ldots \quad [42]$$

when the correlation coefficient decays exponentially with the distance between stations, in the notation of equation [42] Dymond (1982) gives an approximate expression for σ_e, derived by Zawadzki in 1973 as:

$$\sigma_e = 100C_v^{0.88}(1-\rho_n)^{0.44}(C_v^2+1)^{0.06}/(2.23N^{0.46}) \quad \ldots\ldots \quad [43]$$

Thus by knowing the proposed number of raingauges, N, in the proposed reduced network and whether the correlation coefficients decay linearly or exponentially with the distance between stations, the r.m.s. error, σ_e, can be calculated and if tolerable, N will be considered as the optimum number of raingauges.

The above work of Dymond (1982) is useful in an area with a large raingauge density and when the prime objective is to reduce network running costs. In the case of low network densities other methods like the optimum error of interpolation described in the next section can be used.

2.4.2.6 OPTIMUM ERROR OF INTERPOLATION

The problem of finding an optimum network density, resulting in minimum costs and losses can theoretically be solved, but practically, the information needed for a solution to the problem is often difficult to obtain. A purely meteorological formulation of the problem of finding a rational distribution of gauges is often adopted instead of the economic approach. The meteorological approach requires postulation of the expected accuracy in the measurements of the meteorological elements within a given region. Many studies have derived the postulated accuracy levels from the concepts of error of interpolation that should not be exceeded. The error of interpolation has been shown to be uniquely related to the structure function of the given meteorological element (Rodda, 1969, Alaka, 1970, Gandin, 1970, Kagan,(1972a)and(1972b).

It has also been shown that a non-uniform increase in error of interpolation with distance betweeen any two measuring points can be derived from the structure function. The optimum error of interpolation (εopt) at a point equidistant from any two measuring points (point where a maximum error is assumed to occur) can be derived from the principle of optimum weighting. εopt has been given by Gandin (1970), Cislerova and Hutchnison (1974), Hutchnison (1974) and many others as:-

$$\varepsilon opt = \left[1 - \frac{2(r(\frac{1}{2}x))^2}{(r(x)+1+\eta)}\right] \quad \ldots\ldots \quad [44]$$

where: $r(x)$ = the correlation coefficient between points a distance x km apart.

$r(\frac{1}{2}x)$ = the correlation coefficient between points a distance ½x km apart,

and η = the random error at the measuring points.

Thus, the value of εopt given in equation (44) is a function of the correlation coefficient ($r(x)$) and the distance between the points (x). It should be noted that when the distance between the correlated points x = 0, the station is correlated with itself and the theoretical value of the correlation coefficient $r(x) = r(0) = 1$. However, $r(0)$ is often less than 1 due to random errors (η). The value of η has been given by Kagan (1972b), Cislerova and Hutchnison (1974), Hutchnison (1974), Basalirwa (1984) and many others as:

$$\eta = 1 - r(0) \quad \ldots\ldots \quad (45)$$

$r(0)$ can be obtained from the intercept on the graph of $r(x)$ versus x where x = 0. Equation (45) is then used to estimate (η) for the use in estimation of εopt.

It has also been shown by Gandin (1970), Basalirwa (1984), Hutchnison (1974) and many others that the standard error (S_e) at the mid-points of any two measuring stations (at ½x) is of the form:

$$S_e = S\left[1 - \frac{2(r(\frac{1}{2}x))^2}{(r(x)+1+\eta)}\right]^{\frac{1}{2}} \quad \ldots\ldots \quad [46]$$

$$S_e = S(\varepsilon opt)^{\frac{1}{2}} \quad \ldots\ldots \quad [47]$$

where S is the estimated standard deviation of the rainfall estimate at the mid-point. S is generally computed from the square root of the variance.

Equation (47) illustrates that the mean square standard error (S_e^2) cannot exceed the variance estimate (S^2) since $\varepsilon opt \leq 1$, Equation (47) has been widely used to estimate the standard errors (S_e) at the distance $\frac{1}{2}x$ which are located at the mid-points of any two measuring stations. However, since the optimum rainfall estimates at the various mid-points vary in magnitude, the values of S_e at the different mid-points cannot be compared directly without standardization. Optimum rainfall estimates at the individual points have widely been used in the standardization of the computed S_e values. The normalized S_e values are expressed as a percentage (γ) of the optimum rainfall estimate ($\hat{R}_{ij}$) and are of the form:

$$\gamma = \frac{S_e}{\hat{R}_{ij}} \times 100\% \quad \ldots\ldots \quad [48]$$

when equation (47) is used, equation (48) reduces to

$$\gamma = \frac{S(\varepsilon opt)^{\frac{1}{2}}}{\hat{R}_{ij}} \quad \ldots\ldots \quad [49]$$

OR $$\varepsilon opt = \frac{\gamma^2 \hat{R}_{ij}^2}{S^2} \quad \ldots\ldots \quad [50]$$

where $\hat{R}_{ij}$ is the optimum rainfall estimate at the mid-points of stations i and j.

Cislerova and Hutchnison (1974), Hutchnison (1974), Basalirwa (1984) and many others have given the value os $\hat{R}_{ij}$ as:-

$$\hat{R}_{ij} = p\ (R_i + R_j) \quad \ldots\ldots \quad [51]$$

where R_i - the observed rainfall at station i

R_j = the observed rainfall at station j and

p = the optimum weighting probability.

The optimum weighting (p) is usually obtained by assuming that the rainfall field is isotropic (homogeneous) in the region. The value of p at pairs of stations for estimating $\hat{R}_{ij}$ at their mid-points has been given by Gandin (1970), Hutchnison (1974), Basalirwa (1984) and many others as:

$$p = \frac{r(\frac{1}{2}x)}{r(x)+1+\eta} \quad \ldots\ldots \quad [52]$$

where the symbols carry the same meaning as in equation (44).

Thus, the functional relationship between the correlation coefficient (r(x)) and distance (x) can be used to generate p at pairs of observation points. These values of p can then be used to find $\hat{R}_{ij}$ at their mid-points (equation 51).

εopt can be estimated from equation (44) at any mid-point. When this value of εopt is used in equation (49) together with the estimate of $\hat{R}_{ij}$ obtained as above and S estimated from the standard deviation map of the rainfall totals in the region, the existing percentage errors of interpolation (γ) are determined.

To obtain the intergauge distances x; a desired percentage error of interpolation (γ) not to be exceeded at any mid-point in the region is first fixed. Equation (50) can then be used to obtain a value εopt at any mid-point of observation points. When this value of εopt is used in equation (44); it will correspond to a particular value of x. The x value so obtained will be the desired maximum separation between stations to ensure that the particular value of εopt is not exceeded at the point. Thus optimum intergauge distances x are generated at the mid-points for the given γ in the region.

The fundamental method in obtaining these values of x is by finding a relationship between the interstation correlation coefficients (r(x)) and the distances between them in the region. The methods which have been used in finding the relationship between r(x) and x include graphical and regression approaches. These are briefly discussed in the next sections.

2.4.2.6.1 GRAPHICAL METHODS FOR DETERMINING THE CORRELATION-DISTANCE RELATIONSHIP.

Under the graphical approach, a scatter diagram is first drawn. A free hand method is then used to draw the best curve which can be used to describe the functional relationship between the correlation coefficients (r(x)) and the distances (x). This method is very subjective and depends on individual judgement. A better approach to determining this functional relationship is to use regression methods. The linear regression model is briefly discussed in the next section.

2.4.2.6.2 THE LINEAR REGRESSION MODEL

Under this method, a linear relationship is assumed between the correlation coefficient (r(x)) and distance between the stations (x). The linear correlation function model (r(x)) may be expressed as:

$$r(x) = a + bx \qquad \ldots\ldots \qquad [53]$$

where b = the regression coefficient and
a = a constant (intercept at x = 0; r(0)).

The statistical significance of the computed a and b values can be tested using the students t test.

Rodriquez-Iturbe et al (1974), O'connel et al (1977) and many others have noted that an exponential model gives a better representation of the functional relationship between r(x) and x. The exponential model is briefly discussed below.

2.4.2.6.3 EXPONENTIAL REGRESSION MODEL

A two parameter exponential model may be used to express the relationship between r(x) and x as:-

$$r(x) = e^{a+bx} \qquad \ldots\ldots \qquad [54]$$

where a and b carry the same meaning as in equation (53).

Equation (54) can be transformed by taking the natural logarithms of both sides of the equation to give:-

$$\ell n(r(x)) = a + bx \qquad \ldots\ldots \qquad [55]$$

Equation (55) now closely resembles equation (53).

In this study, both linear and exponential models together with the graphical method were used to determine the relationship between r(x) and x values. Analysis of variance techniques were used to determine the best models which could be used to estimate εopt.

Due to large variations between r(x) and x, many authors have attempted to average the correlation coefficient values (r(x)) whithin a given radius of a central distance x (Kagan 1972a). The averaged values of the correlation coefficients ($\hat{r}(x)$) are then plotted against the central distance (x). The purpose of the averaging of the correlation coefficients within a given radius of each central distance is to remove directional effects, topographically induced differences and other effects that affect local rainfall (Kagan, 1972a, Flitcroft et al, 1989). The length of the gradation (Δx, the diameter over which the correlation coefficients are averaged) can be based on the spatial scales of small scale, convective scale and intense precipitation zone scales. This diameter is about 10km for tropical convective systems (Barry and Chorley, 1976). This value (Δx = 10km) was used in the study.

The central distance x for averaging the correlation coefficients can be given as:-

$$x - \frac{\Delta x}{2} \quad \text{to } x + \frac{\Delta x}{2} \qquad \ldots\ldots \qquad [56]$$

where x is the central distance

Δx is the length of the gradation

Finally, the ability of the linear and exponetial regression models given by equations (53) and (55) in representing the functional relationship between the interstation correlations and distance between stations were determined using the analysis of variance technique.

The gauge network optimization methods used in this study included the PCA method, physiographic effects method given in equation (36), Coefficient of variation methods given in equations (37,38, and 39) and optimum error of interpolation. The results from the study are discussed in the next chapter.

CHAPTER 3

RESULTS AND DISCUSSION

3.0 INTRODUCTION

In this chapter, the results are presented and discussed in four parts based on the methods used in deriving the results as outlined in chapter 2.

Part one discusses the results from the estimations of missing data and the quality of the estimated records. The second part examines the quality of the rainfall records which were used in the study.

Areal averaged rainfall values are better indicators of the regional climatic fluctuations and can provide a better representation of the rainfall patterns over the various locations. The third part compares the areal estimates which were derived from the arithmetic mean and PCA methods.

The last part, which forms the main objective of this study constitutes the results obtained for the minimum and optimum raingauge network designs. One of the major principles of the minimum raingauge network design is based on the spatial mapping derived from PCA methods. It is, therefore, necessary to present the results from PCA before discussing the raingauge network designs.

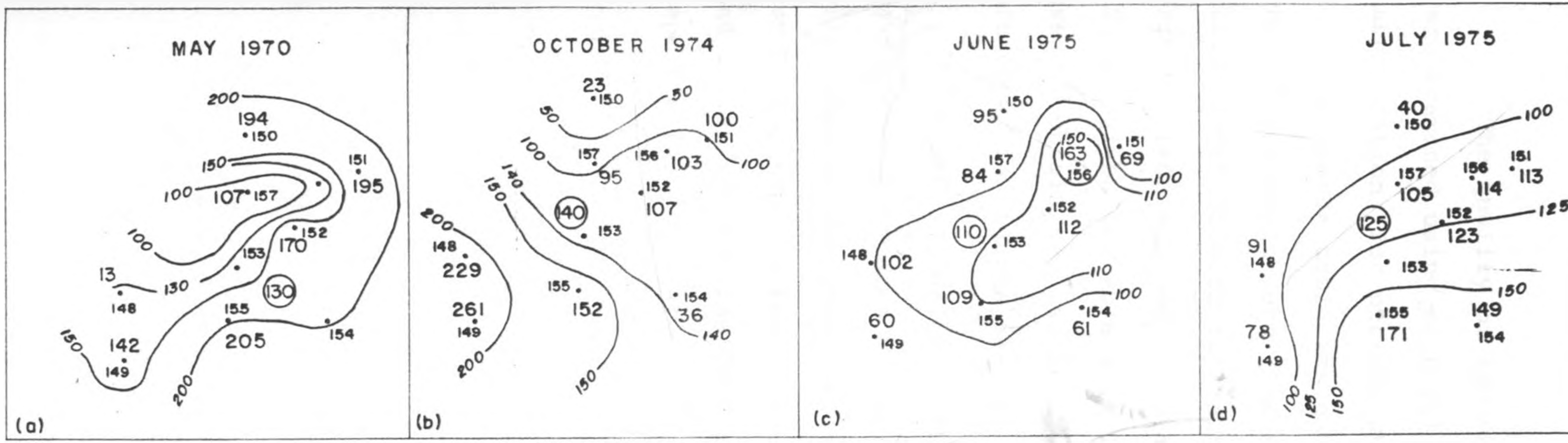
3.1 QUALITY OF THE ESTIMATED RAINFALL RECORDS

Under this section, the results from the estimations of missing data and the quality of the estimated records are presented. The two methods used to estimate the missing data were the isopleth and correlation methods discussed under sections 2.1.1 and 2.1.2 respectively. Examples of the isopleths which were used to extrapolate the missing rainfall records are given in Figures 5a to 5d; for station 153 for some chosen missing data periods. The resulting rainfall estimates are tabulated in Table 3A.

Estimates which were obtained from the correlation method are also given in Table 3A. The station which was correlated with station 153 was station 152. The magnitude of the correlation coefficient between the records of the two locations was 0.66.

Table 3A: Example of the rainfall estimates for station 153 with the isopleth and correlation methods.

Period for Estimation		Estimated Rainfall (mm)	
Year	Month	Isopleth Method	Correlation Method
1970	May	130	180
1974	October	140	114
1975	June	110	94
1975	July	125	124



194 — Observed rainfall (mm) (140) — Estimated rainfall (mm) .152 — Rainfall station

Fig. 5 : Examples of the estimated missing records from the isopleth method.

The quality of the estimated rainfall records were examined using rainfall informations from the same locations for periods which had rainfall records. A random number table was used to select these periods. The periods which were chosen from the random number tables together with the rainfall estimates obtained from the isopleth and correlation methods for station 153 are given in Table 3B. The isopleths which were used to obtain the isopleth estimates given in Table 3B are shown in Figures 6a-6ℓ.

Table 3B indicates that the isopleth method gave better estimates of the rainfall values compared with those of the correlation method. The values of cumulative squared residual errors were about 3.55 and 111.56 for the isopleth and correlation methods repectively.

Table 3B. Observed and Estimated Rainfall at station No.153.

Year	Month	Observed Rainfall O(mm)	Rainfall Estimates					
			Isopleth Method			Correlation Method		
			Estimate E_A (mm)	$O-E_A$	$\frac{(O-E_A)^2}{E_A}$	Estimate E_B (mm)	$O-E_B$	$\frac{(O-E_B)^2}{E_B}$
1949	January	12	10	2	0.4	22	-10	4.5
	April	170	170	0	0	170	0	0
	July	149	150	-1	0.01	192	-43	9.63
	October	85	80	5	0.31	123	-38	11.74
1956	January	30	30	0	0	28	2	0.14
	April	167	170	-3	0.05	203	-36	6.38
	July	72	70	2	0.06	71	1	0.01
	October	139	140	-1	0.01	73	66	59.67
1968	January	0	0	0	0	13	-13	13.0
	April	158	160	-2	0.03	137	21	3.22
	July	89	75	14	2.61	74	15	3.04
	October	133	130	3	0.07	138	- 5	0.18
Totals					3.55			111.56

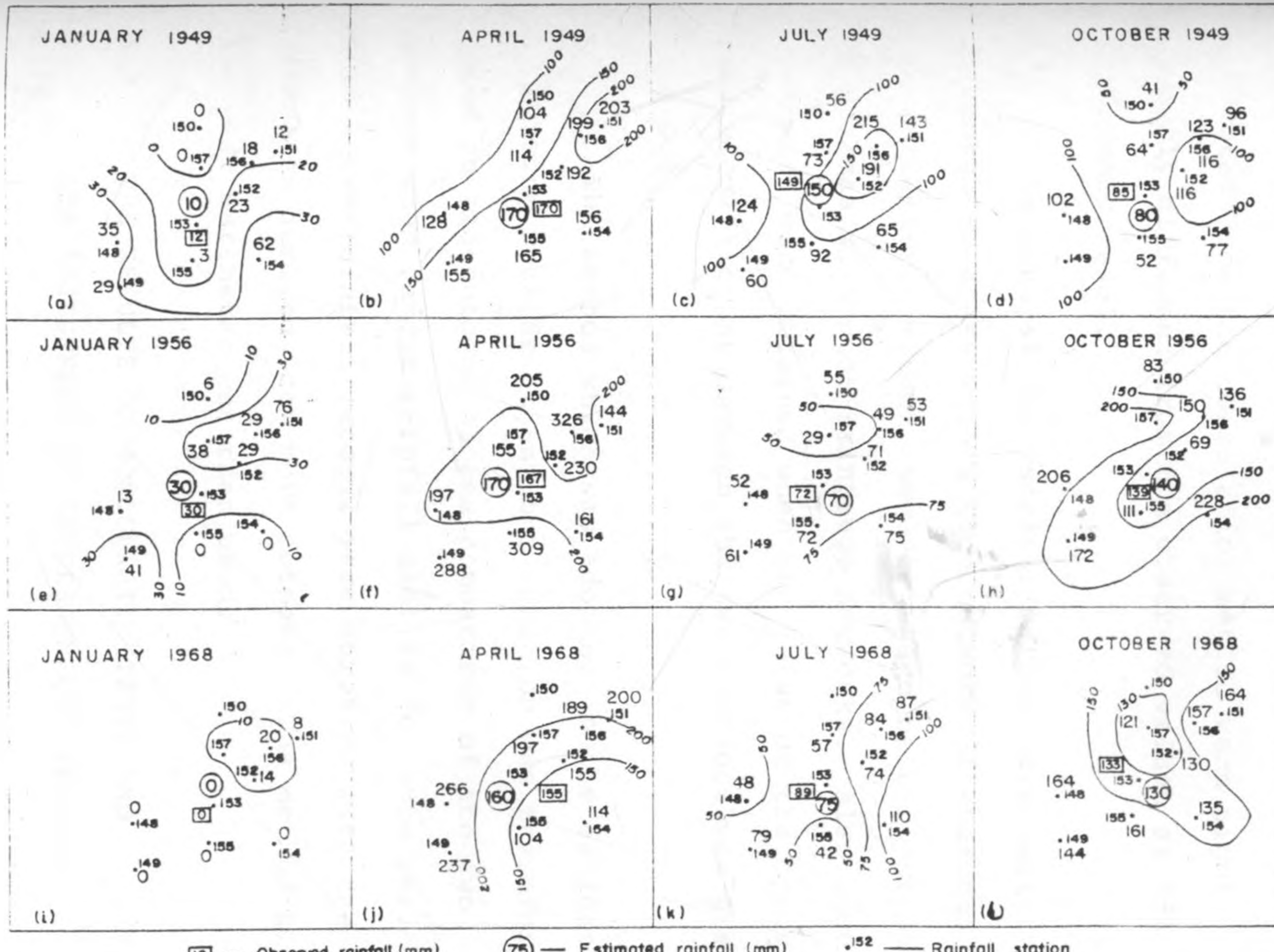


Fig. 6 : Examples of the estimated rainfall records for some selected periods at station 153 using the isopleth method

The isopleth estimates were significant at the 95% confidence level using the chi-square test. The chi-square test further indicated that the values estimated from the correlation method were significantly different from the observed values at 95% confidence level.

In general the isopleth method gave better estimates in areas with a good network of stations. The correlation method, however, gave better estimates over areas with poor raingauge networks. Also good estimates were obtained when the value of the correlation coefficient between the two stations was greater than 0.8.

The method which was adopted in the estimation of missing records at any one location was therefore based on the outcome of the comparison of the two methods when random rainfall samples for some periods with known rainfall records were compared with the estimates derived from both methods. The method giving the best estimates was then used.

3.2 THE RESULTS OF HOMOGENEITY TESTS AND THE ADJUSTMENT OF HETEROGENEOUS RECORDS.

Under this section, the results which were obtained from double mass curves analyses and Wald-Wolfwitz one sample runs test discussed under section 2.3.1 and 2.3.5 respectively are presented.

3.2.1 DOUBLE MASS CURVES RESULTS

Figures 7a to 7d give some examples of the double mass curves which were obtained with heterogeneous records. A single straight line cannot be fitted through the data points in all these graphs. This indicates that station 71 (Mukono) and Station 105 (Rwashamaire) records were heterogeneous. Altogether 8 out of 102 stations tested had heterogeneous records. These included Kembogo (STN.17), Lira (STN.57), Mukono (STN.71), Bakijjula (STN.83), Tororo (STN.99), Rwashamaire (STN.105), Adilang (STN.144) and Masindi (STN.151).

It is evident from Figures 7e and 7f that a single line could be drawn through the data points at all locations with homogeneous records.

The heterogeneous records were adjusted using gradient ratios of the fitted lines as discussed under section 2.3.1. Examples of the patterns of the double mass curves after adjustment of the heterogeneous records are given in Figures 7g and 7h using Mukono (STN.71) data.

3.2.2 WALD-WOLFOWITZ ONE SAMPLE RUNS TEST RESULTS

The Wald-Wolfwitz one sample runs test was used to examine the heterogeneity patterns of the rainfall records.

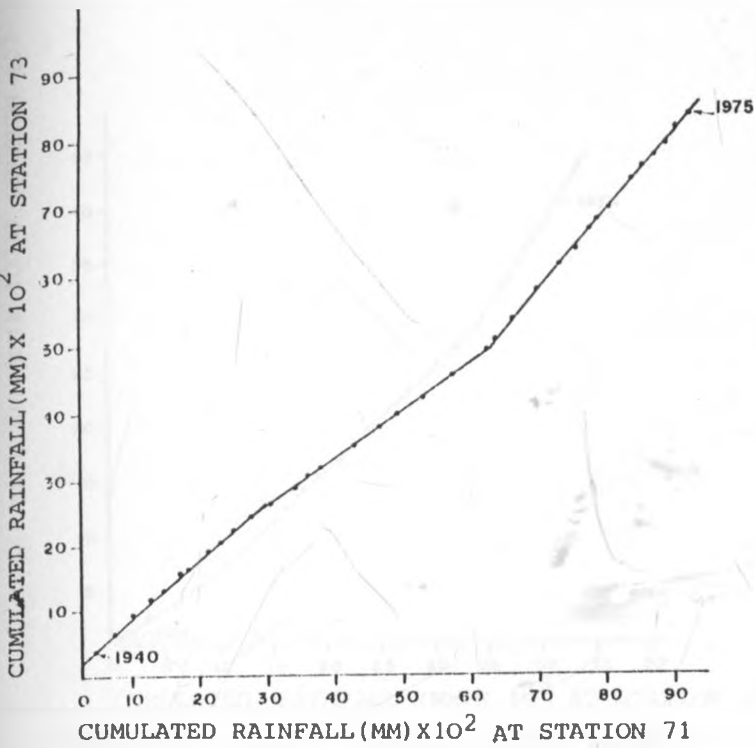


Fig. 7a: season 1 double mass curve for stations 71 and 73

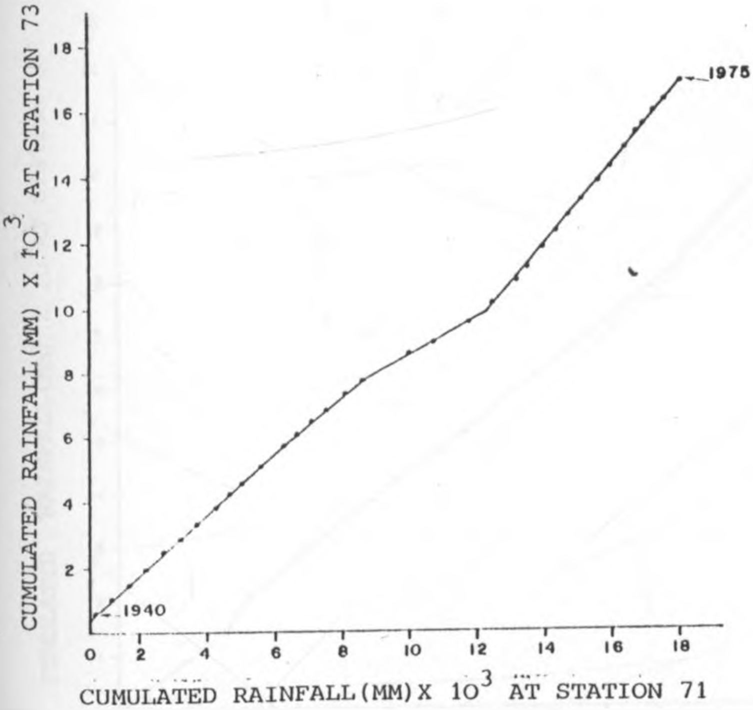


Fig. 7b: season 2 double mass curve for stations 71 and 73.

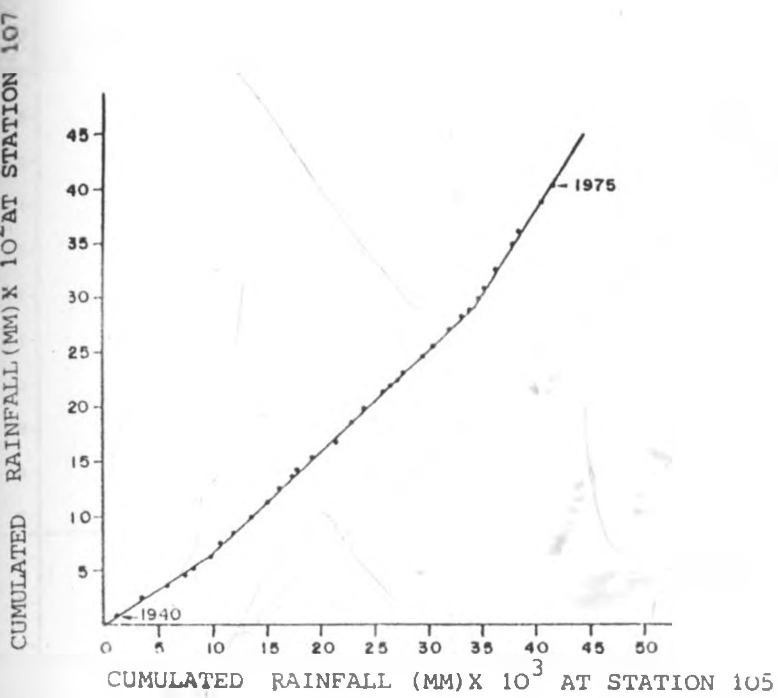


Fig 7c: Season 3 double mass curve for stations 105 and 107

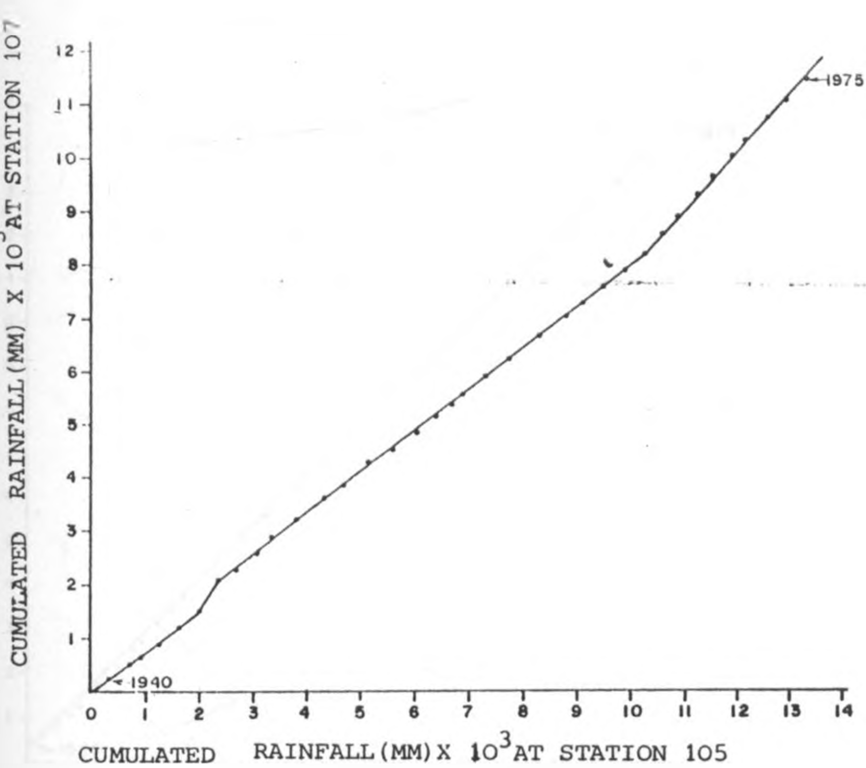


Fig. 7d: Season 4 double mass curve for stations 105 and 107

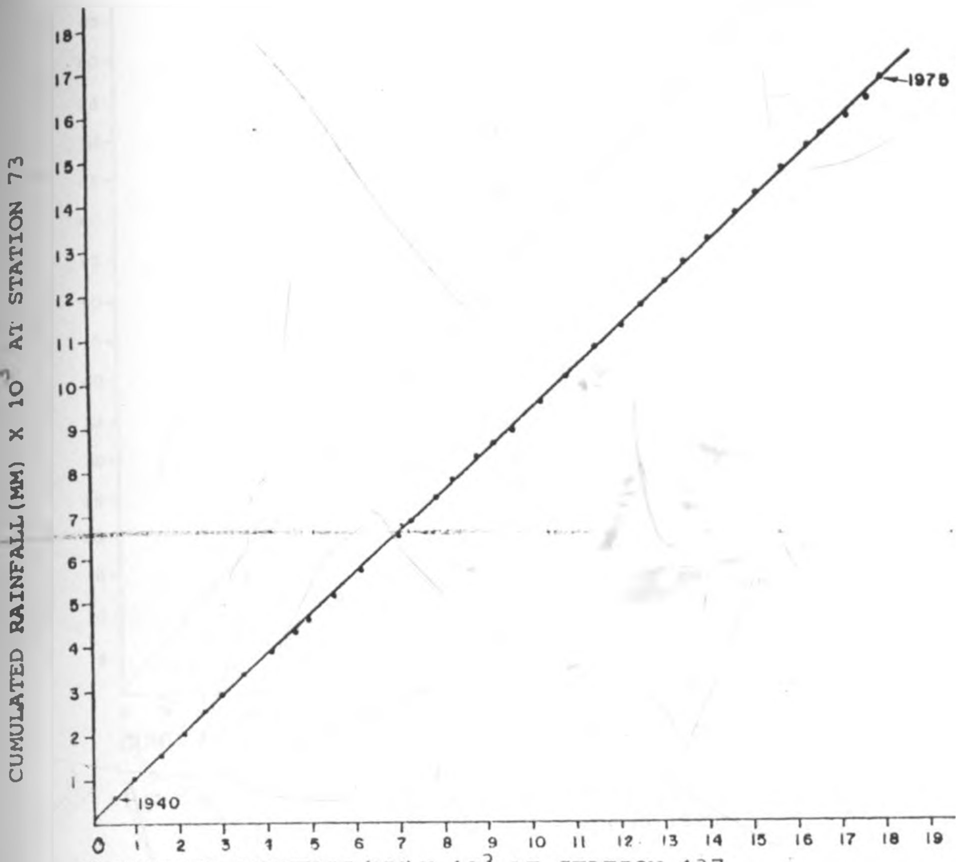


Fig. 7e : season 2 double curve for stations 137 and 73

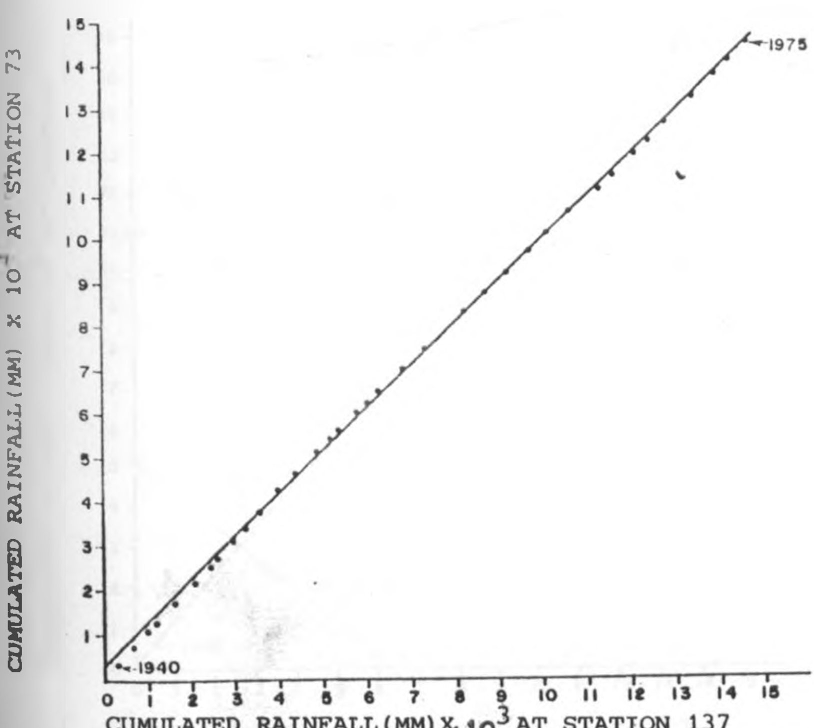


Fig. 7f: season 4 double mass curve for stations 137 and 73

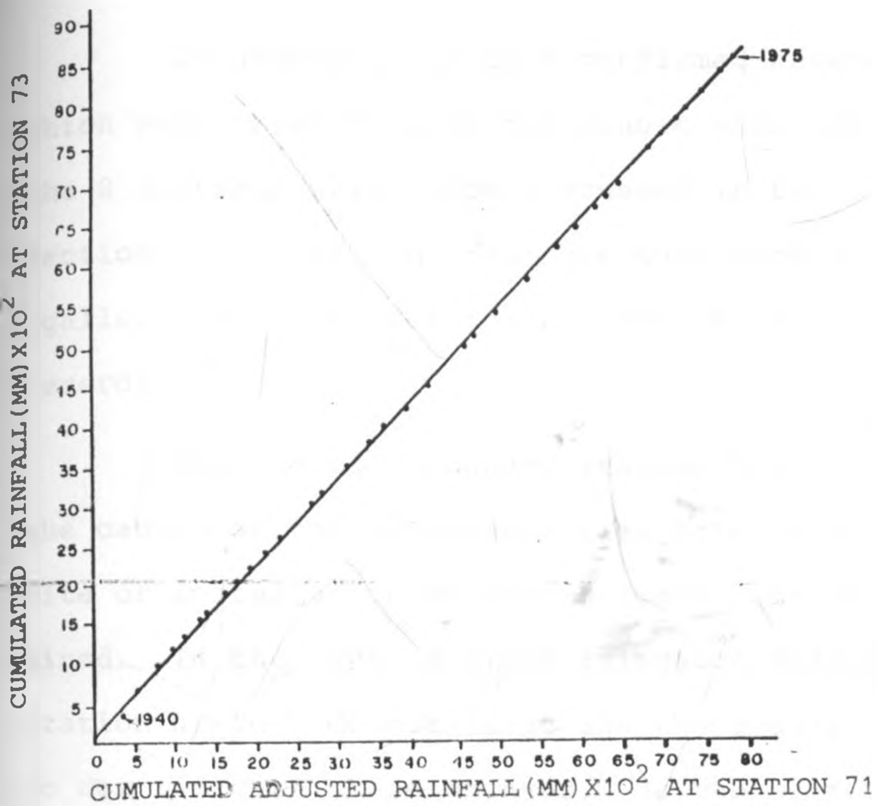


Fig. 7g : season 1 adjusted cumulated rainfall data for station 71 versus that of station 73

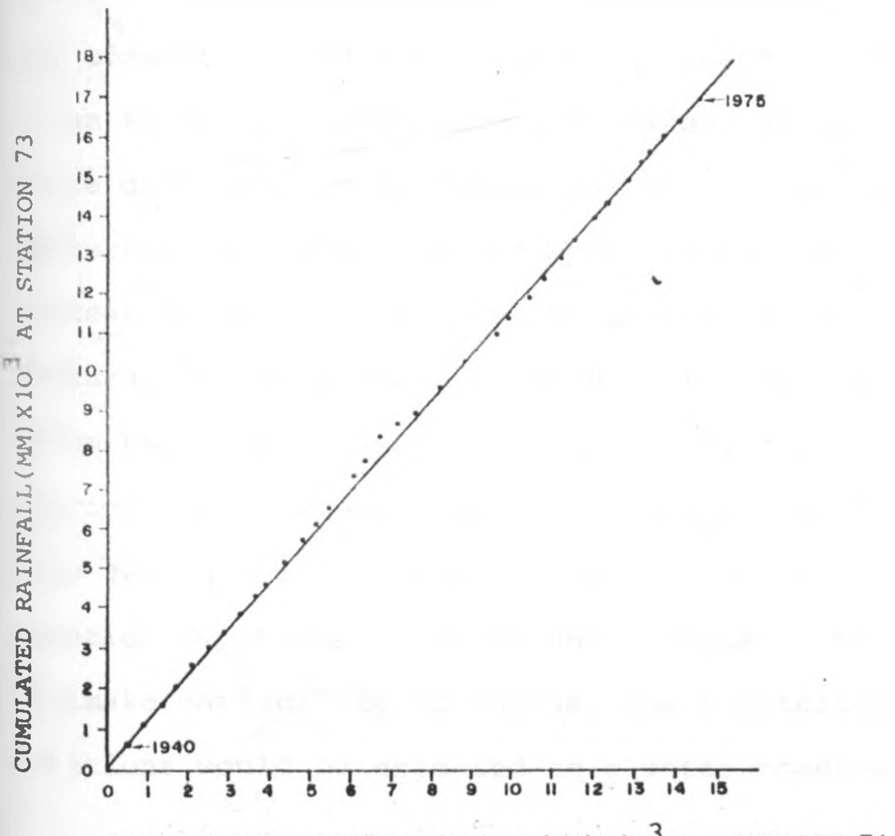


Fig. 7h : season 2 adjusted cumulated rainfall data for station 71 versus that of station 73

In general, the test confirmed heterogeneity which were observed with the double mass curves at the 8 stations which were discussed in the previous section. Similar findings have been reported by Ogallo, (1981) for Tororo (STN.99) annual rainfall records.

Where good raingauge station history exist, the causes of the heterogeneities like change of site or installation of new equipment can be determined. In the case of Uganda raingauge stations, good station history do not exist and the available records do not indicate these changes. At the commencement of these studies, only one inspection report which did not, unfortunately, cover the entire raingauge network had been carried out in the early 1970's. No comments on change of site or equipment other than their serviceability were made. It was therefore difficult to determine the actual causes of the heterogeneity which was observed in the data. The causes therefore could not be linked to any regional natural climatic change. This is further supported from the double mass curves which indicate different periods for the onset of the heterogeneous records. The spatial distribution of the 8 heterogeneous station locations is at random. Normally for any natural climate variability or change, the 8 heterogeneous stations would be expected to cluster together.

The common factors of the stations with heterogeneous records are that they were either small towns and trading centres which have grown since into big towns during the late forties, fifties and sixties or they were farm locations. It is perceivable that in the cases of expanding towns like Lira (STN.57), Tororo (STN.99) and Masindi (STN.151), other developmental demands may have necessitated the movement of these gauges to other sites or changed the exposure patterns of the gauges. In the case of farm locations like Mukono (STN.71) and Kembogo (STN.17), which are not run by the Meteorological Department, it is possible that the exposure of the gauges and/or their locations could have been changed due to changing land use activities and other demands within the farms.

The records declared to be homogeneous formed the fundamental base for all future analyses and discussions. The first part of these discussions will be concentrated on the areal rainfall estimates.

3.3 AREAL RAINFALL ESTIMATES

Areal rainfall estimates are better indicators of climatic fluctuations and changes than point values. Under this section an attempt was made to compare the results of the areal rainfall estimates derived from the arithmetic mean and PCA factor weighting methods and for various homogeneous regions. Details of the derivation of the homogeneous regions used here are, however, given later in the next.

Table 4 gives a summary of the chi-square values which were obtained when areal rainfall estimates from the two methods were compared for the years 1940-75 inclusive over various periods.

Table 4: The Chi-Square values between Arithmetic mean values and PCA Factor weighting averages for various homogeneous regions

Zone	No. of Stations	Chi-square values					Comments
		Dec. - Feb. Season 1	Mar. - May Season 2	Jun. - Aug. Season 3	Sept. - Nov. Season 4	Jan. - Dec. Mean Annual	
A1	9	0.48	0.57	0.83	0.34	1.35	
A2	1	None	None	None	None	None	No comparisons made due to only one station
B	19	0.67	1.82	1.75	1.48	2.23	
C	6	9.38	16.56	11.14	9.09	15.68	
D	7	0.31	0.20	0.16	0.44	0.24	
E	8	0.99	0.82	0.97	1.31	1.68	
F	6	0.14	0.14	0.13	0.12	0.15	
G	5	0.49	0.92	0.72	0.69	1.63	
H	9	0.12	0.10	0.15	0.08	0.18	
I	11	0.33	0.41	1.42	0.46	0.82	
J	5	0.21	0.24	0.32	0.31	0.42	
K	3	None	None	None	None	None	No comparisons due to too few stations
L	10	0.34	0.40	0.24	0.50	0.76	
M	6	4.05	2.47	5.46	3.99	4.70	

$$\chi^2_{0.05,34} = 21.7$$

The chi-square test indicated that at 95% level of confidence, there were no significant differences between the PCA factor weighting and the arithmetic mean areal rainfall averages for the 12 homogeneous regions compared. The other two homogeneous regions had very few stations hence the two methods estimates could not be compared. The close agreement between the two methods may be indicative of the similarities in the rainfall characteristics at the stations within the individual homogeneous regions.

In the next sections, the results of the rain-gauge network designs are presented.

3.4 RESULTS FROM RAINGAUGE NETWORK DESIGN

The results from the minimum and optimum rain-gauge network designs are discussed separately in the following sections.

3.4.1 RESULTS FROM THE MINIMUM RAINGAUGE NETWORK DESIGN.

It was indicated under section 2.4.1 that the concept of a minimum raingauge network is to provide a set of raingauges in the national network which represent all the climatic divisions of the country. These were derived from the spatial mapping of the dominant PCA modes. The characteristics of the PCA modes are therefore presented before they are used to design the minimum network.

3.4.1.1 RESULTS FROM PRINCIPAL COMPOENT ANALYSIS (PCA)

Monthly and seasonal rainfall records were used in PCA analyses to map the spatial patterns of the rainfall over Uganda. The characteristics of the significant PCA modes are highlighted in the following sections. The parameters which were used to describe the time and space patterns of the PCA results included the significance of the generated PCA modes obtained from the rainfall records.

3.4.1.2 SIGNIFICANT FACTORS

Table 5 gives an example of the results which were obtained when climatological records were subjected to PCA. The results show that with the Kaiser criterion of retaining only eigenvalues greater or equal to 1, only three eigenvectors were significant.

Table 5: Example of PCA Results with the mean climatological rainfall records.

FACTOR	EIGENVALUE	% OF VARIANCE	COMULATIVE %
1	19.79	43.0	43.0
2	6.41	13.9	57.0
3	2.61	5.7	62.6
4	0.94	2.0	64.7
5	0.87	1.9	66.6
6	0.84	1.8	68.4
7	0.69	1.5	69.9
8	0.68	1.5	71.4
10	0.61	1.4	74.1

Table 6: Example of significance of sampling errors of the eigenvalues for thé mean climatological rainfall records

FACTOR	EIGENVALUE (λ_i)	SAMPLING ERROR ($\lambda_i(^2/N)^{1/2}$)	$\lambda_{i-1}-\lambda_i$
1	19.79	1.33	
2	6.41	0.44	13.80
3	2.61	0.18	3.80
4	0.94	0.06	1.67<
5	0.84	0.04	0.03

Cut off point<

Results from the scree and logarithm of eigenvalue (LEV) tests are given in Figures 8a and 8b respectively while those from the sampling error test are given in Table 6.

Figures 8a and 8b indicate that four factors were extracted by both the scree and LEV methods, while Table 6 shows that the magnitude of $(\lambda_i (2/N)^{\frac{1}{2}})$ was much less than the magnitude of $(\lambda_{i-1} - \lambda_i)$ for factors 1 to 4 whereas the magnitude of $(\lambda_i (2/N)^{\frac{1}{2}})$ $\approx$ the magnitude of $(\lambda_{i-1} - \lambda_i)$ at factor 5 indicating that the sampling errors test also extracted 4 factors. The largest number of factors derived from the four methods were used in the varimax and oblique rotations, for example, 5 factors were rotated with the climatological records.

Table 5 indicates that with the climatological records, the variance accounted for by the first three factors were about 43%, 14% and 6% respectively. Factor 4, however, accounted for only 2%. The four significant factors therefore accounted for about 65% of the total spatial variance of rainfall over Uganda.

On subjecting the monthly and seasonal rainfall records to PCA, it was noted that the number of significant factors varied from month to month and season to season ranging from 9 to 16. The details of these significant factors are given in Table 7.

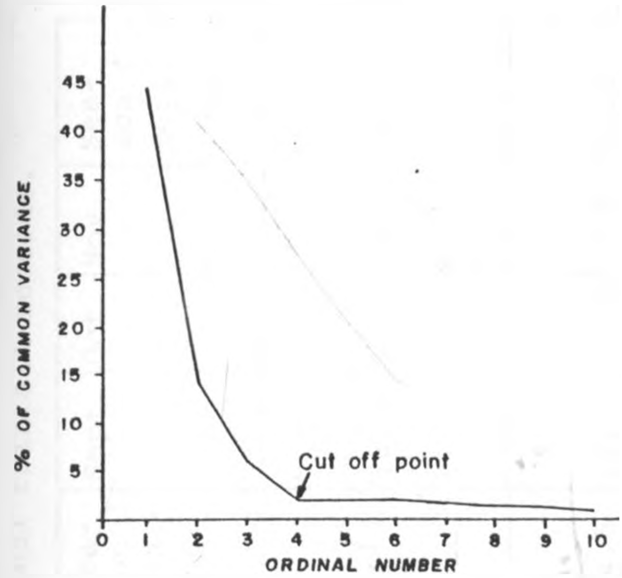


Fig. 8a. Example of the results from the scree test

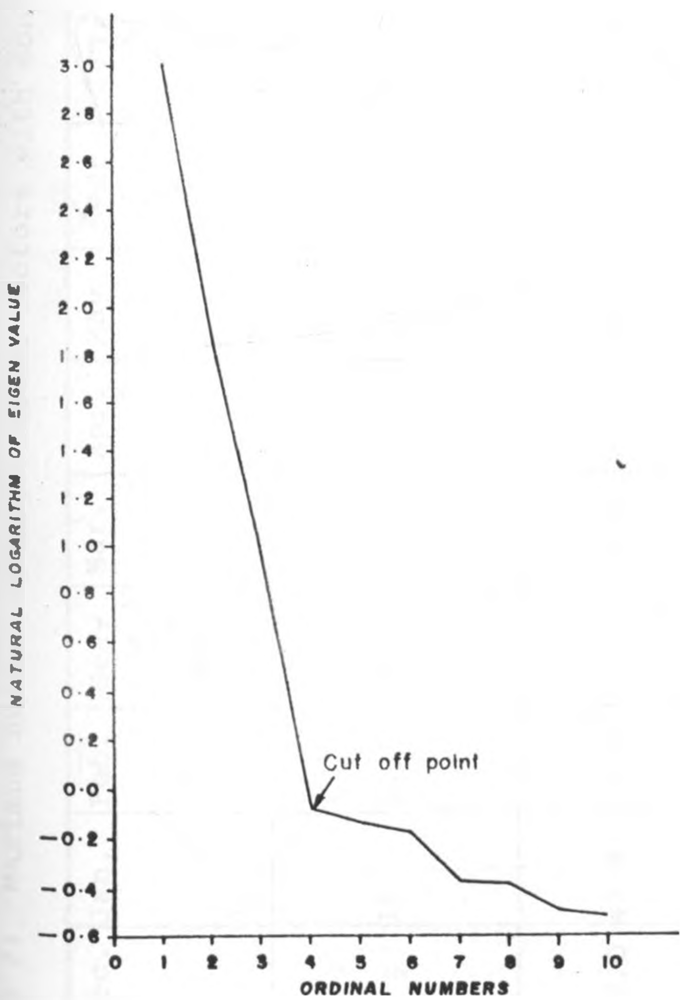


Fig 8b : Example of the results from LEV test

Table 7: **Maximum** number of significant factors with monthly **and** seasonal rainfall records.

	Dec.	Jan.	Feb.	Season 1 Dec.- Feb.	Mar.	Apr.	May	Season 2 Mar- May	Jun.	Jul.	Aug.	Season 3 Jun.- Aug.	Sept.	Oct.	Nov.	Season 4 Sept.- Nov.
Significant number of factors	9	10	11	10	13	15	14	14	14	14	15	14	16	12	9	10
Percentage variance due to factor 1	53.0	43.9	38.8	49.4	35.5	24.8	20.1	26.7	23.6	28.3	23.0	24.5	24.6	32.6	57.6	44.0
Percentage variance accounted for by signifinicant factors	82.1	82.6	81.9	82.1	83.0	86.5	82.8	83.1	82.6	83.4	85.3	82.5	86.5	81.6	83.0	80.7

The above values are as per Kaiser (1959) criterion.

Table 7 shows that the number of significant principal components extracted for seasons 1,2,3 and 4 were 10,14,14 and 10 respectively. The variance explained by the first principal components were 49.4%, 26.7%, 24.5% and 44% respectively in seasons 1,2,3 and 4. The significant PCA modes accounted for over 81% of the rainfall variances during all seasons. The largest number of significant factors and the lowest values of variances extracted by the first significant factors were observed during seasons 2 and 3. This may be due to the large spatial variances in the regional rainfall characteristics during these seasons. Large spatial variations in the rainfall characteristics during the two seasons have also been observed by Ogallo (1989).

Relatively fewer factors were significant during the seasons 1 and 4. The variances extracted by the first factors were also relatively higher during these seasons. This could be associated with the predominance of induced regional systems like land/sea breezes and mountain/valley influences in season 1 and the ITCZ in season 4.

Table 7 further indicates that the number of significant factors extracted from the monthly records were relatively higher for some months than the seasonal values. This may be due to the filtering out of some of the localised factors by the seasonally averaged rainfall records. However, the dominant modes of the monthly records gave similar results to those derived from seasonal and climatological data.

The spatial patterns of the significant PCA modes were used to delineate Uganda into homogeneous rainfall zones. The next section gives the spatial characteristics of the dominant PCA modes together with their use in the delineation of homogeneous climatic zones.

3.4.1.3 THE DELINEATED HOMOGENEOUS ZONES OF UGANDA

Figure 9 gives the homogeneous rainfall zones of Uganda which were delineated based on the spatial patterns of the dominant PCA modes.

The statistical significance of any factor loading at any location was determined using its standard error (E_a) given by equation (31). It was observed that any factor loading $\geq$ 0.13 was significant for climatological rainfall records. A value of a factor loading $\geq$ 0.3 for seasonal/monthly rainfall records was significant when a 5% significance level was used. Thus, threshold values of 0.13 and 0.3 were used for climatological and seasonal/monthly rainfall records respectively.

Examples of the regression equations which were

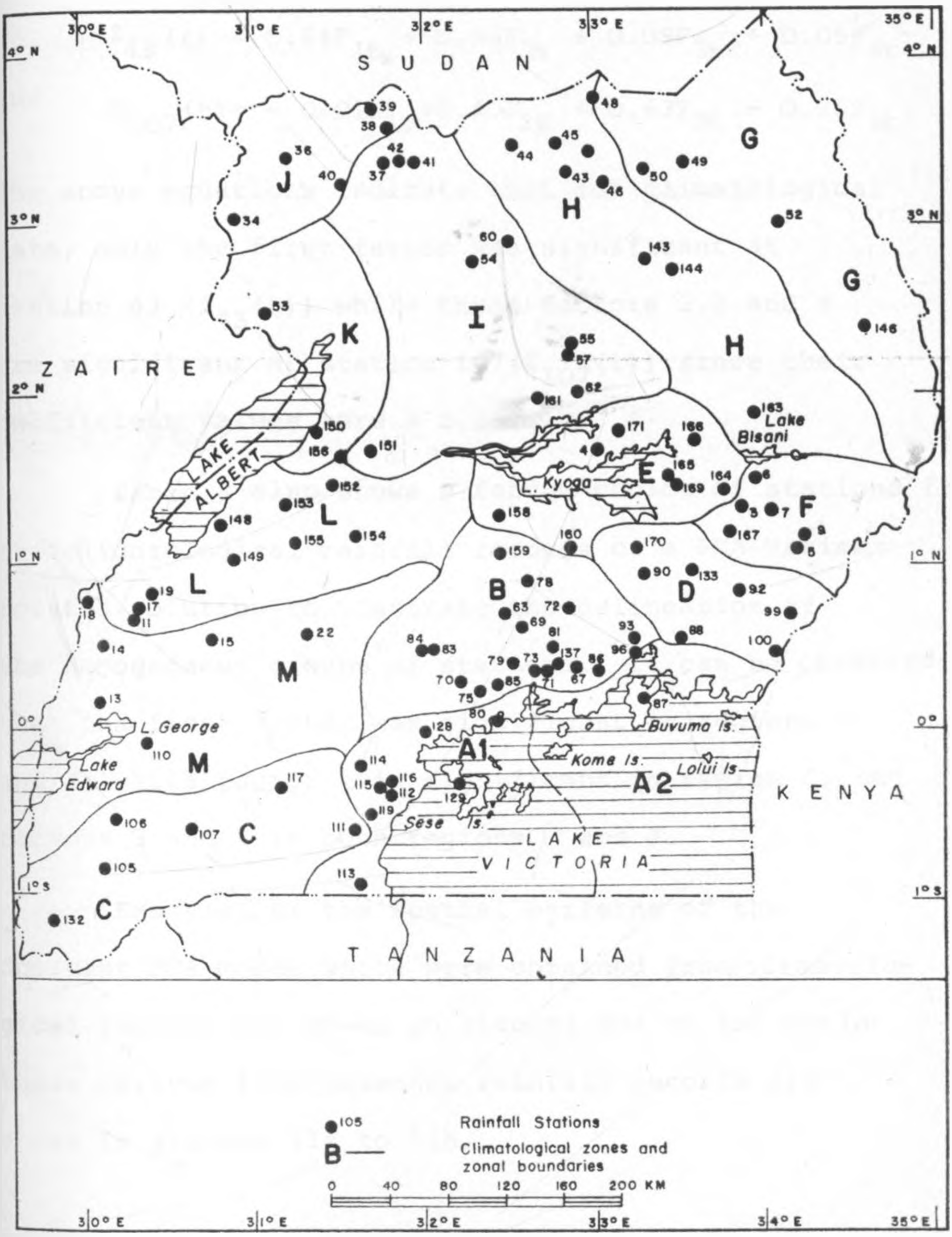


Fig. 9 : The delineated climatological zones of Uganda from PCA results.

obtained from the rotated PCA solution of the climatological data shown in Table 8 can be given by:-

$$Z_{43}(t) = 0.84F_{1t} + 0.06F_{2t} + 0.09F_{3t} + 0.05F_{4t}$$

and

$$Z_{107}(t) = - 0.01F_{1t} + 0.40F_{2t} + 0.63F_{3t} - 0.25F_{4t}.$$

The above equations indicate that for climatological data, only the first factor was significant at station 43 ($Z_{43}(t)$) while three factors 2,3 and 4 are significant at station 107($Z_{107}(t)$) since their coefficient values were $\geq$ 0.13.

Table 8 also shows a random choice of stations from the climatological rainfall records of a PCA varimax rotated solution to ilustrate the delineation of the homogeneous groups of stations. It can be observed that the first factor was significant in regions H and J; while factor 2 is significant in region C; and factors 3 and 4 in both regions C and J.

Examples of the spatial patterns of the dominant PCA modes which were obtained from climatological records are given in Figures 10a to 10d while those derived from seasonal rainfall records are shown in Figures 11a to 11h.

Table 8. Loadings of the four major factors for Regions C, H and J based on climatological records.

REGION	STATION CODE NUMBER	FACTOR LOADINGS			
		FACTOR (1)	FACTOR (2)	FACTOR (3)	FACTOR (4)
C	107	- 0.01	0.40	0.63	- 0.25
	117	- 0.01	0.33	0.61	- 0.27
	132	- 0.12	0.50	0.47	- 0.22
H	43	0.84	0.06	0.09	0.05
	48	0.73	0.01	0.05	0.03
	143	0.84	0.07	0.05	0.10
J	34	0.67	0.10	0.43	0.14
	36	0.72	0.09	0.32	0.19
	39	0.73	0.08	0.29	0.21

In Figure 10b, factor 2 from the climatological data was dominant over the lake victoria region (regions A1 and A2) were the factor regression coefficients were greater than 0.7. The regression coefficients on factor 2 decreased inland from the centre of the regions. Results from the seasonal records, however, indicated some variations in the rainfall characteristics over western and north eastern shores of Lake Victoria. Figures 11a and 11e indicate that one factor was dominant over the western shores of Lake Victoria while more than two factors dominated over the northern to

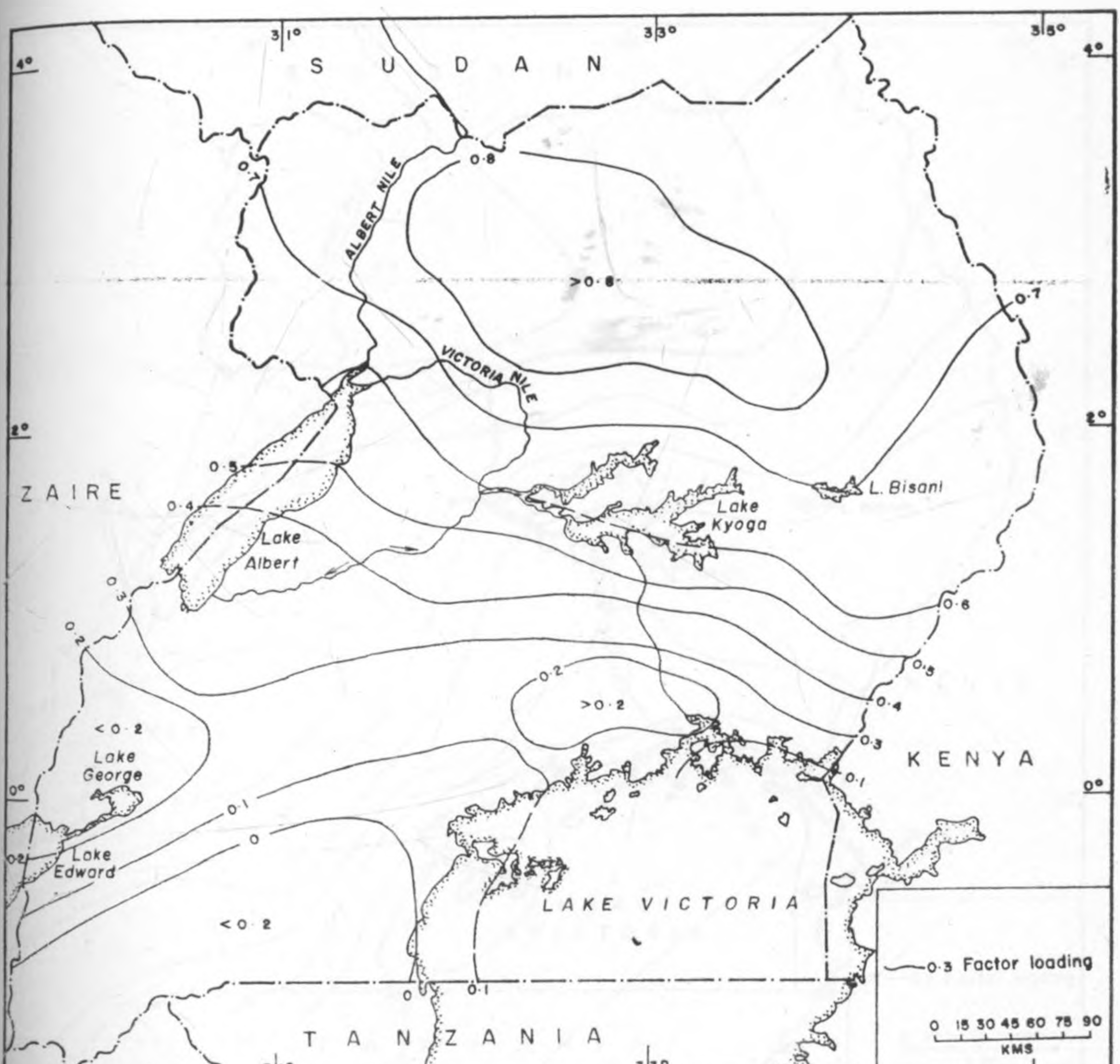


Fig. 10a: Map patterns of factor 1 loadings for climatological records.

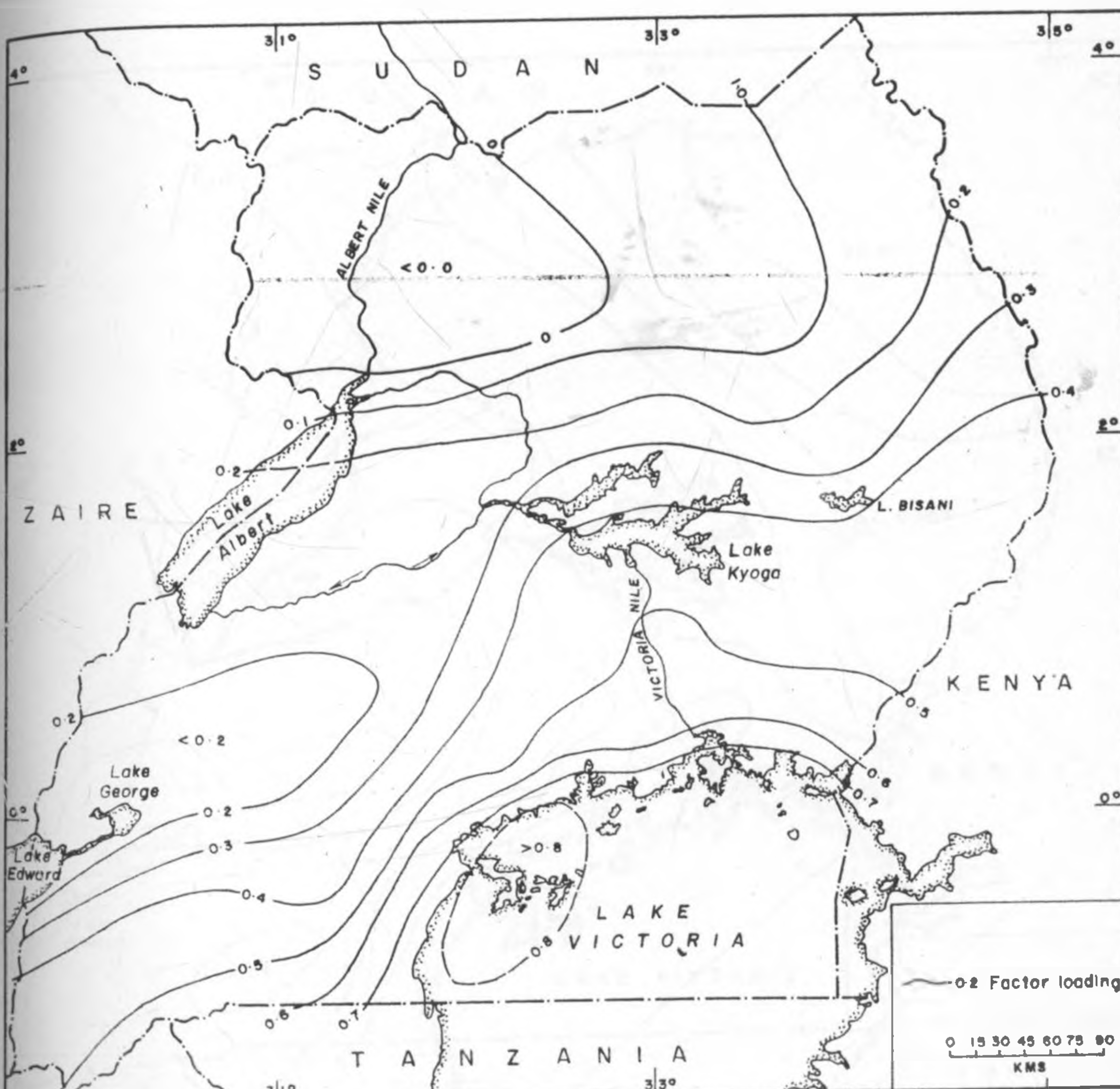


Fig 10b: Map patterns of factor 2 loadings for climatological records.

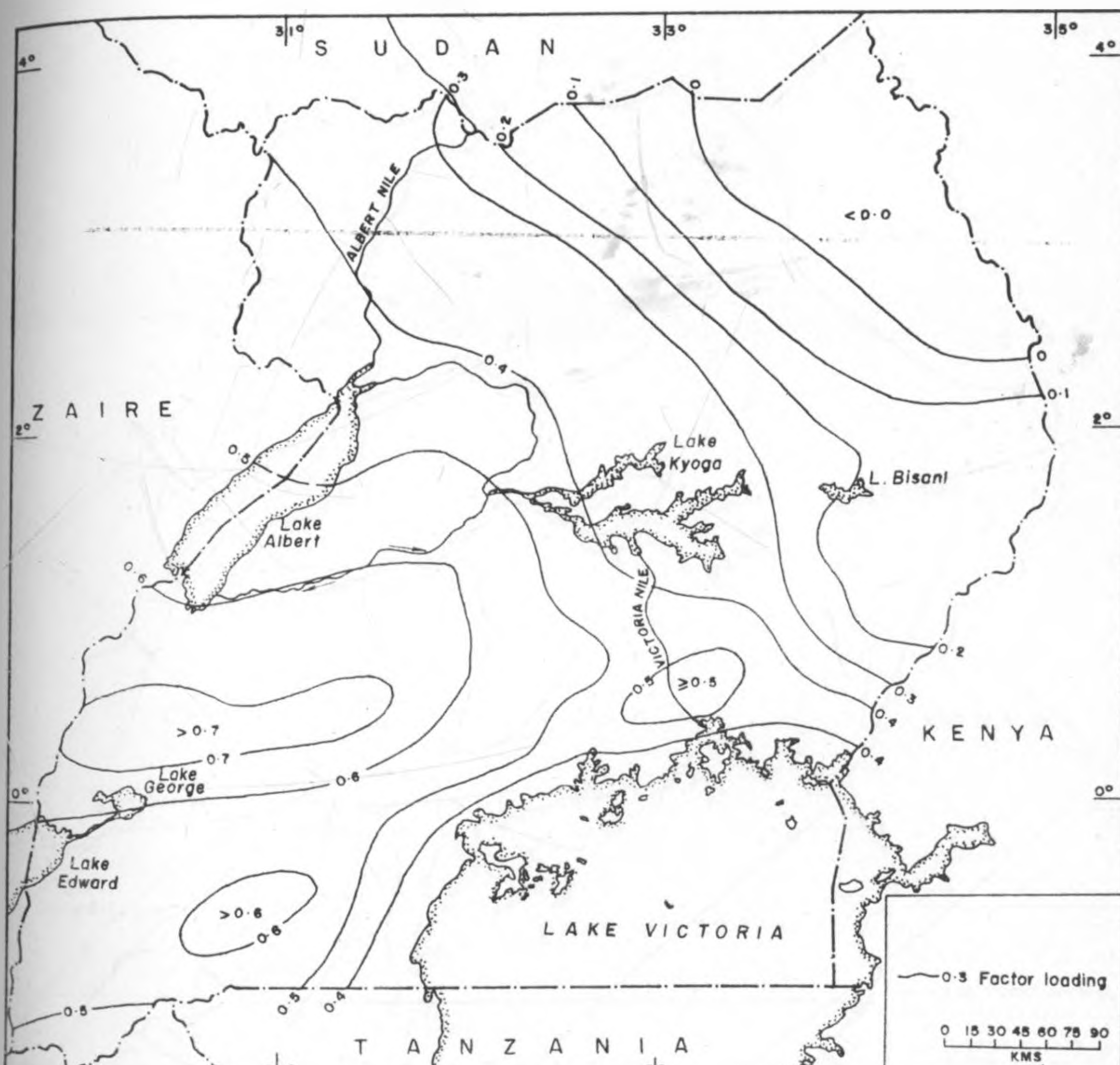


Fig 10c: Map patterns of factor 3 loadings for climatological records.

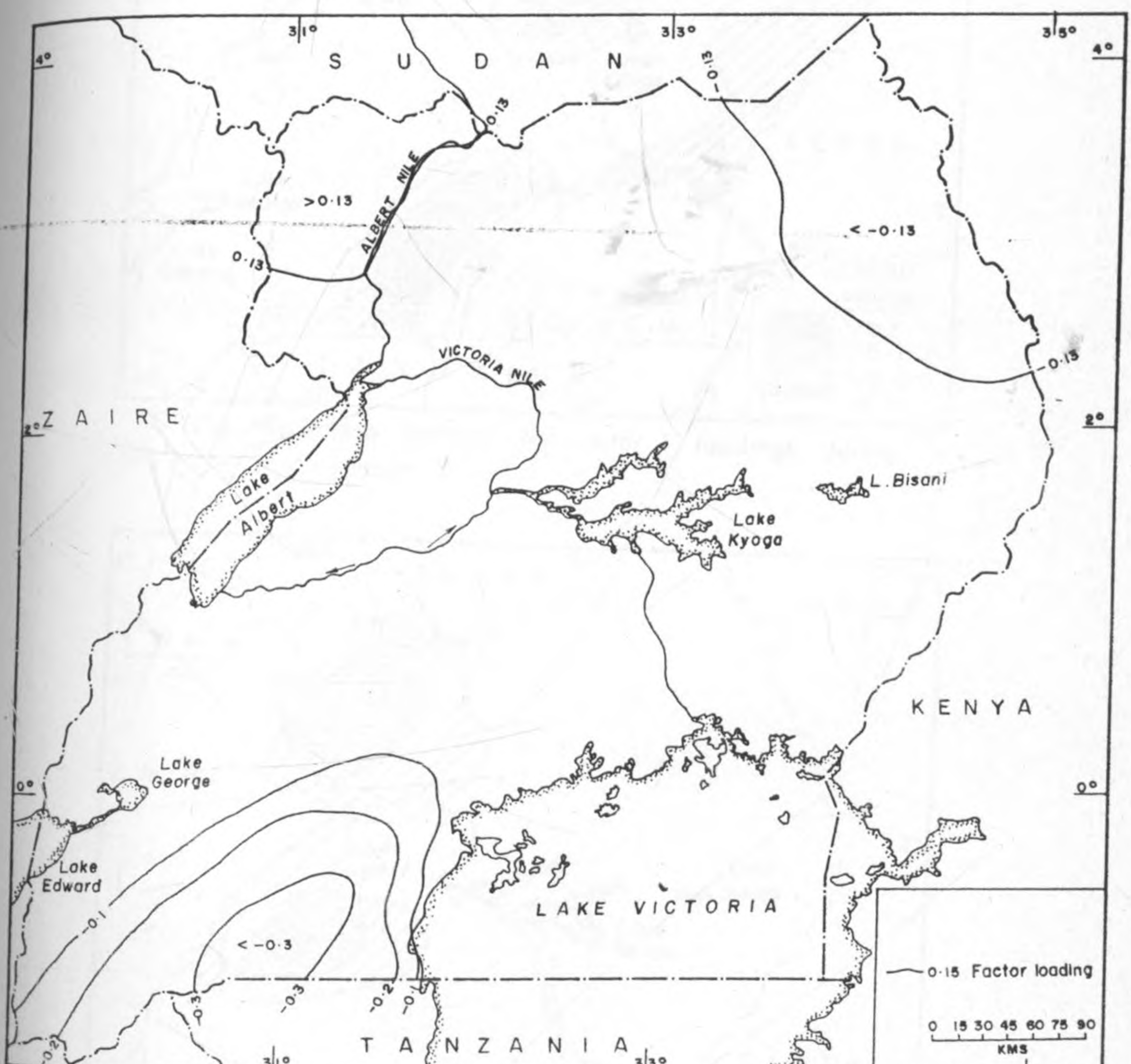


Fig 10d: Map Patterns of factor 4 loadings for climatological records.

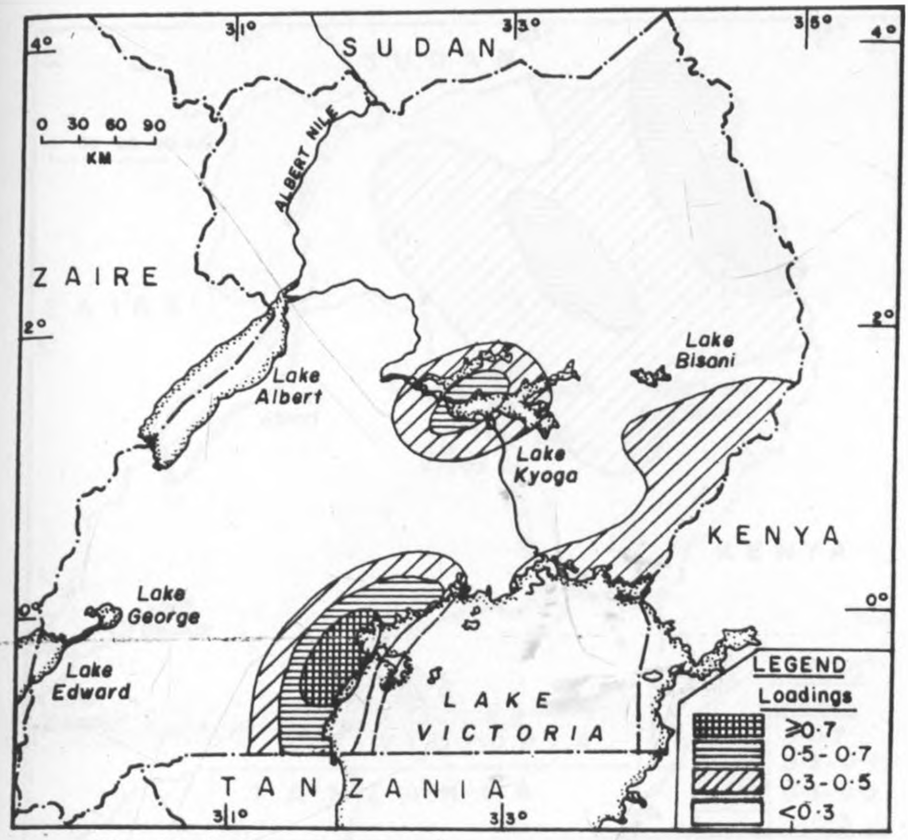


Fig. 11a: Map patterns for factor 1 loadings during season 2.

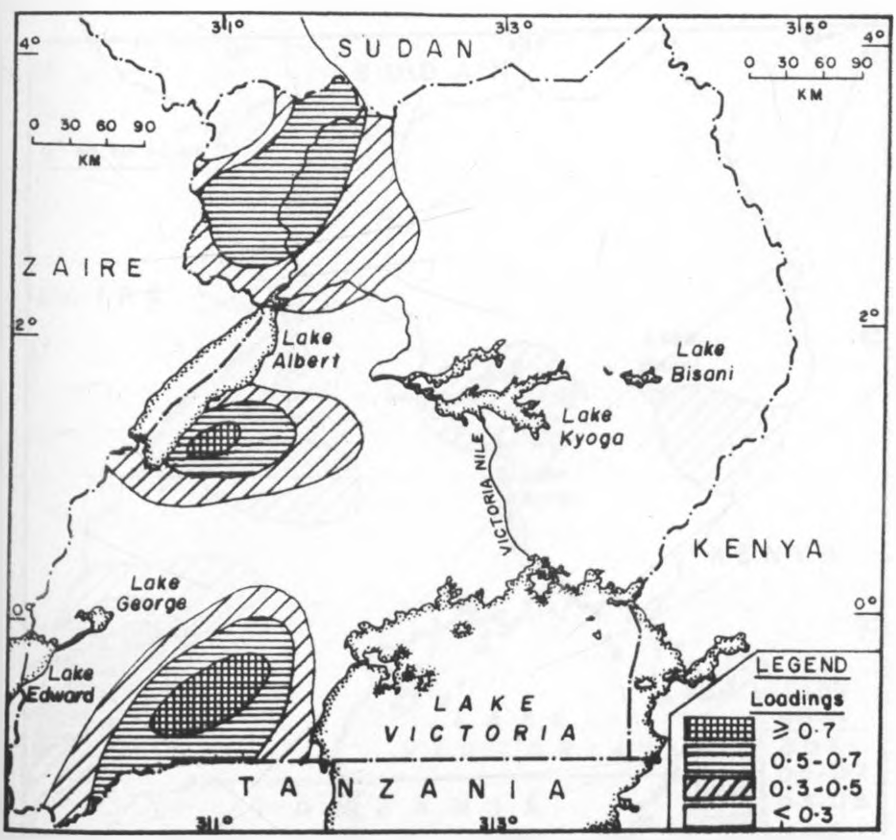


Fig. 11b: Map patterns for factor 2 loadings during season 2.

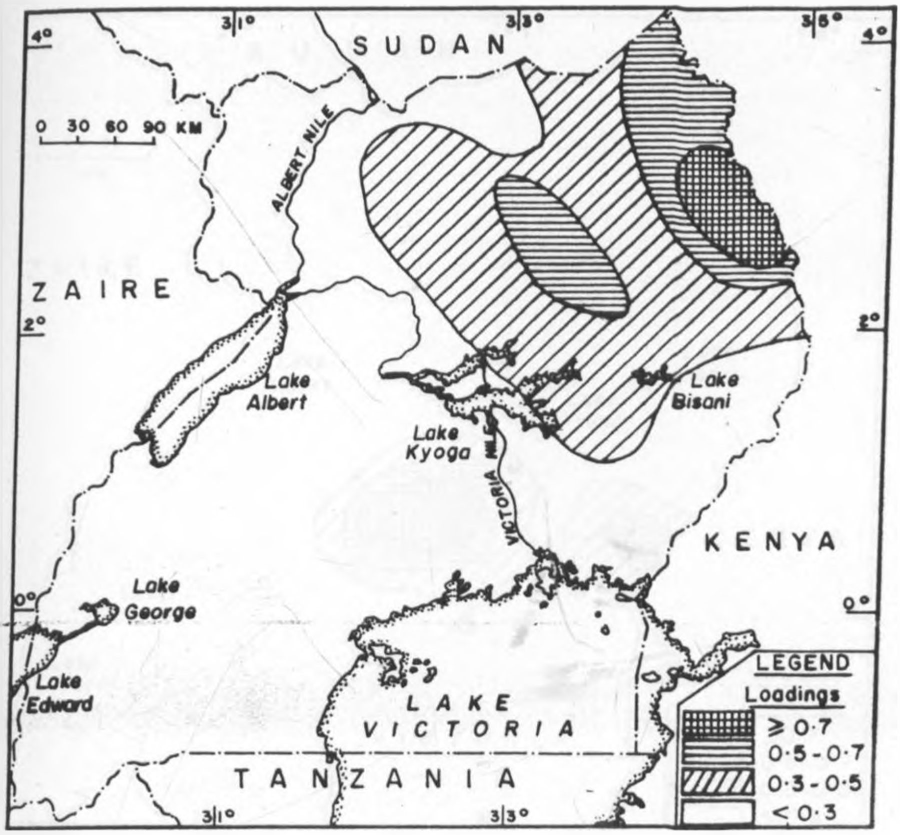


Fig. 11c: Map patterns for factor 3 loadings during season 2.

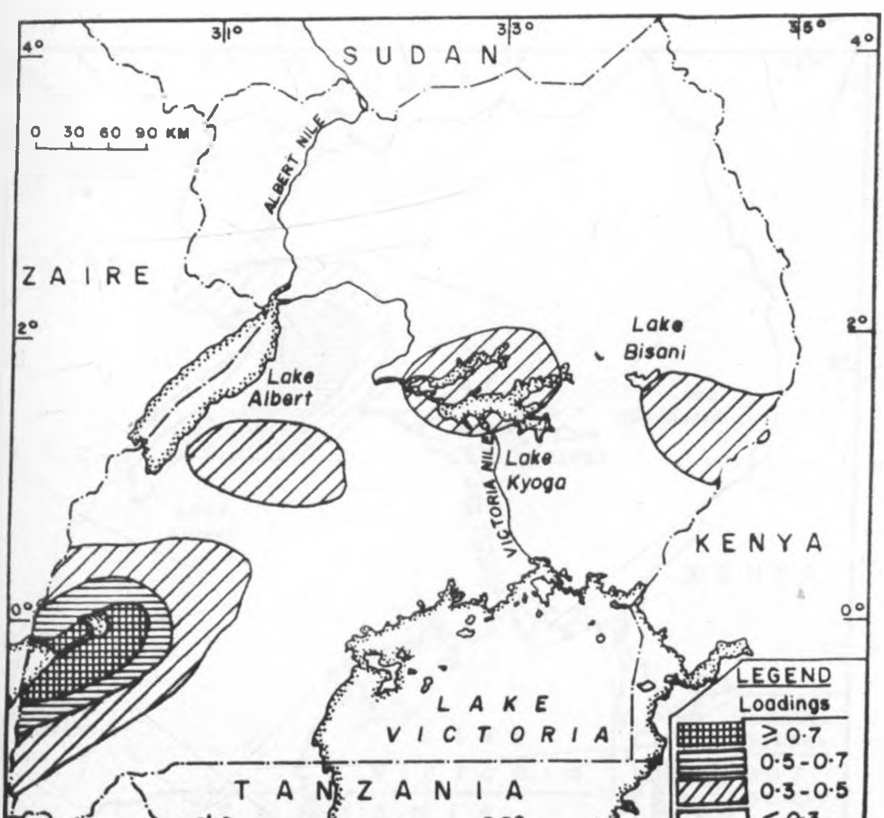


Fig. 11d: Map patterns for factor 4 loadings during season 2.

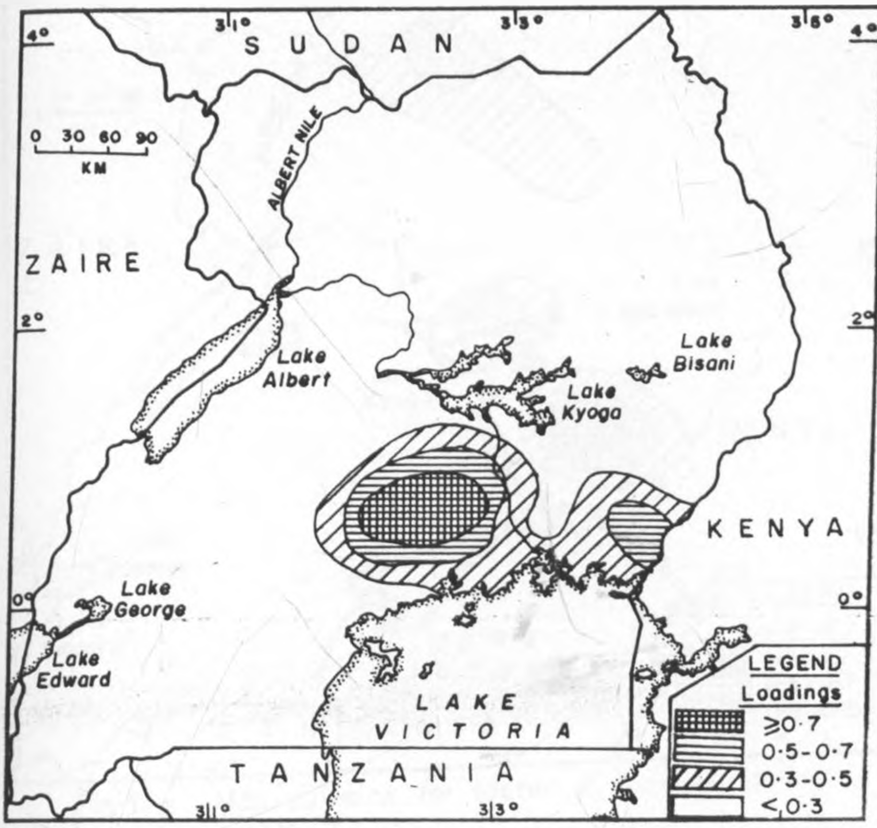


Fig. 11e : Map patterns for factor 9 loadings during season 2

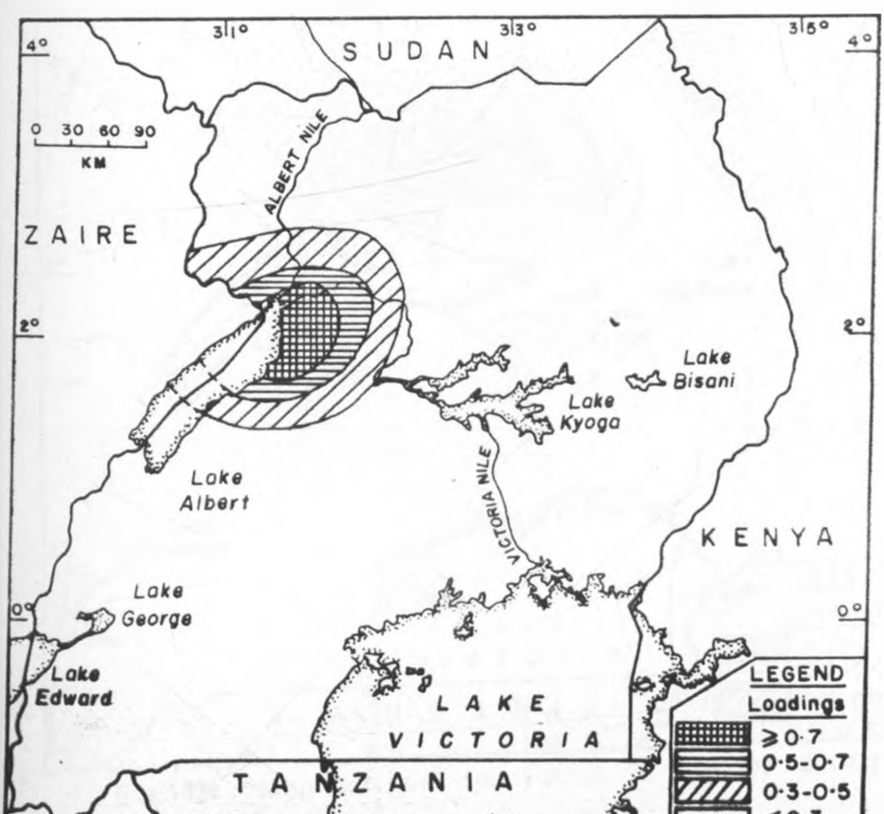


Fig. 11f : Map patterns for factor 11 loadings during season 2 .

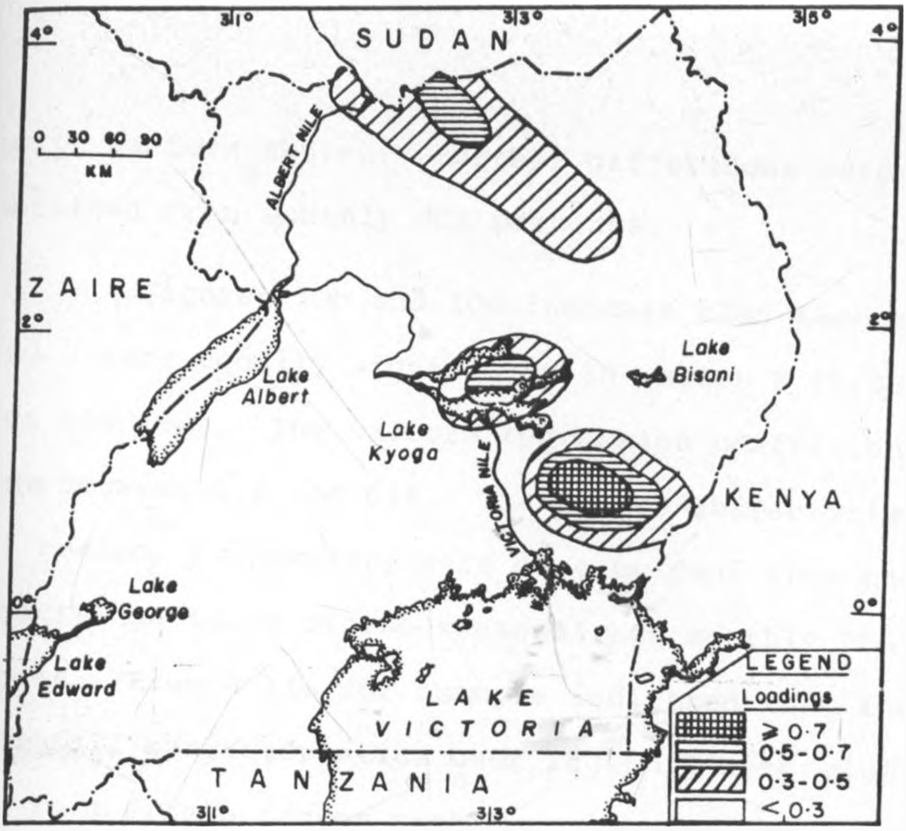


Fig. 11g : Map patterns for factor 2 loadings during season 3

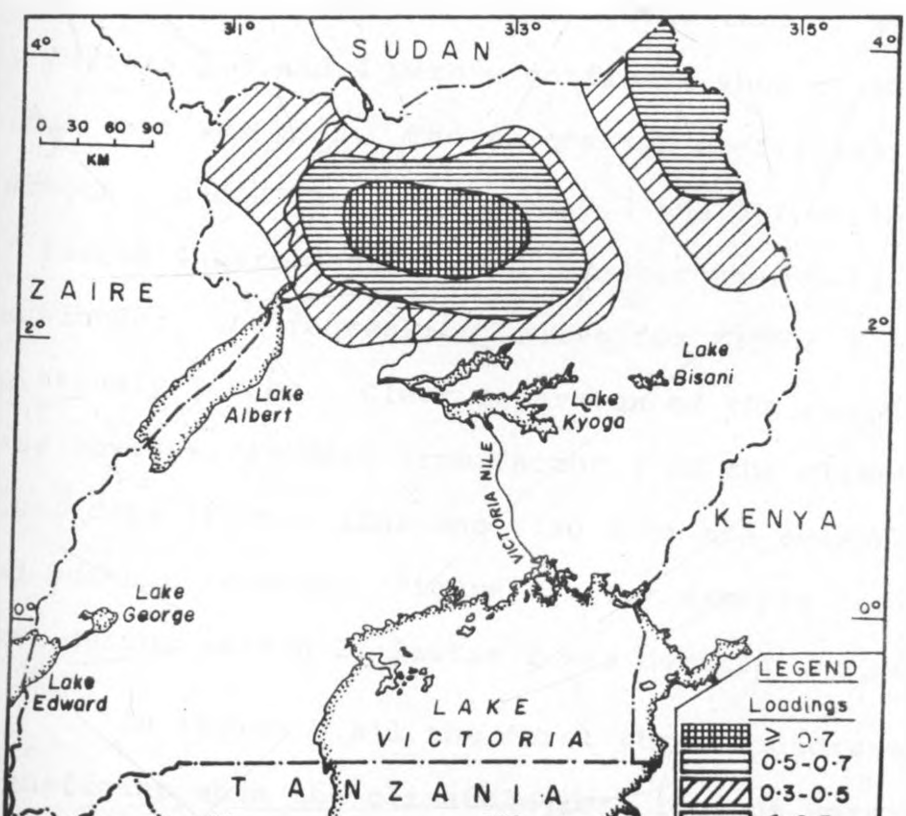


Fig. 11h : Map patterns for factor 3 loadings during season 3 .

north eastern shores. Further differences were also obtained from monthly PCA patterns.

Figures 10b and 10c indicate that factors 2 and 3 were equally significant in region B throughout the year. The factors regression coefficients were between 0.4 and 0.6. The unique characteristics of region, B, however, were more evident from the spatial patterns of the seasonal and monthly PCA modes. Figure 11e for example indicated that the rainfall characteristics over region B were clustered on to a single factor with a regression coefficient greater than 0.7.

Figures 10b, 10c and 10d show that in region C, factors 2,3,and 4 were significant when climatological data was used. The regression coefficient for factors 2 and 3 ranged between 0.3 and 0.7 while those of factor 4 were negative but greater than 0.13 in magnitude. The regression values for factor 3 were relatively larger. Clear separaton of the region C was, however, evident from factor 4 of the climatological data (Figure 10d) and also from the seasonal and monthly records. Figure 11b for exmaple indicated that during season 2, factor 2 was dominant in region C.

In region D all the first three factors were significant when the climatological records were used (Figures 10a, 10b and 10c). the regression coefficients being more than 0.3 in magnitude. The delineation of of zone D was more clear from the seasonal and monthly records. An example of these is given in Figure 11g.

In region E factors 1,2 and 3 were significant with the climatological records, but the first factor was more dominant. The similarity of rainfall characteristics within this region was also more evident with the seasonal and monthly records (Figures 11a and 11d).

Factors 1 and 2 were significant with the mean climatological data in region F. Factor 1 was, however, more dominant. This region was clearly discernible from seasonal and monthly data as shown in Figures 11a and 11d.

In regions G and H, the first factor of the climatological data dominated (Figure 10a) with the regression coefficients being more than 0.7. The maximum values were centred in region H. In region G, however, the fourth factor (Figure 10d) was also significant on the climatological data with the regression coefficients of more than 0.13 in magnitude.

Further differences between regions G and H were evident from the PCA patterns of the monthly and seasonal records. Figures 11c and 11g for example indicate that H was significant on both factors while only one factor dominated over region G.

Region I which is centred around Gulu in northern Uganda had factor 1 dominant with the regression weight greater than 0.7 at the centre on the climatological data. The region was also clearly discernible from seasonal and monthly rainfall records PCA patterns. An example is factor 3 of season 3 (Figure 11h) were the factor loadings are greater than 0.7 at the centre.

Region J like region I had factor 1 dominating throughout the year. However, factor 4 whose regression coefficients were greater than 0.13 in the region was also significant. Other differences between regions I and J can be discerned from PCA patterns of the seasonal and monthly rainfall records for example Figures 11b and 11h.

In region K, factors 1 and 3 were significant on the climatological records having had almost equal regression weights. The spatial delineation of region K can be seen clearly from the seasonal and monthly factor loadings for example as shown in Figures 11b and 11f.

In region L, factor 3 was dominant on the climatological data (Figure 10c) with factor regression coefficients greater than 0.7 at the centre. However, factors 1 and 2 were also significant (Figures 10a and 10b) with regression coefficients greater than 0.20. The spatial extent of region L, however, can be discerned from Figure 11b.

In region M factor 3 was dominant on the climatological data (Figure 10c). The factor loadings were great than 0.6. Factor 4 also explained substantial rainfall variance with season 2 data for region M as shown in Figure 11d.

In general, the delineation of the regions were more distinct from the PCA patterns of the monthly and seasonal records. A summary of the PCA characteristics of the individual homogeneous regions are given in Table 9. The final grouping was based on unrotated, varimax rotated and obliquely rotated solutions. The stability of the delineated homogeneous regions are described in the next sections

Table 9: Summary From Analysis of the Dominant PCA modes From Climatological and Seasonal Rainfall Records for Regions given in Figure 9.

Region	Significant and Dominant spatial Factor Patterns observed
A1	The second Factor was dominant throughout the year. Factor 1 was dominant during season 2.
A2	The second factor was dominant throughout the year. Factors 1 of season 2 and 9 of season 2 were significant in the region.

B	The second and third factors were significant throughout the year. Factor 9 was dominant in in season 2.
C	Factors 2,3 and 4 were significant throughout the year. Factor 2 was dominant during season 2.
D	Factors 1,2 and 3 were sigificant throughout the year. Factor 2 dominated in region during season 3.
E	Throughout the year factors 1, 2 and 3 were significant. Factors 1 and 4 of season 2 were also significant.
F	Factors 1 and 2 were significant throughout the year. Factors 1 and 4 were significant during season 2 in the region.
G	The first factor was dominant throughout the year. However, factor 4 was also significant. During season 2 the third factor was dominant while factor 3 of season 3 was significant.
H	The first factor was dominant throughout the year. Factors 3 and 2 of seasons 2 and 3 respectively were significant in he region.
I	**The first factor was dominant throughout the year. During season 3, factor 3 was dominant in the region.**

J	The first factor was dominant throughout the year in the region. However, factor 4 was also significant. During seasons 2 and 3 factors 2 and 3 were significant in the region.
K	Factors 1 and 3 were significant throughout the year. Factor 11 of season 2 was dominant in the region while factor 2 was also significant during the season.
L	The third factor was dominant throughout the year. However, factor 1 and 2 were also significant. Factor 2 was significant in the region during season 2.
M	Factor 3 dominated the region throughout the year. Factor 4 also dominated during season 2.

3.4.1.4 STABILITY RESULTS OF THE DELINEATED ZONES

In order to verify whether the obtained regions were stable and independent, further PCA and statistical analyses were carried out. These included PCA of the individual groups, interstation correlations, vector clustering of locations and the use of rainfall climatology.

3.4.1.5 THE REGIONAL PCA RESULTS

In principle, PCA on stations within the individual homogeneous regions should have only the same factor(s) dominating. The assumption here is that all stations falling under one homogeneous region have similar temporal characteristics.

Table 10A. The PCA results of a homogeneous group of stations (Region A1 with 9 stations).

Factor	Eigenvalue	PCT of Variance	Comulative PCT
1	6.48	71.9	71.9
2	0.53	5.8	77.8
3	0.43	4.7	82.5
4	0.32	3.6	86.1
5	0.31	3.4	89.5
6	0.28	3.1	92.6
7	0.25	2.8	95.5
8	0.22	2.4	97.9
9	0.19	2.1	100.0

PCT ≡ Percentage

Table 10B. The PCA results of a homogeneous group of stations (Region B with 19 stations).

Factor	Eigenvalue	PCT of Variance	Cumulative PCT
1	11.72	61.7	61.7
2	0.95	5.0	66.7
3	0.75	4.0	70.6
4	0.60	3.2	73.8
5	0.54	2.8	76.6
6	0.52	2.7	79.3
7	0.47	2.5	81.8
8	0.44	2.3	84.1
9	0.40	2.1	86.2
10	0.35	1.9	88.1
11	0.34	1.8	89.9
12	0.33	1.7	91.6
13	0.31	1.6	93.3
14	0.27	1.4	94.7
15	0.25	1.3	96.0
16	0.23	1.2	97.2
17	0.19	1.0	98.2
18	0.18	1.0	99.2
19	0.16	0.8	100.0

PCT ≡ Percentage

Tables 10A and 10B give some examples of the results which were obtained when the climatological rainfall records from regions A1 with 9 stations and B with 19 stations were independently subjected to PCA. Only one factor was extracted in each case from the data implying that the groups were homogeneous.

Factor 1 of zone A1 extracted about 72 percent of spatial rainfall variance at the nine locations while in region B the first factor could account for about 62 percent of rainfall variance at the 19 locations.

One factor also dominated at the other regions with climatological, seasonal and monthly rainfall records.

3.4.1.6 THE USE OF INTERSTATION CORRELATIONS

Another method which has been used widely to test the stability of PCA results is the use of interstation correlations. Examples of the spatial patterns of the interstation correlations coefficients over Uganda are given in Figures 12a and 12b for two homogeneous groups A1 and I shown in Figure 9.

In Figure 12a, the correlation coefficients between station 112 and other locations were plotted at the various locations. It is clear from this figure that there is a close similarity between the

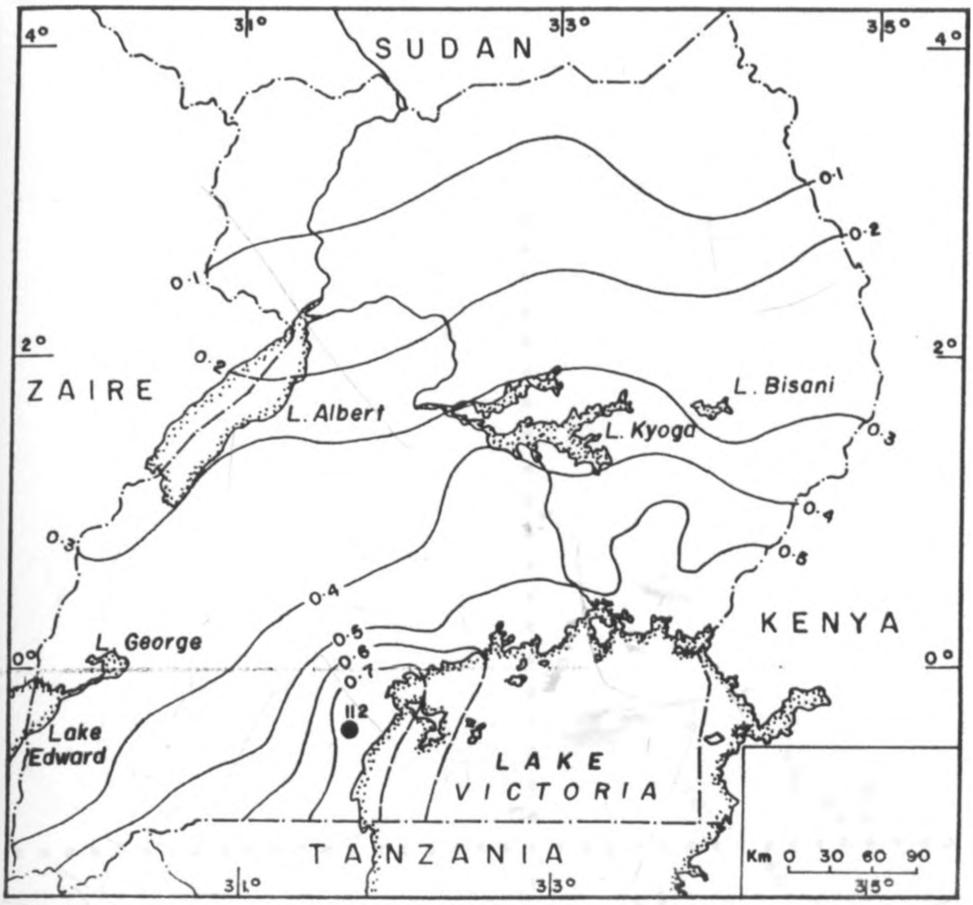


Fig. 12a : The spatial patterns of interstation correlations with station 112.

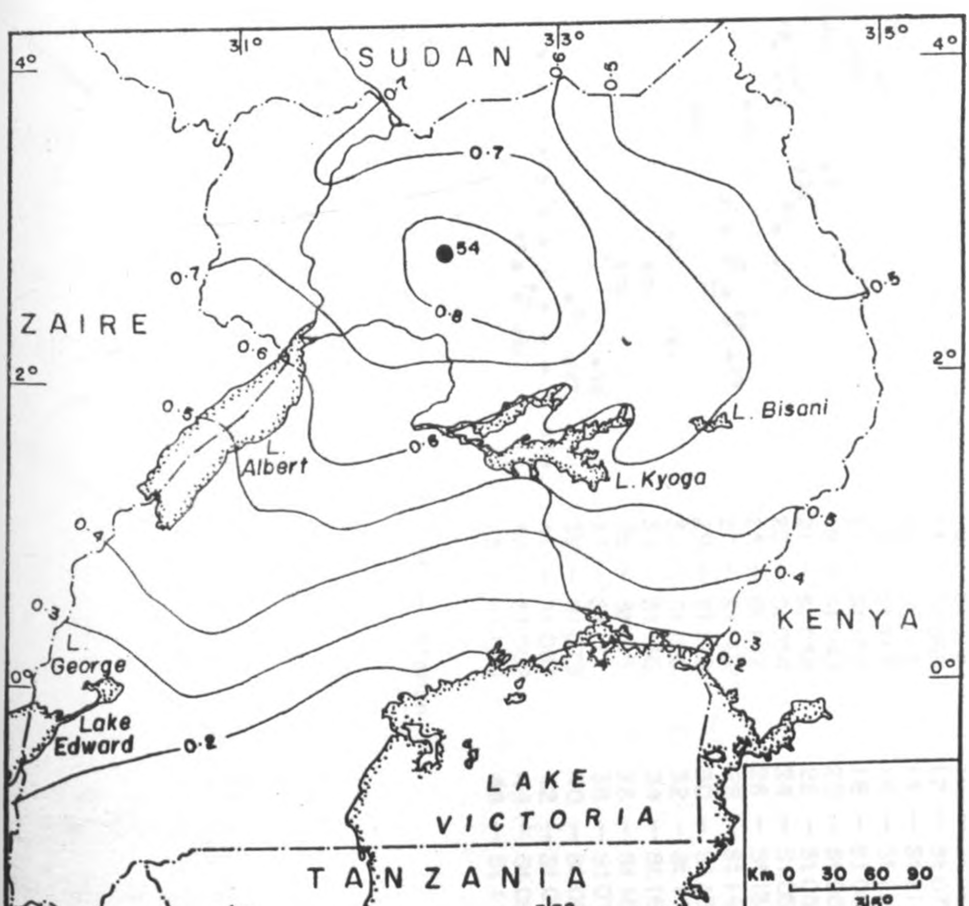


Fig. 12b : The spatial patterns of interstation correlations with station 54.

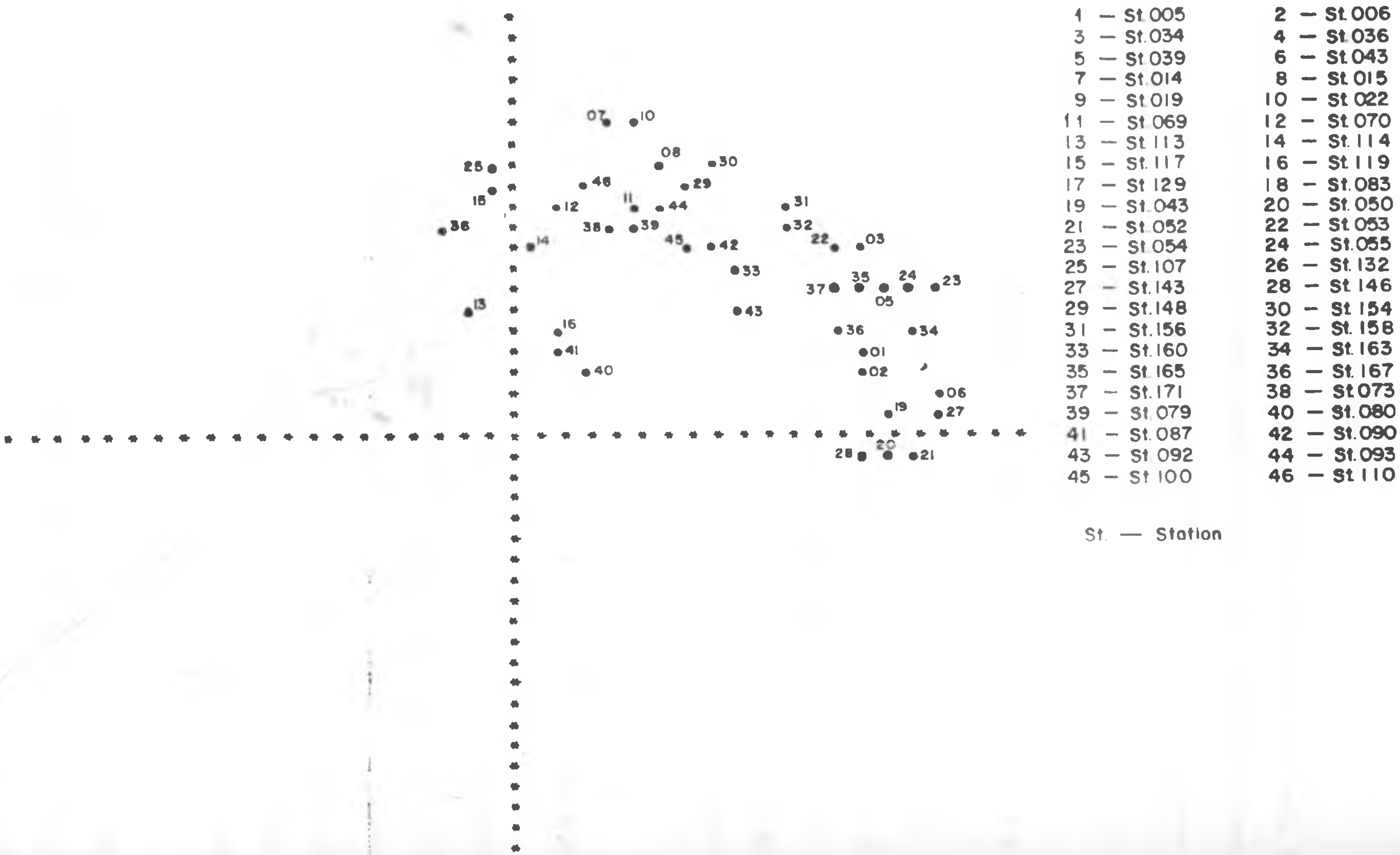


Fig. 13 : Example of a space vector plot for monthly rainfall climatological data

the spatial patterns of the correlation coefficients and the factor 1 loadings for season 2 given in Figure 11a. The correlation coeffients are highest for stations within region A1.

It can also be noted that a similar resemblerance exists between Figures 10a, 11h and 12b of region I. This confirms the high degree of spatial relationship between the locations which were delineated within the individual homogeneous zones. Similar patterns were discernible in the other homogeneous regions.

3.4.1.7 RESULTS OF VECTOR SPACE MAPPING

An example of a vector space mapping of the various locations which were obtained from PCA of the climatological rainfall records from PCA is given in Figure 13. In general, the stations within the individual homogeneous regions clustered together within the vector space of the same vectors.

3.4.1.8 PHYSICAL FACTORS ASSOCIATED WITH THE HOMOGENEOUS REGIONS.

Under this section, an attempt is made to present some of the physical factors which may be associated with the homogeneous regions discussed in the previous sections.

All the 102 station rainfall records were used to obtain mean rainfall maps over the mean annual, seasonal and monthly periods. The derived mean rainfall maps were then compared with the spatial patterns of the homogeneous regions given in Figure 9. The topographic features and the spatial distribution of the major lakes, given in Figure 3 were also compared with the spatial patterns of the homogeneous zones in Figure 9. It was observed from these comparisons that some of the homogeneous regions can be identified with isohyetal patterns, topographic features, large water bodies and other factors.

The mean annual rainfall map given in Figure 14 for example shows that high annual rainfall amounts of over 1500 mm were located over Sese Islands of Lake Victoria (zone A1) and the lake Victoria region A2. The high annual rainfall amounts can be associated with the lake/land breeze effects of lake Victoria.

The centres of regions E,F and I can be delineated by the more than 1400 mm annual rainfall isohyets. Region E's high annual rainfall totals can be attributed to lake Kyoga's mesoscale circulation systems. Region F can be seen to include mountain Elgon peak and its western slopes. This high regional topography may enhance orographic lifting and good rainfall. Similarly in region I,

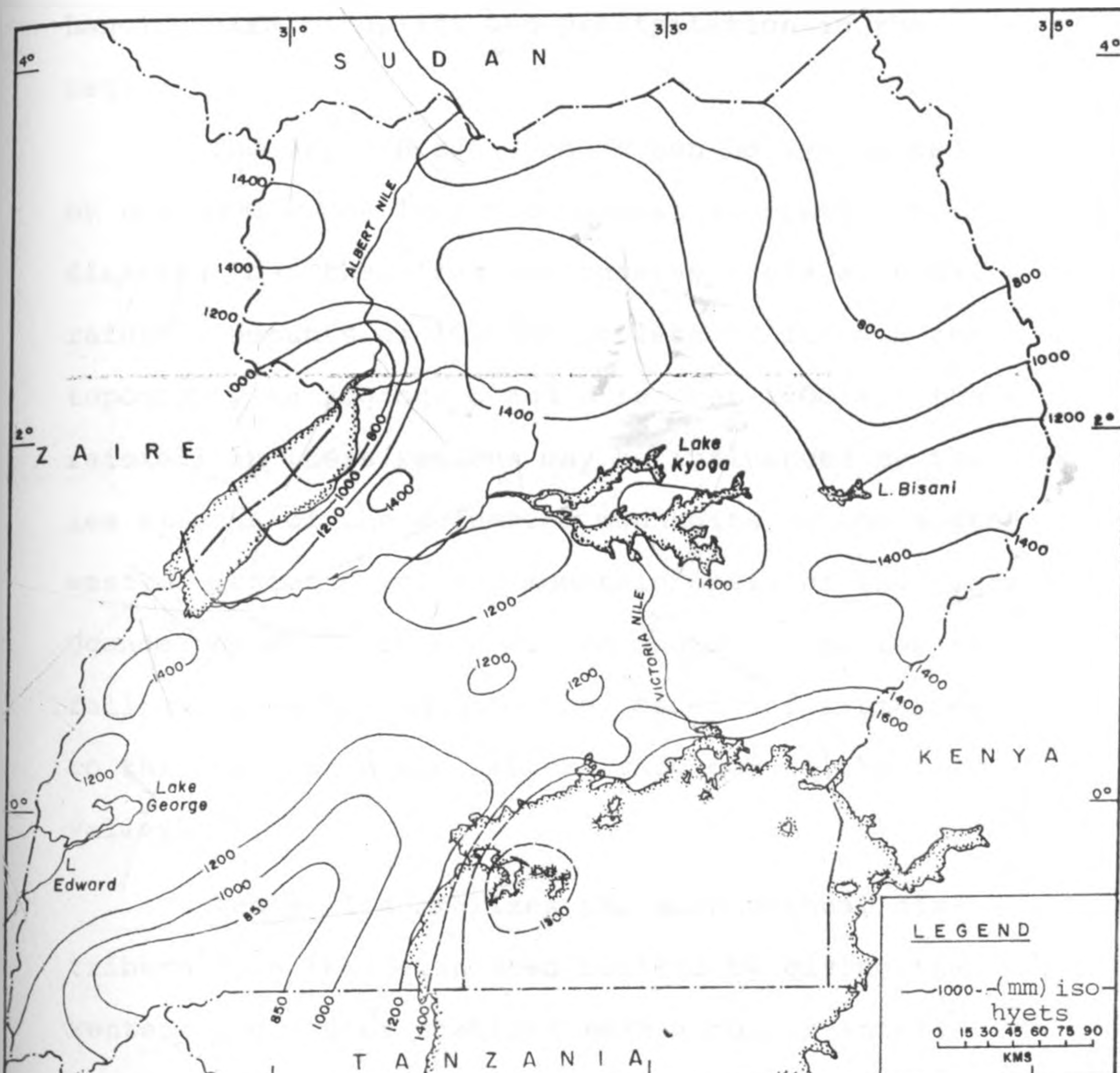


Fig 14: Mean annual rainfall map.

the ridge of high ground joining the hilly peaks of Ngetta, Moru, and Kilak plays a part in enhancing airmass uplift and precipitation in the region.

The dry zones C,G and K can be delineated by the 1000 mm or less mean annual isohyets, indicating that these regions receive average annual rainfall amounts of 1000 mm or less. Although the topography in regions C and G is over 1200 m, the rainfall in these regions may be influenced by the lee effects of the Bufumbiro mountains to the southwest of region C and the mountain chain of the Kenya/ Uganda border in region G. In region K the low rainfall can be attributed to the descent of airmasses into this region in the western shoulder of the rift valley.

Table 11 summarizes the mean annnual distribution in the delineated regions by giving the wettest and driest stations mean annual rainfall values during the period 1940-75.

Table 11: The Mean Annual Rainfall for the wettest and driest stations in the delimited homogeneous regions.

Zone	Wettest Station Name and	Mean Annual Rainfall (mm)	Driest Station Name and Station Code Number	Mean Annual Rainfall (mm)	Comments
A1	Bumangi stn. 129	1965	Katigondo stn. 114	1087	On shore areas generally drier than the Islands
A2	None	None	None	None	Only one station, therefore no wettest/ driest station
B	Moniko Estate stn. 67	1492	Wobusana stn 78	1153	A generally wet region.
C	Kabale stn. 132	1023	Lyantonde stn. 117	823	A dry zone
D	Dabani stn. 100	1660	Namwendwa stn 90	1216	A generally wet region
E	Ochero P.S. stn. 114	1417	Nakasongora stn. 158	1062	Wetter to the East than west of region
F	Bugusege stn. 9	1470	Mbale stn. 5	1167	Wetter towards Mt. Elgon.
G	Madi Opei stn. 9 (boundary with zone H)	935	Kotido stn. 52	699	Very dry region.
H	Kitgum stn. 43	1245	Madi Opei stn. 48 (boundary with zone G)	935	Dry to moderately wet
I	Gulu stn. 54	1582	Ajumani stn. 37	1195	Generally wet
J	Arua stn. 34	1461	Laropi Disp. stn. 38	985	Generally wet
K	Nebbi V.T.C stn. 53	1069	Butiaba stn. 150	753	A dry zone
L	Kijura stn. 19	1481	Kyangwari stn. 148	1113	Wet but gets wetter towards the Rwenzori mountains.
M	Kisomoro stn. 14	1477	Bugoye stn. 13	1055	Wet.

stn. = Station

Table 11 illustrates that beside the dry zones C,G and K, there is an almost uniform range between the wettest and driest stations. Thus mean annual rainfall totals may be misleading in that similar annual rainfall amounts could be interpreted to mean similarity in the spatial and temporal rainfall characteristics which may not be the case. The rainfall distribution is quite different in the individual regions throughout the year. Some of the regions could not, however, be identified from the annual rainfall map in Figure 14. The seasonal rainfall maps given in Figures 15a to 15d show that some of these regions can be delineated from the seasonal rainfall patterns.

Seaon 1 (December to February) is a dry season throughout the country with most of the rainfall concentrated in regions surrounding large water bodies like zones A1 and A2. Region C can be delineated by the 200 mm or less isohyets to the south (Figure 15a).

Season 2 (March to May) is the major rainfall season for most of the country associated with the ITCZ which migrates northward with the overhead sun. The seasonal rainfall map for the season which is given in Figure 15b, shows high rainfall totals in excess of 300 mm for most of the country except in

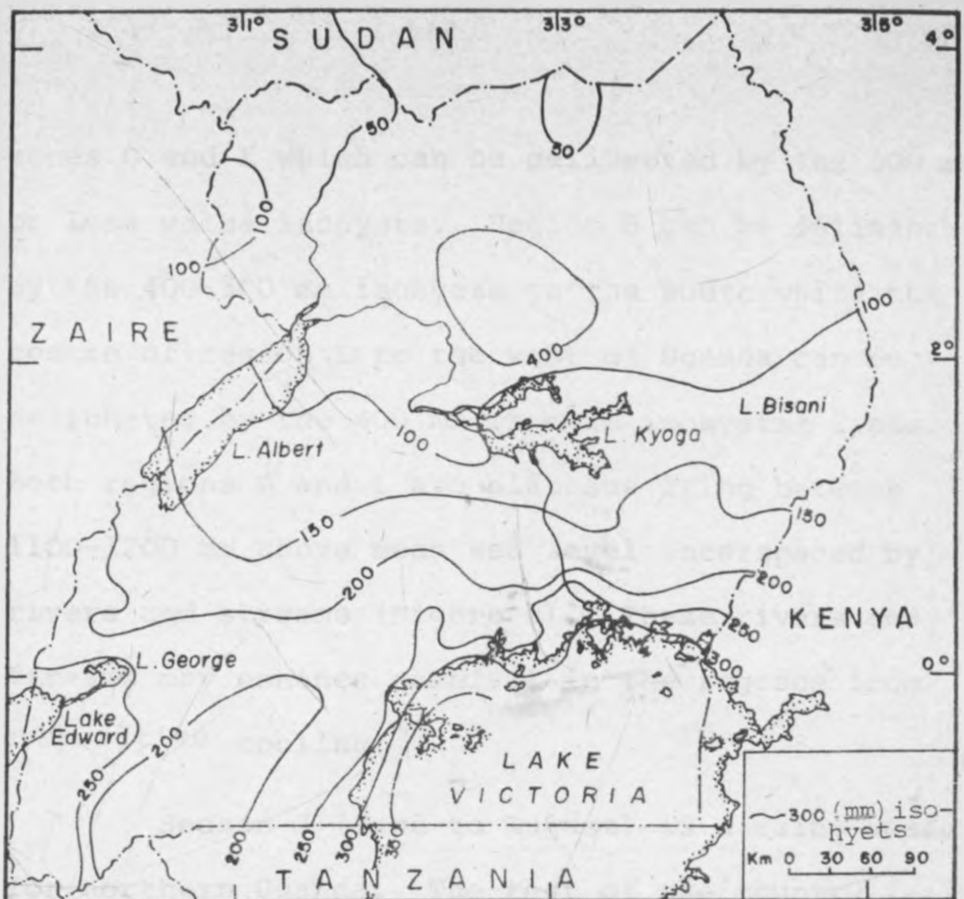


Fig. 15a : Season 1 (December – February) mean rainfall map

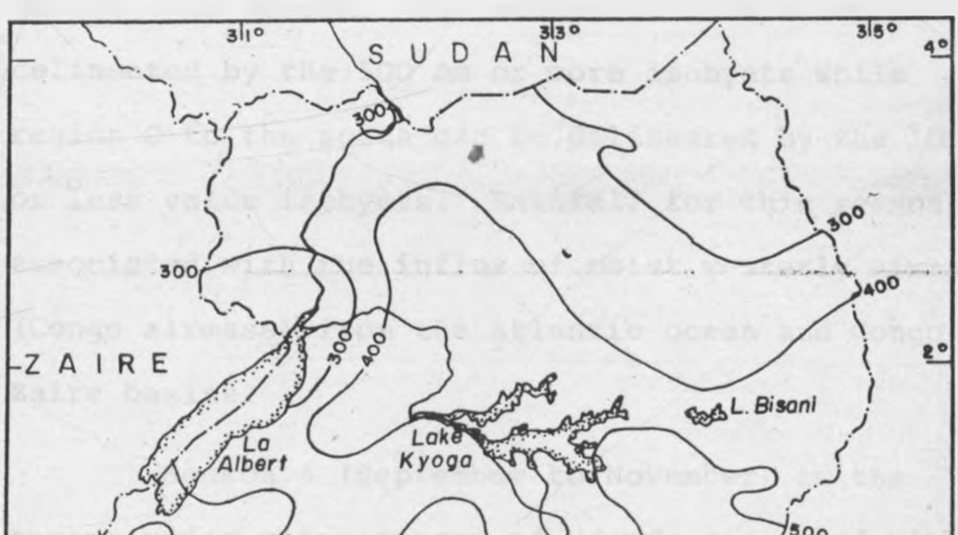


Fig. 15b : Season 2 (March – May) mean rainfall map zones G and K which can be delineated by the 300 mm or less value isohyets. Region B can be delineated by the 400-500 mm isohyets to the south while the centre of region L to the west of Uganda can be delineated by the 400 mm or more isohyetal lines. Both regions B and L are plateaus lying between 1100-1200 mm above mean sea level interspaced by rivers and streams (Figure 3). These rivers and streams may enhance rainfall in the regions from evaporative cooling.

Season 3 (June to August) is a rainy season for northern Uganda. The rest of the country is, however, generally dry (Figure 15c). Most regions of northern Uganda receive more than 300 mm of rain during this season. The centre of zone I can be delineated by the 500 mm or more isohyets while region C to the south can be delineated by the 100 mm or less value isohyets. Rainfall for this season is associated with the influx of moist westerly airmass (Congo airmass) from the Atlantic ocean and Congo/ Zaire basins.

Season 4 (September to November) is the second major rainy season of Uganda connected with the second passage of the ITCZ from the northern hemisphere. During this period most of the country receives more than 300 mm in the three months except zones G to the north and C to the south (Figure 15d).

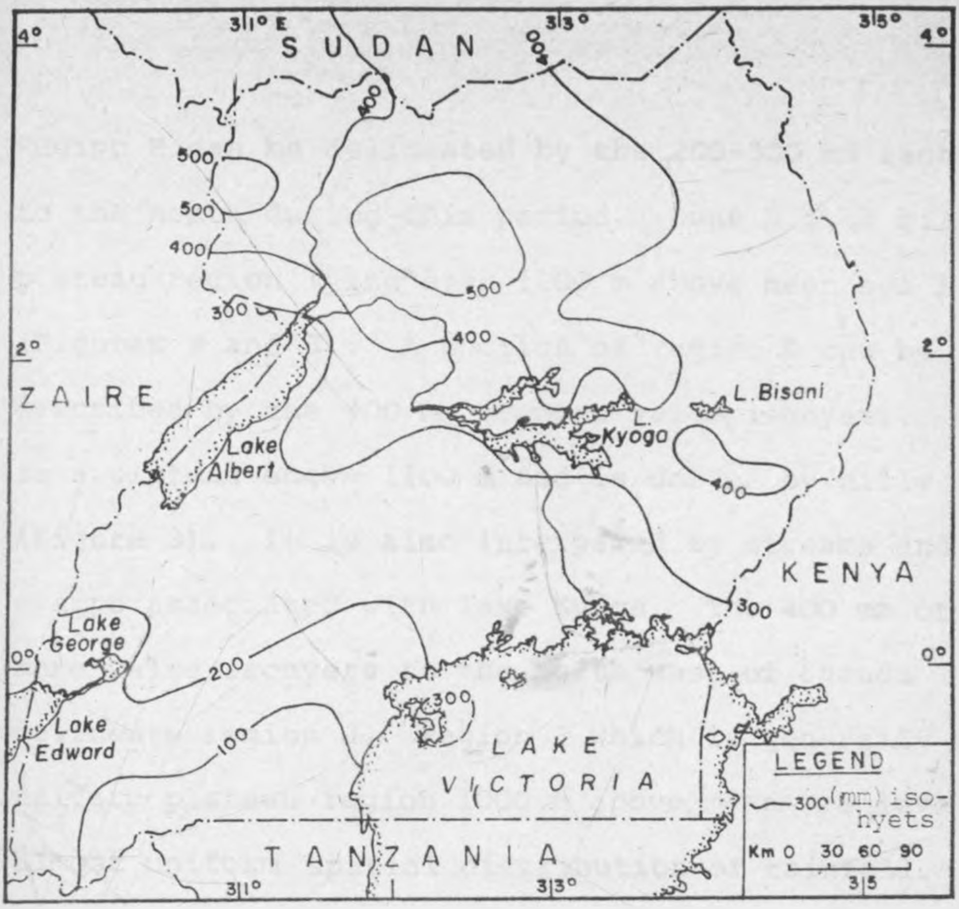


Fig. 15c : Season 3 (June– August) mean rainfall map

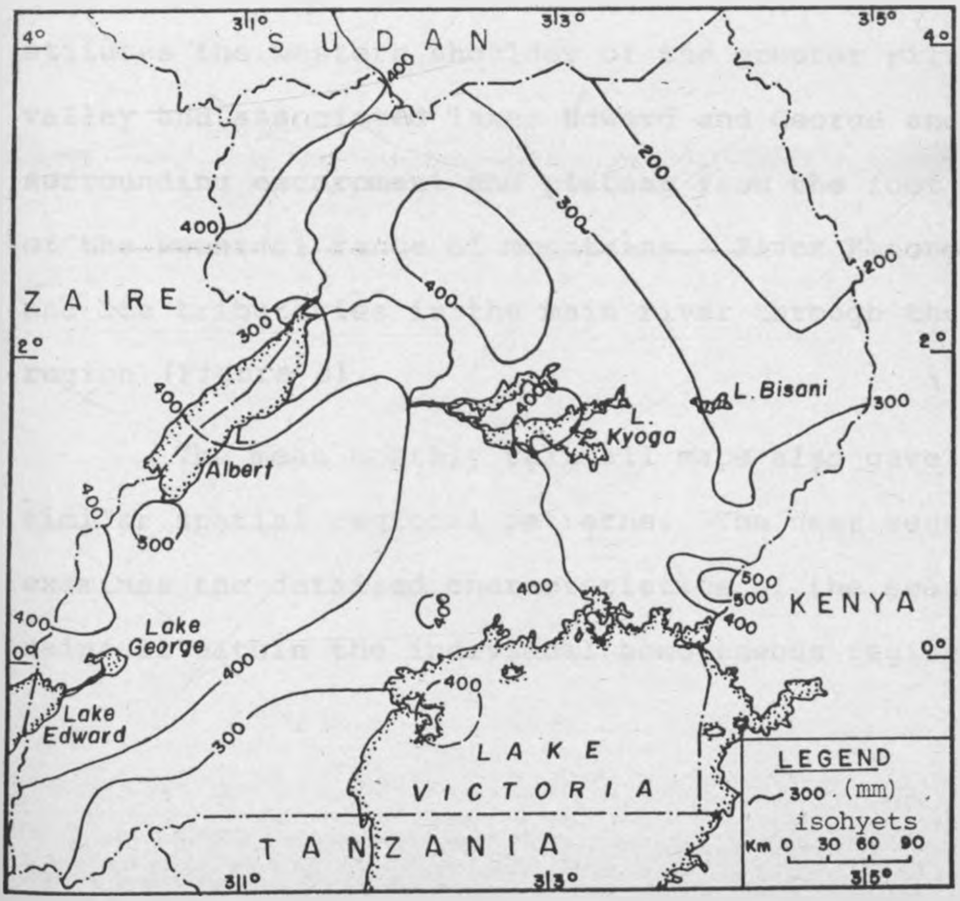


Fig. 15d : Season 4 (September — November) mean rainfall map

Region H can be delineated by the 200-300 mm isohyets to the north during this period. Zone H is a hilly plateau region lying over 1100 m above mean sea level (Figures 9 and 3). A portion of region D can be described by the 400 mm or more value isohyets. Zone D is a plateau above 1100 m and is dotted by hilly peaks (Figure 3). It is also interpaced by streams and swamps associated with lake Kyoga. The 400 mm or more value isohyets to the north west of Uganda delineate region J. Region J which is generally a uniform plateau region 1000 m above mean sea level has almost uniform spatial distribution of rainfall. Regions L and M are delineated by the 400 mm or more isohyetal lines to the west of Uganda. The relief of region L has been discussed above. Region M constitutes the western shoulder of the greater rift-valley and associated lakes Edward and George and the surrounding escarpment and plateau from the foot hills of the Rwenzori range of mountains. River Katonga and its tributaries is the main river through the region (Figure 3).

The mean monthly rainfall maps also gave similar spatial regional patterns. The next section examines the detailed characteristics of the seasonal rainfall within the individual homogeneous regions.

3.4.1.9 THE MEAN SEASONAL RAINFALL CHARACTERISTICS OVER THE INDIVIDUAL HOMOGENEOUS REGIONS

Under this section an attempt is made to explain some of the unique seasonal characteristics over each of the homogeneous zones with the help of seasonal rainfall graphs.

3.4.1.9.1 REGION A1

Figure 16 for Kyanamukaka (station 119) gives the typical seasonal rainfall characteristics of region A1. The region experiences a bimodal rainfall structure centred around March to May and September to November months. Substantial rainfall totals in excess of 50 mm are, however, received throughout the year, except during July and August when 30 to 50 mm of rainfall are realised.

These rainfall characteristics are associated with the seasonal migrations of the overhead sun inducing two (Long and Short) rainy seasons over the region. The rainfall amount received during the short rainy season (September to November) is, however, relatively lower than the amounts received during the long rainy season (March to May). This can be associated with the relatively continental dry northern hemisphere (N.H) arrivals which are advected in the region during the N.H. autumn.

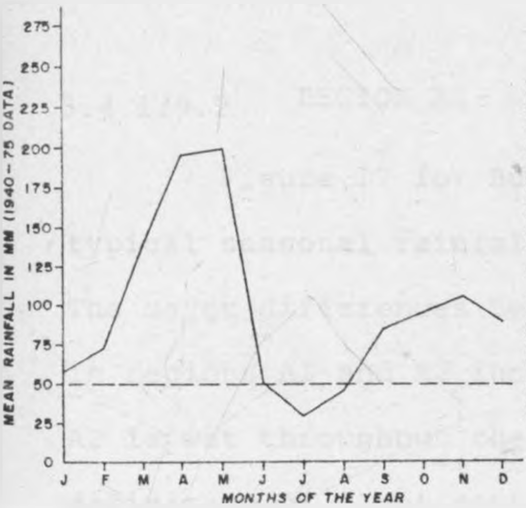


Fig 16 Seasonal rainfall distribution at Kyanamukaka (Station 119) in region A1

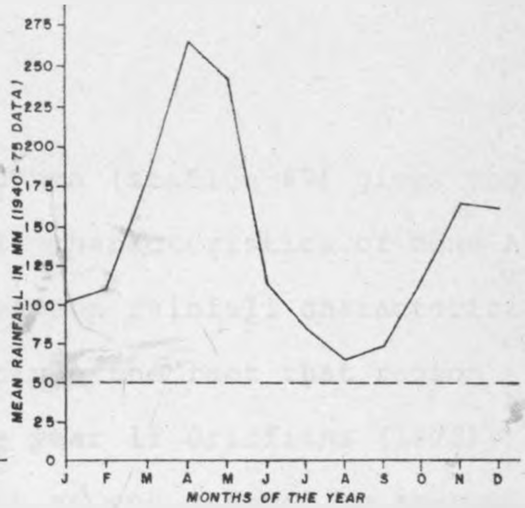


Fig. 17 : Seasonal rainfall distribution at Buvuma (Station 87) in region A2

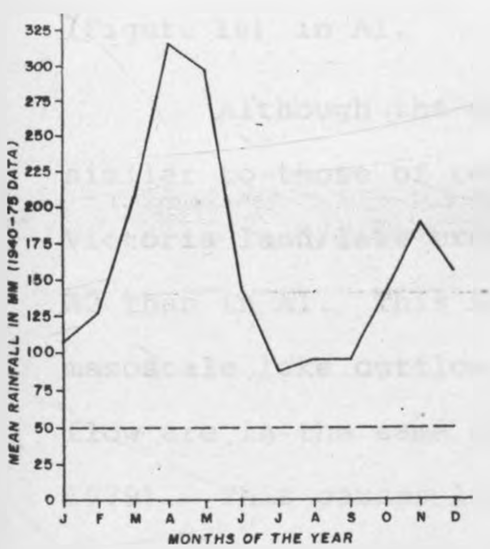


Fig 18 : Seasonal rainfall distribution at Bumangi (Station 129) in region A1

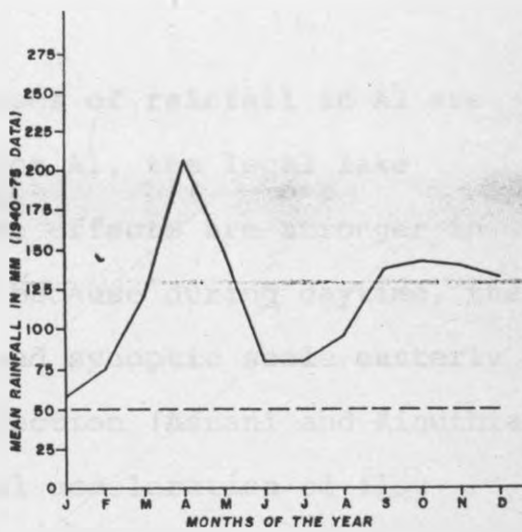


Fig.19 : Seasonal rainfall distribution at Mutai Forest Station (Station 93) in region B

The rainfall received during the N.H. summer and winter seasons is associated with the Lake Victoria effects.

3.4.1.9.2 REGION A2

Figure 17 for Buvuma (station 87) gives the typical seasonal rainfall characteristics of zone A2. The major differences between rainfall characteristics in regions A1 and A2 include the fact that region A2 is wet throughout the year if Griffiths (1972) definition of a wet month as one assumed to receive 50 mm of rainfall or more is used. The relatively dry months of July and August in region A2 had over 60 mm compared to 30 - 50 mm realised at Kyanamukaka (Figure 16) in A1.

Although the causes of rainfall in A2 are similar to those of region A1, the local lake Victoria land/lake breeze effects are stronger in A2 than in A1. This is because during daytime, the mesoscale lake outflow and synoptic scale easterly flow are in the same direction (Asnani and Kinuthia, 1979). This causes local acceleration of flow leading to horizontal velocity divergence and subsidence over the western lake shore areas (region A1) during day time. The dominant flow over region A2 is southerly to westerly during daytime. Although this effect in region A1 would be reversed at night when the land inflow near the west coast and the synoptic flow are in opposite directions causing horizontal velocity convergence, vertical motion, cooling and precipitation, A1 benefits marginally from it as the maximum effect is on the lake rather than the shore areas. Figure 18, the seasonal rainfall plot for Bumangi (station 129) on Sese Islands in the Lake Victoria confirms this by having all year around rainfall.

3.4.1.9.3 REGION B

Figure 19 gives the seasonal plot of the mean rainfall of Mutai (station 93) in region B. The region is between 1200 - 1300 m above mean sea level (AMSL) as can be seen from the relief map of Figure 3. Zone B is also interspaced by swamp and river systems.

Figure 19, indicates that region B receives rainfall throughout the year like regin A2 with a double maximum centred around March to May and September to November. These maxima like in regions and A1 and A2 can be associated with the seasonal migrations of the overhead sun when the ITCZ is in the region during the two rainy seasons. The mean seasonal rainfall, however, for the drier months were greater than 50 mm. The all year round wetness can be associated with the moderating influence of the swamps and rivers from evaporative cooling during the northern summer and winter months. The smaller rainfall peak which is observed during the N.H. autumn is associated with the relatively dry continental north easterlies from the northern hemisphere anticyclones, while the observed relative break of the dry season in August may be associated with the moist Congo air incursions during the month. The effects of the moist Congo air in the region is relatively small owing to the relatively low altitude of the area.

3.4.1.9.4 REGION C

Figure 20 gives the mean seasonal rainfall characteristics for Mbarara (station 107) in region C. It is a typical bimodal rainfall structure with double maximum centred around March to May and September to November. This can be attributed to the influence of the ITCZ in the region as it follows the overhead sun north and south respectively. There are two distinct dry seasons December to February and June to July when the ITCZ is outside the region.

Zone C is mostly high ground lying above 1500 mm AMSL (Figure 3). The region however, lies on the leeward side of the Bufumbiro mountains to

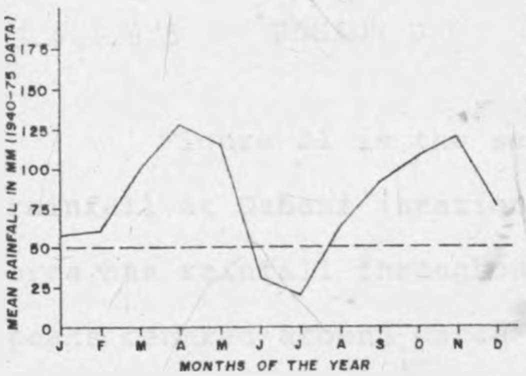


Fig 20. Seasonal rainfall distribution at Mbarara (Station 107) in region C

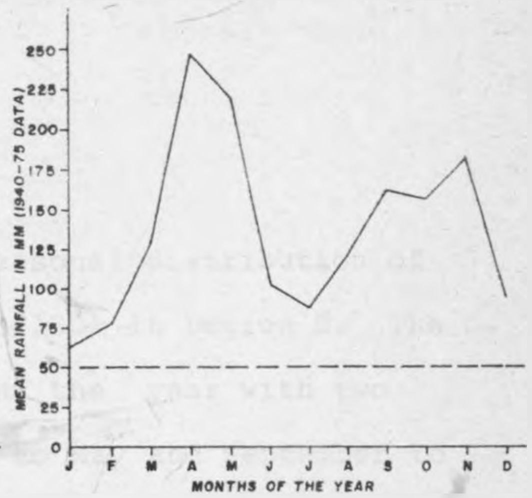


Fig 21 : Seasonal rainfall distribution at Dabani (Station 100) in region D

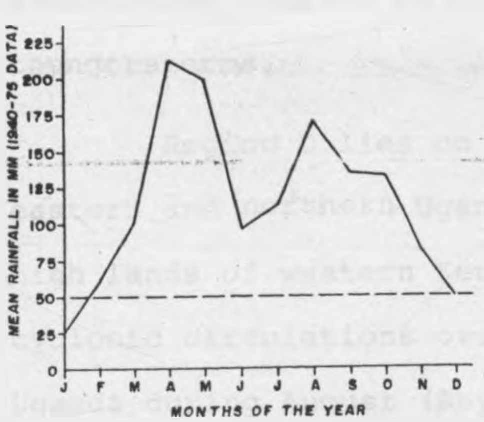


Fig 22: Seasonal rainfall distribution at Serere (Station 165) in region E

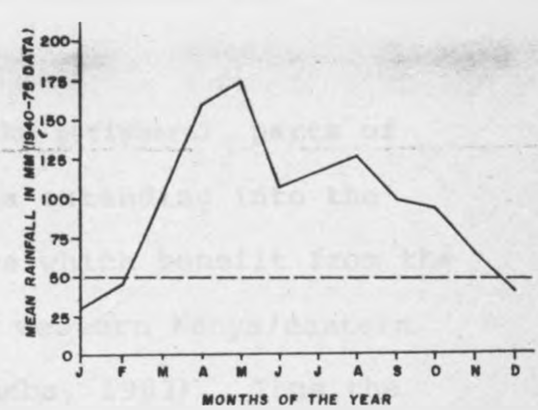


Fig 23 : Seasonal rainfall distribution at Mbale (Station 5) in region F the south west hence little or no precipitation is realised from the moist Congo air incursions during July and August.

3.4.1.9.5 REGION D

Figure 21 is the seasonal distribution of rainfall at Dabani (station 100) in region D. The area has rainfall throughout the year with two peaks centred around March to May and September to November associated with the seasonal migrations of the overhead sun when the ITCZ is the area.

Although the rains decrease during the N.H. summer and winter months, there is no dry month with less than 50 mm of rainfall. This may be due to the many swamps and streams in the area which cause evaporative cooling to give afternoon showers and thunderstorms.

Region D lies on the peripheral parts of eastern and northern Uganda extending into the high lands of western Kenya which benefit from the cyclonic circulations over western Kenya/eastern Uganda during August (Anyamba, 1983). Thus the short rains, of September to November start early in August giving a small peak in September at the start of the short rains, before their main peak in November.

3.4.1.9.6 REGION E

Figure 22 gives the seasonal rainfall distribu-distribution of Serere (station 165) in region E. The rainfall structure is unimodal unlike that of zone B to the south. The two observed rainfall peaks during March to May and September to November can be attributed to the double passage of the ITCZ following the overhead sun north and south respectively. The third peak during July/September can be associated with the incursions of the moist Congo air mass during July and August. The dry season during the N.H. winter months is somewhat moderated by the lake Kyoga lake/land breeze effects and only January month can be said to be dry with less than 50 mm mean rainfall.

3.4.1.9.7 REGION F

Figure 23 gives the seasonal rainfall characteristics of zone F for Mbale (station 5) of region F. The rainfall characteristics are unimodal with peak rainfall values concentrated between March and November. Two rainfall maximum are, however, observed in May and August. The basic difference with zone E lies in the length of the dry season which extended from December to February and the topography patterns (Figure 3). Region F has a generally higher altitude, the region surrounding mountain Elgon.

3.4.1.9.8 REGION G

Figure 24 gives the seasonal rainfall

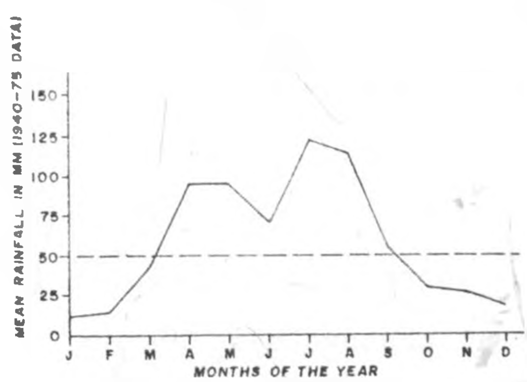


Fig 24. Seasonal rainfall distribution at Kotido (Station 52) in region G

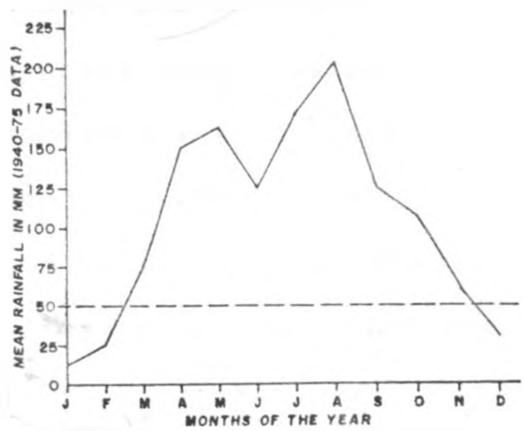


Fig 25: Seasonal rainfall distribution at Kitgum (Station 43) in region H

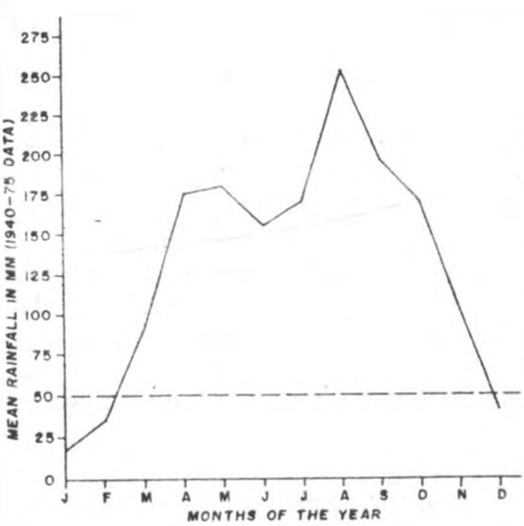


Fig. 26: Seasonal rainfall distribution at Gulu (Station 54) in region I

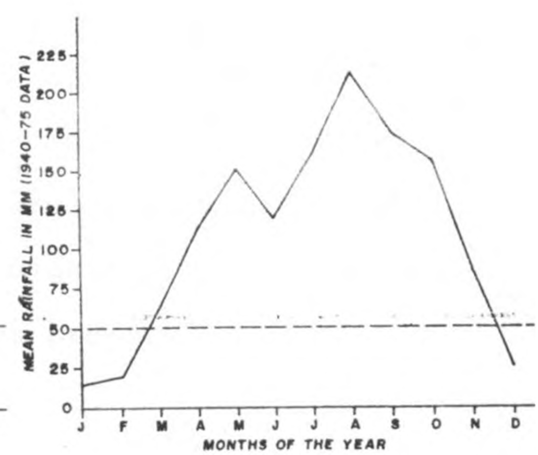


Fig. 27: Seasonal rainfall distribution at Moyo (Station 39) in region J

characteristics of Kotido (station 52) in zone G. The rainfall is basically unimodal and like in region F to the south has two peaks centred around April/May and July/August associated with the seasonal migrations of the overhead sun when the ITCZ is in the area. The major difference with all the other zones so far discussed lies in the fact that the dry season here extends from October to March a total of almost 6 months when the monthly rainfall is below 50 mm. This is a consequence of the 'rain-shadow effect' of the mountain chain on the Kenya/Uganda border and the orography of the area itself.

Region G is the driest part of Uganda and is a typical example of those regions in the tropics described by Trewartha (1961) as the 'problem climates' which have prolonged dry spells of five months or more, much of the area a genuine desert yet occupying latitudinal and geographical locations where normally humid climates prevail.

3.4.1.9.9 REGION H

Figure 25 gives the seasonal rainfall distribution at Kitgum (station 43) in zone H. Like in region G, the rainfall distribution is unimodal centred around March to November. There are two distinct peaks during April/May and July/September. The April/May peak can be associated with the presence of the ITCZ in the area, while the July/ September peak may be due to westerly influx of the moist airmass from the Atlantic and Zaire/Congo basin (Congo airmass) and the presence ITCZ as it moves southward in August giving a higher peak than the April/May peak of the long rains. These differences differentiate H from G to the East.

3.4.1.9.10 REGION I

Figure 26 gives the seasonal rainfall distribution of Gulu (station 54) in region I. Figure 3 shows zone J to be traversed from south east to north west by a ridge of high ground associated with the hilly peaks fo Ngetta (1252m), Moru (1234 m) and Kilak (1376 m). These high grounds enhance precipitation in the region from orographic lifting especially when there is a surge of moist Congo air converging with the synoptic easterlies in July/ August making zone I the wettest region of northern Uganda.

The seasonal rainfall distribution of I is quite similar to that of H to the east of it (Figure 25) except that I enjoys relatively higher precipitation compared to H and hence a lower degree of temporal variability in the rainfall characteristics.

3.4.1.9.11 REGION J

Figure 27 for Moyo (station 39) gives the typical seasonal rainfall characteristics of region J. The region experiences a unimodal rainfall structure centred around March to November. Like region I to the south, the region has two rainfall maxima during May and August.

The May peak can be associated with the presence of the ITCZ as it moves northward, while the August peak, which is substantially higher than that of May, can be attributed to the westerly influx of moist airmass (Congo airmass).

Regions I and J have similar rainfall characteristics. However, the two regions are separated by a drier Albert Nile valley of the western shoulder of the rift valley. Other observed differences can be attributed to orographic disparities and locally induced effects.

3.4.1.9.12 REGION K

Figure 28 gives the seasonal rainfall distribution of Nebbi (station 53) in region K. Region K consists of the eastern and northern shores of lake Albert; and is hence part of the western shoulder of the great rift valley and the escarpment surrounding it.

Fig 28 Seasonal rainfall distribution at Nebbi (Station 53) in region K

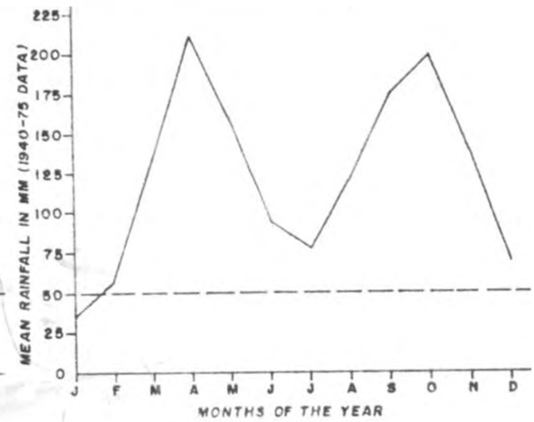


Fig 29 Seasonal rainfall distribution at Kijura (Station 19) in region L

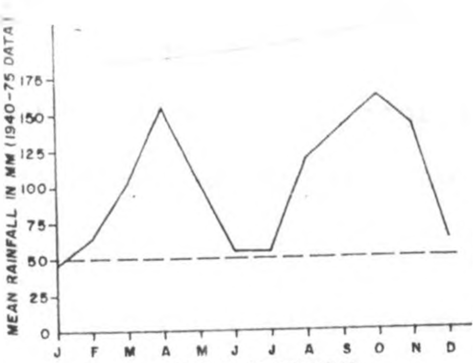


Fig 30 Seasonal rainfall distribution at Mubende (Station 22) in region M

As this area is a valley, it is apparent that airmass flowing in the region warm up as they descend and deposit very little moisture here compared to region I to the east and zone L to the south of it, making K one of the very dry zones of Uganda.

The rainfall characteristics of K are, however, similar to those of I and the major differences could be due to orography and other thermally induced regional systems.

3.4.1.9.13 REGION L

Figure 29 gives the seasonal rainfall characteristics observed at Kijura (station 19) in region L. The rainfall distribution is unimodal starting February to December with Janaury being the sole dry month of the area. There are two peaks associated with the migrations of the ITCZ north and south during March to May and September to November.

Unlike other unimodal regions already described, the surge in moist Congo airmass of July/August is not noticeable here. This could be due to the Lee effect of the Rwenzori mountain chain to the southwest and the high-grounds in eastern Zaire to the west. The drier seasons during the N.H. summer and winter months enjoy more than

50 mm of rainfall except in January when this total is about 30 mm.

A notable characteristic of the seasonal rainfall distribution in zone L (Figure 29) is the apparent equal effects of the long and short rains as reflected in the rainfall peaks during the March to May and September to November months.

3.4.1.9.14 REGION M

Figure 30 gives the seasonal rainfall characteristics observed at Mubende (station 22) in region M. Zone M has the typical bimodal rainfall characteristics associated with tropical regions, with rains during March to May and September to November associated with the presence of the ITCZ.

Although zone M surrounds lakes Edward and George, their influence cannot be differentiated because of their relatively limited circulations as compared to that of lake Victoria for example. However, these lakes probably influene the moderation of the dry months of June and July and December to February to be around 50 mm compared to region C where June and July months have rainfall totals below 50 mm.

It may be concluded from all these rainfall analyses that the patterns of the delineated homogeneous divisions represented realistic spatial modes of rainfall over Uganda. In the next section, the minimum network design which was derived from these homogeneous divisions is presented.

3.4.10 MINIMUM RAINGAUGE NETWORK DESIGN DERIVED FROM PCA

From PCA results discussed in the previous sections, a minimum of 14 homogeneous regions were identified over

Uganda. The concept of a minimum raingauge network design is to ensure that each homogeneous region is represented by at least one raingauge. A minimum raingauge network design for Uganda would therefore consist of at least 14 raingauge stations when each of the 14 homogeneous groups is represented in the network by at least one station.

The communality of a variable (station) represents the degree of association the station has with all the other stations both within and outside the homogeneous group (equation 30). A high communality value therefore represents a high degree of association. This concept was used to select the best representative station for each homogeneous region. The communalities used were derived from the PCA solutions of the mean climatological records for all the stations in each homogeneous region. Where two or more stations in a homogeneous region (group) had comparable highest communalities, other factors like the centrality of a station in the region were used to determine which of these stations best represented the region.

The fourteen (14) stations chosen using the communality concept explained above formed the minimum network design for Uganda and included Buwunga in region A1, Buvuma in region A2, Kivulu in region B, Mbarara in region C, Vukula in region D, Serere in region E, Mukongoro in region F, Kotido in region G, Kitgum in region H, Gulu in region I, Yumbe in region J, Nebbi in region K, Kijura in region L and Kisomoro in region M. The code numbers of these stations are given in Table 12A while their spatial distribution is shown in Figure 31.

The minimum network design stations would constitute the best synoptic network for Uganda since each homogeneous climatological region of the country would be represented by at least one station. However, out of the above minimum network stations, only Gulu and Mbaraka are included in Uganda's current synoptic stations network. Most of these synoptic stations, out of necessity, were started near the major national development administrative centres or towns. Thus, the current synoptic stations of Uganda may not give the best climatological representation of the rainfall characteristics of the country.

In order to determine whether the chosen stations gave good representations of the spatial rainfall characteristics at the individual homogeneous regions, areal rainfall estimates at the individual homogeneous regions were correlated with the corresponding records from the representative stations as discussed under section (2.4.1.5). The regional areal rainfall estimates which were correlated with the representative station rainfall totals were obtained from the simple arithmetic average and PCA factor weighting methods as explained under section (3.3) for each of the homogeneous regions. The magnitude of the correlation coefficient squared (r^2) is a measure of the variance of the areal rainfall which can be explained by the representative station (section 2.4.1.5). The value of r^2 was computed for each of the homogeneous regions.

Table 12A: The correlation coefficients (r) and their squares (r^2) between areal rainfall and representative station rainfall totals in the various homogeneous regions.

Region	Representative Station Code Number	Season 1 (Dec. - Feb.)		Season 2 (Marc. - May)		Season 3 (Jun. - Aug.)		Season 4 (Sept. - Nov.)		Mean Annual (Jan. - Dec.)		Comments
		r	r^2	r	r^2	r	r^2	r	r^2	r	r^2	
A1	112	0.81	0.66	0.87	0.76	0.25	0.06	0.79	0.62	0.82	0.67	Very poor correlation relationship in season 3
A2	87	No	data	for	compari	son						One station
B	137	0.88	0.77	0.82	0.67	0.60	0.36	0.88	0.77	0.86	0.76	Fairly correlated throughout the seasons.
C	107	0.78	0.61	0.82	0.67	0.85	0.72	0.77	0.59	0.79	0.62	Good relationship
D	133	0.92	0.85	0.79	0.62	0.58	0.34	0.86	0.74	0.89	0.79	Low correlation during season 3
E	165	0.87	0.76	0.72	0.52	0.72	0.52	0.85	0.72	0.70	0.49	Fairly well correlated over all periods.
F	3	0.89	0.79	0.81	0.66	0.70	0.49	0.81	0.66	0.83	0.69	Good relationship
G	52	0.74	0.55	0.63	0.40	0.62	0.38	0.74	0.55	0.54	0.29	Generally poor relation.
H	43	0.75	0.56	0.45	0.20	0.72	0.52	0.72	0.52	0.57	0.32	Correlation relationship poor in season 3
I	54	0.82	0.67	0.76	0.58	0.74	0.55	0.74	0.55	0.69	0.48	Generally fair correlation relationships
J	36	0.83	0.69	0.57	0.32	0.83	0.69	0.80	0.64	0.68	0.46	Low correlation relationship in season 2
K	53	0.77	0.59	0.63	0.40	0.78	0.61	0.73	0.53	0.51	0.26	Low correlation on the annual total.
L	19	0.82	0.67	0.71	0.50	0.74	0.55	0.82	0.67	0.77	0.59	Fairly good correlation relationship.
M	14	0.72	0.52	0.80	0.64	0.61	0.37	0.69	0.48	0.72	0.52	Generally good correlation relationships.

Table 12B: Regression parameters for the linear regression model (equation 33)

Region	Representative station code Number	Season 1 (Dec. - Feb.)		Season 2 (Marc. - May)		Season 3 (Jun. - Aug.)		Season 4 (Sept. - Nov.)		Mean Annual (Jan. - Dec.)	
		a	b	a	b	a	b	a	b	a	b
A1	112	89	0.80	238	0.60	149	0.18+	140	0.62	677	0.54
A2	87	No	data	for cal	culatio	n (one sta	tion)				
B	137	53	0.62	153	0.57	152	0.34	113	0.65	460	0.57
C	107	100	0.60	142	0.63	43	0.56	129	0.61	417	0.61
D	133	49	0.78	167	0.66	164	0.42	112	0.71	266	0.80
E	165	21	0.74	132	0.66	128	0.62	122	0.62	598	0.50
F	3	38	0.65	198	0.55	198	0.49	35	0.84	504	0.59
G	52	27	0.61	191	0.44	243	0.46	99	0.83	639	0.41
H	43	26	0.52	208	0.35	182	0.52	95	0.59	561	0.44
I	54	40	0.58	209	0.46	224	0.38	152	0.56	726	0.4o
J	36	28	0.69	204	0.40	211	0.51	99	0.73	595	0.52
K	53	36	0.79	201	0.47	101	0.58	149	0.61	754	0.32
L	19	39	0.67	247	0.34	151	0.46	182	0.54	651	0.45
M	14	80	0.51	215	0.38	125	0.48	240	0.37	701	0.38

+ Coefficient not significant at 95% level of Confidence

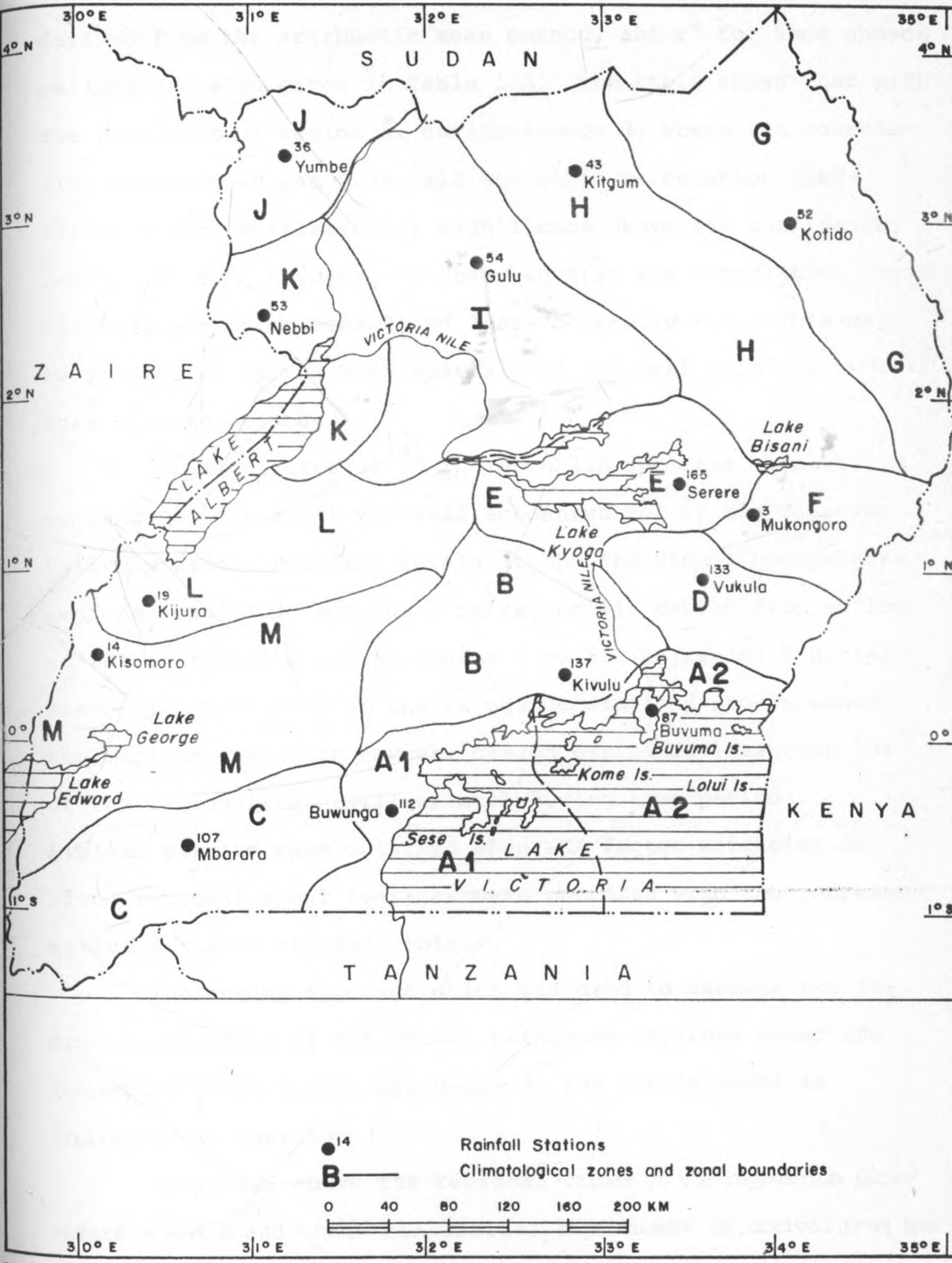


Fig. 31 : Map of the minimum raingauge network design generated from the study.

The calculated correlation coefficients (r) between the representative station totals and the regional areal totals derived from the arithmetic mean method, and r^2 for some chosen periods are also given in Table 12A. The table shows that with the exception of region A1 during season 3, where the correlation coefficient was 0.25, all the other correlation coefficients were statistically significant above 95% confidence level. It can, however, be observed that the correlation coefficients during seasons 2 and 3 are generally low. This may be attributed to the high spatial and temporal rainfall variability during these seasons.

The magnitudes of r^2 in Table 12A indicate that the variance of the areal rainfall accounted for by the representative stations rainfall totals at the individual homogeneous regions (excluding seasons 3 in region A1) varied from as low as 20% in region H during season 2 to 85% in region D during season 1. Thus some of the raingauge stations chosen under the minimum raingauge network design explained less than 50% of the regional rainfall variance during some periods. Similar results were obtained when PCA factor weighting derived regional areal averages were compared with the representative stations rainfall totals.

The second approach which was used to examine the representativeness of the chosen raingauge stations under the minimum network design was based on the linear model as indicated in equation (33).

Table 12B shows the regional values of the regression parameters a and b and their statistical significance as derived from the analysis of variance. Table 12A indicated that apart from region A1 during season 3, the computed regression parameters were statistically significant at the other 12 homogeneous regions over all the periods. Analysis of variance techniques indicated that when the areal rainfall values were derived from the representative stations totals with such a simple linear regression model, the same levels of significance of the variance were extracted (Table 12B) as in Table 12A.

It is further evident from Figure 31 that vast areas in the central parts of Uganda and several other areas are not represented by any raingauges under the minimum raingauge network design.

Thus, while the minimum raingauge network design may represent all the different spatial and temporal rainfall characteristics over Uganda and would constitute the best synoptic network stations, they may not necessarily provide the optimum rainfall informations to the various users. The next sections will be devoted to the results from the optimum raingauge network designs.

3.4.2 OPTIMUM RAINGAUGE NETWORK DESIGN RESULTS

In the previous sections, it was shown that a minimum raingauge network design for Uganda would consist of at least 14 raingauges, one raingauge in each homogeneous region. This network, however, was shown to be sometimes inadequate (section 3.4.1.10). In this section, the results of optimization of a raingauge network to adequately provide rainfall informations to the various users over the country are presented. The methods used here have included the PCA method, physiographic effects, coefficient of variation methods and optimum error of interpolation techniques. The results from these methods are, however, discussed separately in the next subsections.

3.4.2.1 PCA REGIONAL NETWORK OPTIMIZATION RESULTS

Under this method, raingague stations were added to the network of the individual homogeneous regions as given in eqution (34) until an optimum number of stations (m) which could account for at least 90% of the areal variance were obtained. Error levels of 10% in the rainfall parameters have been used by many researchers including Dymond (1982), Kohler (1958), Ahuja (1959) among several others in the derivation of results. Hence 90% of the regional areal rainfall variance was assumed to give an optimum representation of the regional areal rainfall variance. Analysis of variance techniques were used to determine the rate of increase in the areal rainfall variance when any new station was added stepwise as shown

Table 13A: The Regression parameters for the PCA optimum network design (model equation 34) For season 1.

Region	Season 1 (Dec. - Feb.)						
	m	a	b_1 (stn)	b_2 (stn)	b_3 (stn)	b_4 (stn)	b_5 (stn)
A1	3	28	0.32 (119)	0.38 (128)	0.19 (129)		
A2		No	Sufficient data for analysis				
B	2	39	0.35 (67)	0.51 (159)			
C	4	17	0.29 (107)	0.30 (105)	0.16 (111)	0.20 (132)	
D	2	39	0.46 (133)	0.41 (170)			
E	3	6	0.40 (158)	0.31 (171)	0.27 (160)		
F	1	27	0.87 (7)				
G	3	6	0.29 (49)	0.32 (52)	0.23 (146)		
H	3	11	0.27 (46)	0.35 (44)	0.18 (144)		
I	2	4	0.51 (156)	0.40 (42)			
J	3	2	0.33 (40)	0.36 (39)	0.20 (34)		
K		No	sufficient data for analysis				
L	3	30	0.26 (152)	0.30 (149)	0.27 (153)		
M	3	21	0.28 (110)	0.30 (14)	0.27 (15)		

Note: m = Optimum number of stations

(stn) = Station code number

a = Constant

b,d,---.b_m = Regression Coefficient

Table 13B:

The Regression parameters for the PCA optimum network design (model equation 34) for season 2.

Region	Season 2 (Mar. - May)						
	m	a	b_1 (stn)	b_2 (stn)	b_3 (stn)	b_4 (stn)	b_5 (stn)
A1	3	122	0.42 (114)	0.30 (112)	0.12 (129)		
A2			No sufficient data for analysis				
B	4	27	0.28 (137)	0.24 (75)	0.22 (93)	0.15 (83)	
C	3	49	0.34 (107)	0.26 (132)	0.33 (117)		
D	4	39	0.29 (88)	0.25 (133)	0.19 (90)	0.23 (99)	
E	4	21	0.31 (164)	0.25 (171)	0.23 (165)	0.12 (4)	
F	3	15	0.36 (7)	0.36 (9)	0.23 (167)		
G	3	36	0.27 (146)	0.26 (49)	0.28 (50)		
H	4	57	0.24 (143)	0.26 (48)	0.11 (46)	0.23 (163)	
I	4	46	0.40 (55)	0.17 (42)	0.15 (151)	0.12 (156)	
J	3	64	0.36 (38)	0.23 (40)	0.25 (39)		
K	No	sufficient data for analysis					
L	4	89 4	0.27 (154)	0.19 (19)	0.20 (152)	0.14 (11)	
M	4	56	0.18 (14)	0.31 (15)	0.14 (110)	(16)	

Note: m = Optimum number of stations

(stn) = Station code number

a = Constant

$b, b, ---, b_m$ = Regression Coefficient

Table 13C :

The Regression parameters for the PCA optimum network design (model equation 34). for season 3.

Region	Season 3 (Jun. - Aug.)						
	m	a	b_1 (stn)	b_2 (stn)	b_3 (stn)	b_4 (stn)	b_5 (stn)
A1	5	22	0.15 (116)	0.13 (80)	0.17 (119)	0.13 (129)	0.16 (114)
A2	No suffi		cient	data for	analysis		
B	4	66	0.23 (73)	0.17 (81)	0.16 (78)	0.18 (84)	
C	3	12	0.33 (107)	0.22 (106)	0.23 (111)		
D	4	55	0.26 (99)	0.25 (170)	0.17 (90)	0.16 (88)	
E	3	62	0.35 (169)	0.17 (4)	0.25 (165)		
F	3	64	0.34 (9)	0.26 (5)	0.21 (7)		
G	4	13	0.21 (50)	0.33 (146)	0.19 (49)	0.21 (48)	
H	4	83	0.22 (45)	0.20 (143)	0.17 (51)	0.19 (43)	
I	4	86	0.19 (54)	0.25 (161)	0.21 (42)	0.13 (57)	
J	3	76	0.41 (36)	0.37 (39)	0.40 (34)		
K	No suffi		cient	data	for anal	ysis	
L	3	56	0.22 (153)	0.28 (19)	0.21 (156)		
M	4	39	0.25 (15)	0.21 (16)	0.19 (110)	0.18 (13)	

Note: m = Optimum number of stations

(stn) = Station code number

a = Constant

$b, b, ---, b_m$ = Regression Coefficient

Table 13D:

The Regression parameters for the PCA optimum network design (model equation 34) for season 4.

Region	Season 4 (Sept. - Nov.)						
	m	a	b_1 (stn)	b_2 (stn)	b_3 (stn)	b_4 (stn)	b_5 (stn)
A1	3	48	0.25 (129)	0.30 (112)	0.20 (80)		
A2		No sufficient data for analysis					
B	2	81	0.38 (86)	0.40 (73)			
C	3	53	0.32 (107)	0.26 (111)	0.32 (105)		
D	3	32	0.36 (133)	0.25 (100)	0.22 (88)		
E	3	44	0.32 (165)	0.33 (166)	0.22 (160)		
F	2	54	0.46 (9)	0.31 (6)			
G	3	29	0.25 (49)	0.23 (50)	0.47 (52)		
H	3	47	0.35 (48)	0.35 (45)	0.19 (143)		
I	3	74	0.29 (156)	0.30 (42)	0.22 (62)		
J	3	26	0.29 (40)	0.29 (39)	0.29 (34)		
K			No sufficient data for analysis				
L	4	89	0.38 (19)	0.27 (149)	0.31 (156)	0.28 (154)	
M	4	48	0.29 (16)	0.22 (14)	0.24 (22)	0.13 (110)	

Note: m = Optimum number of stations

(stn) = Station code number

a = Constant

$b, b, ---, b_m$ = Regression Coefficient

Table 13E:

The Regression parameters for the PCA optimum network design (model equation 34) for the mean annual values.

Region	Mean Annual (Jan.- Dec.)						
	m	a	b_1 (stn)	b_2 (stn)	b_3 (stn)	b_4 (stn)	b_5 (stn)
A1	4	191	0.20 (112)	0.17 (129)	0.23 (114)	0.26 (128)	
A2		No	sufficient data for analysis				
B	3	165	0.32 (137)	0.24 (96)	0.24 (67)		
C	4	127	0.33 (107)	0.18 (117)	0.18 (111)	0.26 (106)	
D	3	120	0.40 (133)	0.25 (88)	0.22 (100)		
E	4	183	0.23 (166)	0.37 (158)	0.10 (4)	0.26 (171)	
F	3	138	0.22 (167)	0.28 (3)	0.32 (9)		
G	4	92	0.20 (50)	0.20 (49)	0.36 (52)	0.18 (48)	
H	4	180	0.20 (45)	0.24 (43)	0.22 (144)	0.19 (51)	
I	5	163	0.32 (55)	0.16 (42)	0.19 (161)	0.11 (60)	0.11 (41)
J	3	213	0.36 (40)	0.27 (34)	0.20 (38)		
K		No	sufficient data	for	analysis		
L	4	267	0.19 (19)	0.14 (156)	0.27 (154)	0.20 (17)	
M	4	190	0.20 (16)	0.26 (15)	0.21 (14)	0.14 (110)	

Note: m = Optimum number of stations

(stn) = Station code number

a = Constant

b , b ,----, b_m = Regression Coefficient

Table 13F: Regional Areal Variances ($x10^2$) Explained at Each Regression Step for season 1.

Re-gion	Exi-siting No. of Stations	Season 1 (Dec. - Feb.)						M_R
		m	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	
A1	9	3	0.74	0.88	0.93			7
A2	1	No	sufficient data for analysis					1
B	19	2	0.79	0.93				12
C	6	4	0.61	0.81	0.87	0.93		6
D	7	2	0.84	0.92				6
E	8	3	0.76	0.88	0.92			8
F	6	1	0.91					6
G	5	3	0.56	0.82	0.98			5++
H	9	3	0.77	0.89	0.94			8++
I	11	2	0.85	0.93				10*
J	5	3	0.74	0.89	0.95			5
K	2	No	sufficient data for analysis					2 *
L	10	3	0.77	0.88	0.93			7*
M	6	3	0.63	0.85	0.91			6
Totals		32 (+3)						89

Note: s_1^2, s_2^2, ---s_i^2, --- s_m^2 = Explained variance at the ith step.
m = Seasonal optimum number of stations
M_R = Idealized optimum regional number of stations.

Table 13G:

Regional Areal Variances (x10 ^{2}) Explained at Each Regression Step for season 2

Region	Exisiting No. of Stations	Season 2 (Mar. - May)						M_R
		m	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	
A1	9	3	0.78	0.88	0.92			7
A2	1	No sufficient data for analysis						1
B	19	4	0.67	0.80	0.89	0.92		12
C	6	3	0.66	0.84	0.91			6
D	7	4	0.68	0.82	0.88	0.92		6
E	8	4	0.61	0.80	0.85	0.90		8
F	6	3	0.66	0.88	0.94			6
G	5	3	0.46	0.74	0.90			5++
H	9	4	0.48	0.72	0.83	0.90		8++
I	11	4	0.67	0.83	0.89	0.93		10*
J	5	3	0.65	0.83	0.92			5
K	2	No sufficient data for analysis						2*
L	10	4	0.59	0.81	0.86	0.90		7*
M	6	4	0.63	0.77	0.87	0.93		6
Totals		43 (+3)						89

Note: s_1^2, s_2^2, ---s_i^2 , ---s_m^2 = Explained variance at the ith step.
m = Seasonal optimum number of stations
M_R = Idealized optimum regional number of stations.

Table 13H

Regional Areal Variances (x10 ^{2}) Explained at Each Regression Step for season 3.

Re-gion	Exi-siting No. of Stations	Season 3 (Jun. - Aug.)						M_R
		m	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	
A1	9	5	0.48	0.68	0.80	0.87	0.92	7
A2	1	No sufficient data for analysis						1
B	19	4	0.57	0.72	0.82	0.91		12
C	6	3	0.72	0.84	0.92			6
D	7	4	0.54	0.79	0.87	0.91		6
E	8	3	0.63	0.86	0.91			8
F	6	3	0.67	0.84	0.91			6
G	5	4	0.43	0.70	0.87	0.97		5++
H	9	4	0.66	0.80	0.89	0.93		8++
I	11	4	0.55	0.81	0.89	0.92		10*
J	5	3	0.70	0.80	0.91			5
K	2	No sufficient data for analysis						2*
L	10	3	0.60	0.80	0.90			7*
M	6	4	0.45	0.79	0.86	0.92		6
Totals		44 (+3)						89

Note: s_1^2 , s_2^2 , ---s_i^2 ,---s_m^2 = Explained variance at the ith step.

m = Seasonal optimum number of stations

M_R = Idealized optimum regional number of stations

Table 13I:

Regional Areal Variances (x10 ^{2}) Explained at Each Regression Step for season 4 .

Region	Exisiting No. of Stations	Season 4 (Sept. - Nov.)						M_R
		m	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	
A1	9	3	0.71	0.89	0.94			7
A2	1		No sufficient data for analysis					1
B	19	2	0.82	0.93				12
C	6	3	0.60	0.84	0.93			6
D	7	3	0.74	0.89	0.95			6
E	8	3	0.72	0.86	0.92			8
F	6	2	0.79	0.91				6
G	5	3	0.64	0.81	0.94			5^{++}
H	9	3	0.65	0.85	0.91			8^{++}
I	11	3	0.72	0.87	0.93			10*
J	5	3	0.69	0.86	0.95			5
K	2		No sufficient data for analysis					2*
L	10	4	0.68	0.82	0.89	0.93		7*
M	6	4	0.55	0.73	0.85	0.91		69
Totals		36 (+3)						89

Note: s_1^2 , s_2^2 , ---s_i^2 , ---s_m^2 = Explained variance at the ith step.

m = Seasonal optimum number of stations

M_R= Idealized optimum regional number of stations.

Table 13J:

Regional Areal Variances (x10 ^{2}) Explain at Each Regression Step for the mean annual values.

Region	Existing No. of Stations	Mean Annual (Jan. - Dec.)						M_R
		m	s_1^2	s_2^2	s_3^2	s_4^2	s_5^2	
A1	9	4	0.68	0.82	0.89	0.93		7
A2	1		No suf	ficient	data	for ana	lysis	1
B	19	3	0.75	0.87	0.93			12
C	6	3	0.75	0.87	0.93			6
D	7	4	0.62	0.78	0.86	0.93		6
E	8	4	0.54	0.77	0.84	0.90		8
F	6	3	0.76	0.86	0.94			6
G	5	4	0.38	0.69	0.87	0.95		5++
H	9	4	0.58	0.75	0.85	0.94		8++
I	11	5	0.51	0.72	0.84	0.88	0.92	10*
J	5	3	0.66	0.83	0.90			5
K	2	No	suffici	ent	data for	analys	is	2*
L	10	4	0.59	0.75	0.84	0.93		7*
M	6	4	0.56	0.72	0.81	0.91		6
Totals		45 (+3)						89

Note: s_1^2, s_2^2 , ---s_i^2 , ---s_m^2 = Explained variance at the ith step.

m = Seasonal optimum number of stations

M_R = Idealized optimum regional number of stations.

in equations (34). The regression parameters from the optimization scheme at each homogeneous region over the selected periods are given in Tables 13A to 13E while the corresponding increases in the explained areal rainfall variance at each regression step are given in Tables 13F to 13J.

Tables 13A to 13E indicate that the optimum number of raingauges (m) required at the individual homogeneous regions ranged from one in region F during season 1 to five stations in regions A1 and I during season 3/annual period respectively.

The Tables indicate that in a number of cases, different stations were picked up in the optimum network during the various seasons/periods. The code numbers of the stations which were chosen at each regression step are bracketed. For example in region A1, stations 119, 128 and 129 formed the optimum network in season 1; stations 114, 112 and 129 in season 2; stations 116, 80, 119, 129 and 114 in season 3; stations 129, 112 and 80 in season 4 and stations 112, 129, 114 and 128 for the annual period (Tables 13A to 13E).

The corresponding regression parameters of the optimum network derived from equation (34) are also given in the tables. In region A1 for example in season 1, the constant (a) was 28 mm and the regression coefficients (b_i, i = 1,2,....,m) of 0.32, 0.38 and 0.19 for stations 119, 128 and 129 respectively were obtained (Table 13A). In season 2, the constant (a) was 122 mm and the regression coefficients (b_i) were 0.42, 0.30 and 0.12 for stations 114, 112 and 129, respectively (Table 13B). A constant (a) of 22mm and regression coefficients (b_i) of 0.15, 0.13, 0.17, 0.13 and 0.16 for stations 116, 80, 119, 129 and 114 respectively were obtained for season 3 (Table 13C). In season 4, the constant (a) was 48 mm while the regression coefficients (b_i) were 0.25, 0.30 and 0.20 for stations 129, 112 and 80 respectively (Table 13D). A constant (a) of 191 mm and regression coefficients (b_i) of 0.20, 0.17 0.23 and 0.26 for stations 112, 129, 114 and 128 respectively were obtained for the annual period (Table 13E). The regression coefficients (b_i) of the optimum network stations at each homogeneous region for each season/period were found to be significant above 95% confidence level when tested using the student test. This confirmed the goodness of fit of the optimum network stations in estimating the areal rainfall at the individual regions for the various periods/ seasons.

Details of the optimum network stations and the corresponding regression parameters derived from the equation (34) for the other regions over the various seasons/periods are given in Tables 13A to 13E.

The proportion of the regional areal rainfall variance explained when each of the stations in the optimum network was included in the regression schemes are shown in Tables 13F to 13J. In region A1 for example in season 1, station 119 accounted for 74% stations 119 and 128, 88% while the optimum network consisting of stations 119, 128 and 129 accounted for a total of 93% of the regional areal rainfall variance (Table 13F). Similarly, in season 2, station 114 accounted for 78%; stations 114 and 112, 88% and the optimum network of stations 114, 112 and 129 92% of the regional areal rainfall variance (Table 13G). Station 116 accounted for 48%; stations 116 and 80, 68%; stations 116, 80 and 119, 80%; stations 116, 80 119 and 129, 87% and the optimum network consisting of stations 116, 80, 119, 129 and 114 accounted for 92% of the regional areal rainfall variance in season 3 (Table 13H). In season 4, station 129 accounted for 71%, stations 129 and 112, 89%; and the optimum network stations consisting of stations 129, 112 and 80 accounted for 94% of the regional areal rainfall variance (Table 13I). In the annual period, station 112 accounted for 68%; stations 112 and 129, 82% stations 112, 129 and 114, 89% and stations 112, 129, 114 and 128 forming the optimum network accounted for 93% of the regional areal rainfall variance (Table 13J).

Details of the proportions of the regional areal rainfall variances explained at each step (S_i^2, i=1,2,...,m) for the other regions are given in Tables 13F to 13J.

An optimum network must be able to adequately represent the unique rainfall characteristics of each of the seasons/periods. In this study, the optimum stations (m) picked up in each season/period were considered as subsets of the idealized set of optimum stations. The mathematical concept of the union of sets (Nahikian, 1966) was finally used to determine the idealized optimum raingauge network design for each of the homogeneous regions. The optimum number of stations (M_R) for region A1 for all seasons/periods (Tables 13A to 13E) for example may be expressed as:-

$$M_R = \{119,128,129\} \cup \{114,112,129\} \cup \{119,80,129,114\} \cup \{129,112,80\} \cup \{112,129,114,128\} \equiv \{80,112,114,116,119,128,129\}$$

a set of 7 stations,

where U ≡ union.

These stations formed the optimum network design for region A1. The regional optimum number of stations (M_R) which were obtained at the other regions using the union of sets concept are shown in Tables 13F to 13J. In total 89 stations formed the optimum gauge network design for Uganda. It should be noted that this number (89) includes the 1 and 2 stations in regions A2 and K respectively which could not be subjected to optimization analyses. The spatial distribution of the 89 stations are shown in figure 32.

Figure 32 gives a better spatial raingauge distribution than the minimum network which was given in Figure 31. All the minimum network design stations however, form part of the optimum network design.

One advantage of this method of optimization of a raingauge network is that the generated numbers of gauges were not influenced by the magnitudes of the rainfall totals but by their temporal characteristics. Another advantage is that all the 89 stations included in the optimum network design are already established and require no relocation. Thus station siting problems and station establishment costs may be avoided. A disadvantage, however, is that in the derivation of the PCA homogeneous regions, the unique variances of the individual locations were neglected. However, the unique variances are generally small compared to the common variances on which the PCA regionalization were based as discussed earlier in the text. Any changes in the regional PCA patterns which may affect the derived optimum network design would, therefore, be expected to be small.

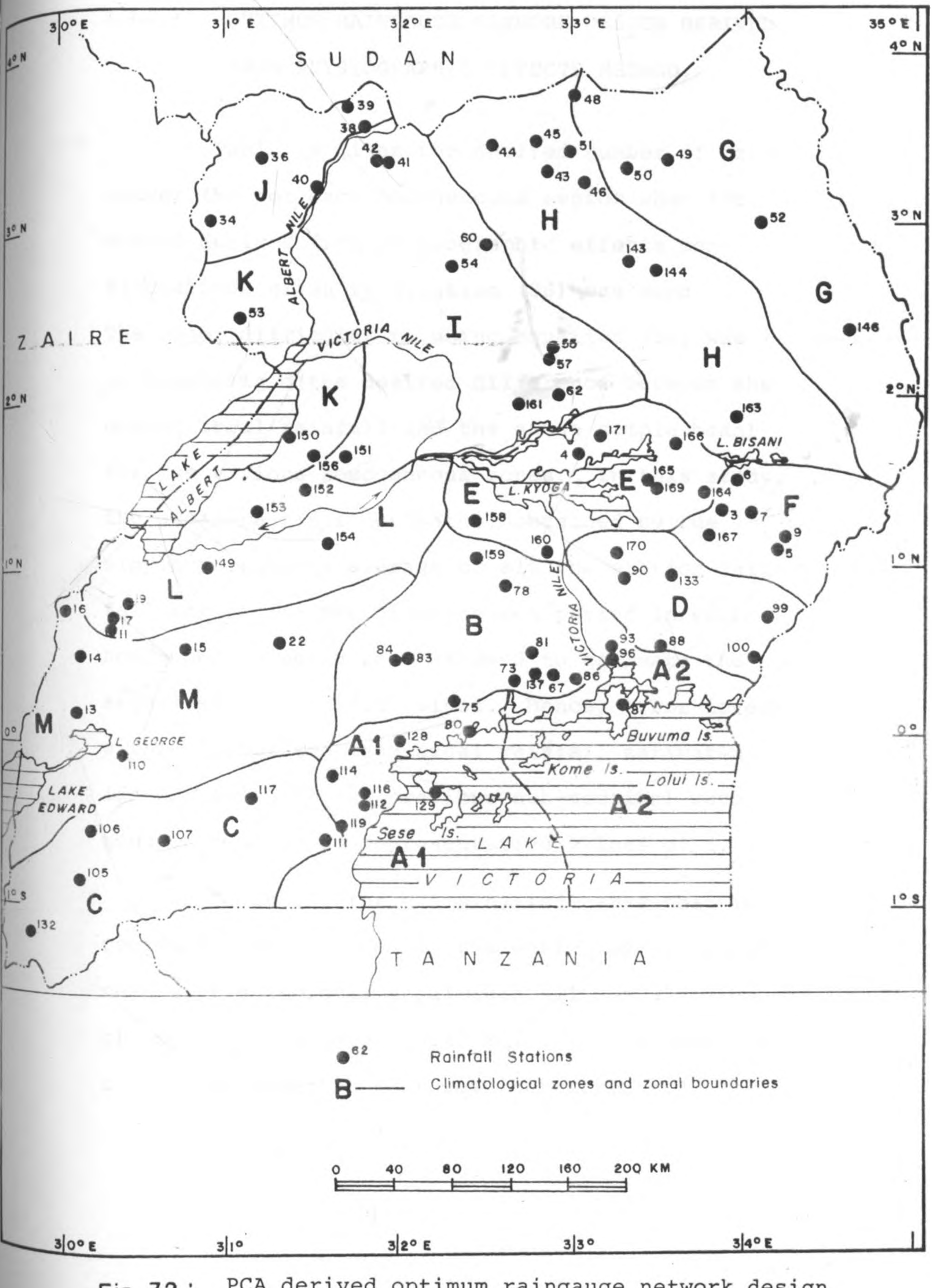


Fig. 32 : PCA derived optimum raingauge network design

3.4.2.2 OPTIMUM RAINGAUGE NETWORK DESIGN DERIVED FROM PHYSIOGRAPHIC EFFECTS METHOD.

Table 14 gives the desired number of raingauges (N) for each homogeneous region when the method derived from physiographic effects considerations given in equation (36) was used. The major difficulty in using equation (36) was in obtaining d(the desired difference between the actual areal rainfall and the gauge sample mean) for the various homogeneous zones. In this study, the areal rainfall estimates, obtained by the simple arithmetic average of all the station rainfall totals for the given season/period in each homogeneous region were assumed to estimate the actual areal rainfall values. Hence, error levels within 10% of the mean areal rainfall estimates (as explained in the previous sub-section) were assumed to be good representative values of d.

Table 14, for example, indicates that in region E during season 4, the estimated arithmetic mean of the regional areal rainfall was 354.7 mm, giving a 10% value of 35.47 mm. The estimate for d given in equation (36) was therefore 35.47 mm;

Table 14: The Desired number of gauges (N) derived from the use of the Physiographic effects method (equation 36).

Region	Existing No. of Gauges	Season 1 (Dec.-Feb.)			Season 2			Season 3			Season 4			Mean Annual (Jan.-Dec.)			Optimum Regional Gauges (M_R)
		Mean R/F (mm)	S (mm)	N	Mean R/F (mm)	S (mm)	N	Mean R/F (mm)	S (mm)	N	Mean R/F (mm)	S (mm)	N	Mean R/F	S (mm)	N	
A1	9	248.5	86.56	50	557.5	117.14	19	171.5	35.84	18	306.2	88.73	35	1285.8	171.57	8	50
A2	1	368.0	177.46	96	690.0	204.54	37	261.0	131.32	105	356.0	174.85	100	1682.0	504.69	38	105
B	19	216.9	75.94	51	435.0	78.86	14	240.1	39.87	12	380.7	110.85	35	1272.8	171.74	8	51
C	6	215.2	58.14	30	341.0	77.75	22	106.5	38.80	55	324.1	70.13	20	989.2	132.11	8	55
D	7	183.4	82.67	84	514.0	104.60	18	298.7	53.87	14	383.9	121.78	42	1381.3	209.32	10	84
E	8	124.1	65.11	114	464.8	89.07	16	363.5	76.97	19	354.7	102.13	35	1306.5	135.45	5	114
F	6	136.1	81.20	147	488.9	104.17	19	394.1	64.53	12	304.6	94.48	40	1324.7	180.59	8	147
G	5	57.7	38.69	185	294.0	67.57	22	383.7	91.60	24	190.6	78.47	70	925.9	145.48	11	185
H	9	63.9	41.55	175	340.8	61.70	14	438.1	90.01	18	268.7	80.73	38	1111.2	132.12	7	175
I	11	95.3	47.97	105	412.8	68.91	12	440.9	71.98	11	403.9	98.66	25	1352.9	127.12	4	105
J	5	65.6	43.34	180	332.7	67.78	18	449.4	80.12	14	401.6	112.47	33	1249.6	153.7	7	180
K	3	96.4	54.99	134	337.9	69.41	18	298.9	72.72	25	368.6	111.78	38	1101.4	158.2	9	134
L	10	148.7	59.14	66	419.1	57.72	8	284.1	58.30	18	459.1	85.15	15	1310.8	129.61	5	66
M	6	192.6	76.60	66	384.6	75.75	16	249.5	58.22	23	445.0	89.48	17	1268.5	170.42	8	66
Totals				1483			253			368			543			136	1517

Note: R/F = Rainfall (mm); S = Standard deviation; N = desired number of raingauges.

while the sample standard deviation (S) was 102.12 mm. For a 95% confidence inverval ($\alpha = 0.05$) taking its upper limit for 36 years of observations, the students t-statistic with 35 degrees of freedom is about 2.032. When these values were substituted in equation (36) the desired number of raingauges (N) obtained was 35. The other values of N were as given in Table 14 for the various seasons/periods and homogeneous regions. It should be noted that in Table 14, the gauge estimates for regions A2 and K are based on the available 1 and 3 stations respectively.

Table 14 further shows that a maximum of 1517 raingauges would be required country-wide. This total number of 1517 was derived from the sum of the largest number of raingauges which were obtained at each of the homogeneous regions. Table 14 indicates that the highest number of stations included in the optimum network design was during season 1. As already discussed in the previous sections season 1 is generally dry over many parts of Uganda (Figure 15a). Rainfall during season 1 is due to local factors like land/lake breeze effects or orographic lifting. There is wide spatial and temporal variability in the localised rainfall generating systems. The seasonal areal amounts are therefore generally small (50 - 200 mm, Table 14). The definition of d (the acceptable 10% difference between the actual and the estimated rainfall that is measured by the sample raingauges used here) over any region is, therefore, generally small (about 5-20 mm). A dense raingauge network would therefore be required to detect these localized storms.

During seasons 2, 3 and 4 when the spatial rainfall variations are governed by synoptic scale systems, as discussed earlier in the text, fewer numbers of raingauges of 253, 368 and 543 respectively were generated with equation (36). The annual records generated the smallest optimum network of 136 gauges. This is because annual rainfall totals are substantial (generally over 1000 mm) and have relatively small spatial variation within the individual regions.

If the estimates of raingauges (N) for regions A2 and K are neglected as being based on inadequate data, it can be observed from the gauge estimates of the other regions that the drier zones with low seasonal/period rainfall amounts require denser raingague networks than the wetter ones. A good example is region G which is the driest part of Uganda.

It can, therefore be observed from the results of this method that the dry periods and regions need the highest raingague densities for the optimum raingauge networks. This is a disadvantage of this method because the economic value of the optimum raingauge network design in dry regions and of other areas of extreme climates like over the lakes, mountains and dense forests which have relatively low habitation is small compared to the network operating costs. It should also be noted that water dependent economic activities are limited in the dry areas and during the dry seasons making large numbers of raingauges uneconomical to operate.

3.4.2.2 OPTIMUM NETWORK DESIGNS DERIVED FROM COEFFICIENT OF VARIATION METHODS

Table 15A gives the results of the desired numbers of raingauges for the homogeneous regions as derived from the coefficient of variation model given in equation (37). A 10% coefficient of variation as the desired level of accuracy (as explained under section 3.4.2.1) was used in the derivation

The results indicate that in region M for example, during season 2, the existing coefficient of variation was about 19.7%. The desired level of accuracy is a coefficient of varition of 10%; the existing number of gauges (n) is 6 giving a raingauge network of about 24 stations. The number of

Table 15A: The Results from the use of coefficient of variation (model equation 37).

Region	Existing No. of rain-gauges (n)	Season 1 (Dec. - Feb.)		Season 2 (Mar - May)		Season 3 (Jun. - Aug.)		Season 4 (Sept. - Nov.)		Mean Annual (Jan. - Dec.)		Optimum regional No. of rain-gauges (M_R)
		Existing C_v	N	Existing C_v	N	Existing C_v	N	Existing C_v	N	Existing C_v	N	
A1	9	34.83	110	21.01	40	20.90	40	28.98	76	13.34	16	110
A2	1	48.22	24	29.79	9	50.31	26	49.12	25	30.01	10	26
B	19	35.01	233	18.13	63	16.61	53	29.12	162	13.49	35	233
C	6	27.02	44	22.80	32	36.43	80	21.64	29	13.36	11	80
D	7	45.08	143	20.35	29	18.03	23	31.72	71	15.15	17	143
E	8	52.47	221	19.16	30	21.17	36	28.79	67	10.37	9	221
F	6	59.66	214	21.31	28	16.37	17	31.08	58	13.63	12	214
G	5	67.05	225	22.98	27	23.87	29	41.17	85	15.71	13	225
H	9	65.02	381	18.10	30	20.55	39	30.04	82	12.43	14	381
I	11	50.34	279	16.69	31	16.33	30	24.43	66	9.40	10	279
J	5	66.07	219	20.37	21	17.83	16	28.01	40	12.30	8	219
K	3	57.04	98	20.54	13	24.33	18	30.33	28	14.36	7	98
L	10	39.77	159	13.77	19	20.52	43	18.55	35	9.89	10	159
M	6	39.77	95	19.70	24	23.33	13	20.11	25	13.43	11	95
Totals			2445		396		483		849		183	2483

Note: C_v = Coefficient of variation
N = Desired number of Raingauges for a given season
M_R = Optimum Regional Number of Gauges

raingauges generated in the network for seasons 1,3 and 4 and the annual period were 95, 33, 25 and 11 respectively. Details of the desired numbers of gauges (N) obtained at the other homogeneous regions with this method are given in Table 15A.

The requirement of a denser network of gauges during the dry seasons and at dry regions is still evident as observed with the physiographic effects method. However, the desired number of gauges (N) are much higher in Table 15A than in Table 14 for all regions except regions A2 and K. This is because the desired number of gauges (N) derived from equation [37] also depend on the exisiting number of gauges (n) in a region. Areas with large n values will therefore have denser networks with this model if their existing coefficients of variation is above 10%.

The raingauge networks which were obtained with the coefficient of variance models given in equations (38) and (39) are given in Tables 15B and 15C respectively.

It can be observed from Table 15B that although there were seasonal and spatial variations in the total numbers of derived gauges (N), the number of raingauges obtained from equation (38) were relatively fewer than those from equation (37). The large values fo N are, however, still concentrated in the dry areas.

Table 15B: The Results from the use of coefficient of variation (model equation 38, P = 10%).

Region	Existing No. of raingauges (n)	Season 1 (Dec. - Feb.)		Season 2 (Mar. -- May)		Season 3 (Jun. - Aug.)		Season 4 (Sept. - Nov.)		Mean Annual (Jan. - Dec.)		Optimum regional No. of raingauges (M_R)
		Existing C_v	N	Existing C_v	N	Existing C_v	N	Existing C_v	N	Existing C_v	N	
A1	9	34.83	13	21.01	5	20.90	5	28.98	9	13.34	2	13
A2	1	48.22	24	29.79	9	50.31	26	49.12	25	30.01	10	26
B	19	35.01	13	18.13	4	16.61	3	29.12	9	13.49	2	13
C	6	27.02	8	22.80	6	36.43	14	21.64	5	13.36	2	14
D	7	45.08	21	20.35	5	18.03	4	31.72	11	15.15	3	21
E	8	52.47	28	19.16	4	21.17	5	28.79	9	10.35	2	28
F	6	59.66	36	21.31	5	16.37	3	31.08	10	13.63	2	36
G	5	67.05	45	22.98	6	23.87	6	41.17	17	15.71	3	45
H	9	65.02	43	18.10	4	20.55	5	30.04	10	12.43	2	43
I	11	50.34	26	16.69	3	16.33	3	24.43	6	9.40	1	26
J	5	66.07	44	20.37	5	17.83	4	28.01	8	12.30	2	44
K	3	57.04	33	20.54	5	24.33	6	30.33	10	14.36	3	33
L	10	39.77	16	13.77	2	20.52	5	18.55	4	9.89	1	16
M	6	39.77	16	19.70	4	23.33	6	20.11	5	13.43	2	16
			366		67		95		138		37	374

Note: C_v = Coefficient of variation

N = Desired number of Raingauges for a given season

M_R = Optimum Regional Number of Gauges

Table 15C: The Results from the use of coefficient of variation (model equation 39, D = 10%).

Region	Existing No. of raingauges (n)	Season 1 (Dec. - Feb.)		Season 2 (Mar. - May)		Season 3 (Jun. - Aug.)		Season 4 (Sept. - Nov.)		Mean Annual (Jan. - Dec.)		Optimum regional No. of raingauges (M_R)
		Existing C_v	N	Existing C_v	N	Existing C_v	N	Existing C_v	N	Existing C_v	N	
A1	9	34.83	49	21.01	18	20.90	18	28.98	34	13.34	8	49
A2	1	48.22	94	29.79	36	50.31	102	49.12	97	30.01	37	102
B	19	35.01	50	18.13	14	16.61	12	29.12	34	13.49	8	50
C	6	27.02	30	22.80	21	36.43	54	21.64	19	13.36	8	54
D	7	45.08	82	20.35	17	18.03	14	31.72	41	15.15	10	82
E	8	52.57	111	19.16	15	21.17	18	28.79	34	10.35	5	111
F	6	59.66	143	21.31	19	16.37	11	31.08	39	13.63	8	143
G	5	65.02	180	22.98	22	23.87	23	41.17	68	15.71	10	180
H	9	65.02	170	18.10	14	20.55	17	30.04	37	12.43	7	170
I	11	50.34	102	16.69	12	16.33	11	24.43	24	9.40	4	102
J	5	66.07	175	20.37	17	17.83	14	28.01	32	12.30	7	175
K	3	57.04	131	20.54	17	24.33	24	30.33	37	14.36	9	131
L	10	39.77	64	13.77	8	20.52	17	18.55	14	9.89	4	64
M	6	39.77	64	19.70	16	23.33	22	20.11	17	13.43	8	64
Totals			1445		246		156		527		133	1477

Note: C_v = Coefficient of variation
N = Desired number of raingauges
M_R = Optimum Regional Number of Gauges

The results from Bleasdale method (equation (39)) which are given in Table 15C, indicates that when the value of the student t statistic in equation (36) is close to 2, equation (36) and equation (39) predict virtually similar desired numbers of raingauges (N) for a given confidence level. For example, at 95% confidence level, the value of the two sided student t statistic with 35 degrees of freedom is approximately 2.03. The predicted numbers of gauges (N) in Tables 14 and 15C are approximately the same.

Thus, in general, the coefficient of variation methods discussed above, like the method of physiographic effects discussed under section 3.4.2.2 predict high desired gauge densities during the dry seasons and in dry areas. This, as discussed earlier, is a disadvantage because given the limited levels of rain dependent economic activities during the dry seasons and in dry zones, large raingauge networks are uneconomical to run.

3.4.2.4 OPTIMUM RAINGAUGE NETWORK DESIGN DERIVED FROM ERROR OF INTERPOLATION METHOD

This method is based on the relationship between interstation correlation coefficients (r(x)) and distances between the stations (x). The functional relationships between r(x) and x as obtained from graphical and regression methods are independently presented in the following sections.

3.4.2.4.1 RESULTS FROM GRAPHICAL METHODS

Figures 33a and 33b are typical examples of the scatter diagrams of the relationship between r(x) and x which were obtained with the seasonal rainfall records. It can be observed from these figures that the relationships between r(x) and x are quite complex in Uganda. These can be associated with the high spatial and temporal variability of tropical showers and thunderstorms, topographically induced differences, microclimatic factors, etc. (Lifiga, 1979, Flitcroft et al, 1989). These factors introduce spurious values in correlation versus distance relationships. Similar patterns of interstation correlations versus distances for tropical rainfall have been osberved by Jackson (1969, 1977), Snijders (1985) and many others. The scatter diagrams illustrate some of the major difficulties which were faced with the use of free hand methods of finding a relationship between correlation coefficients and distances between stations.

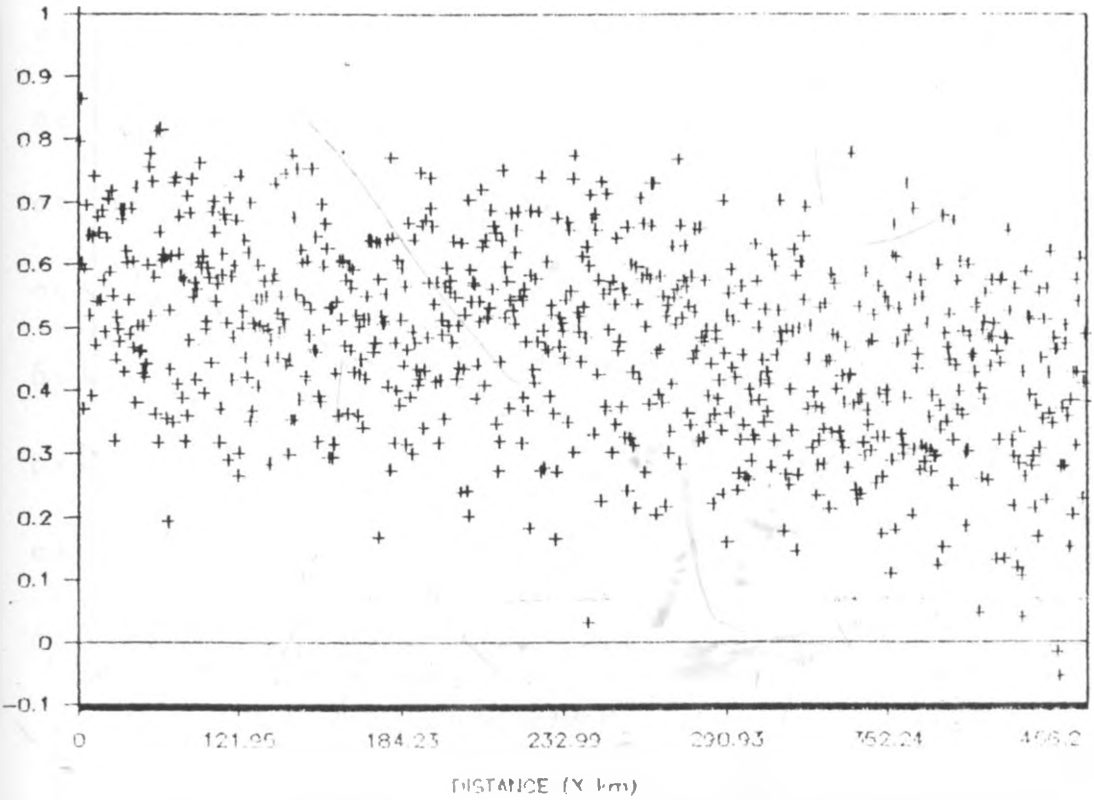


Fig. 33a: Examples of the relationship between inter-station correlation and distance during season 1.

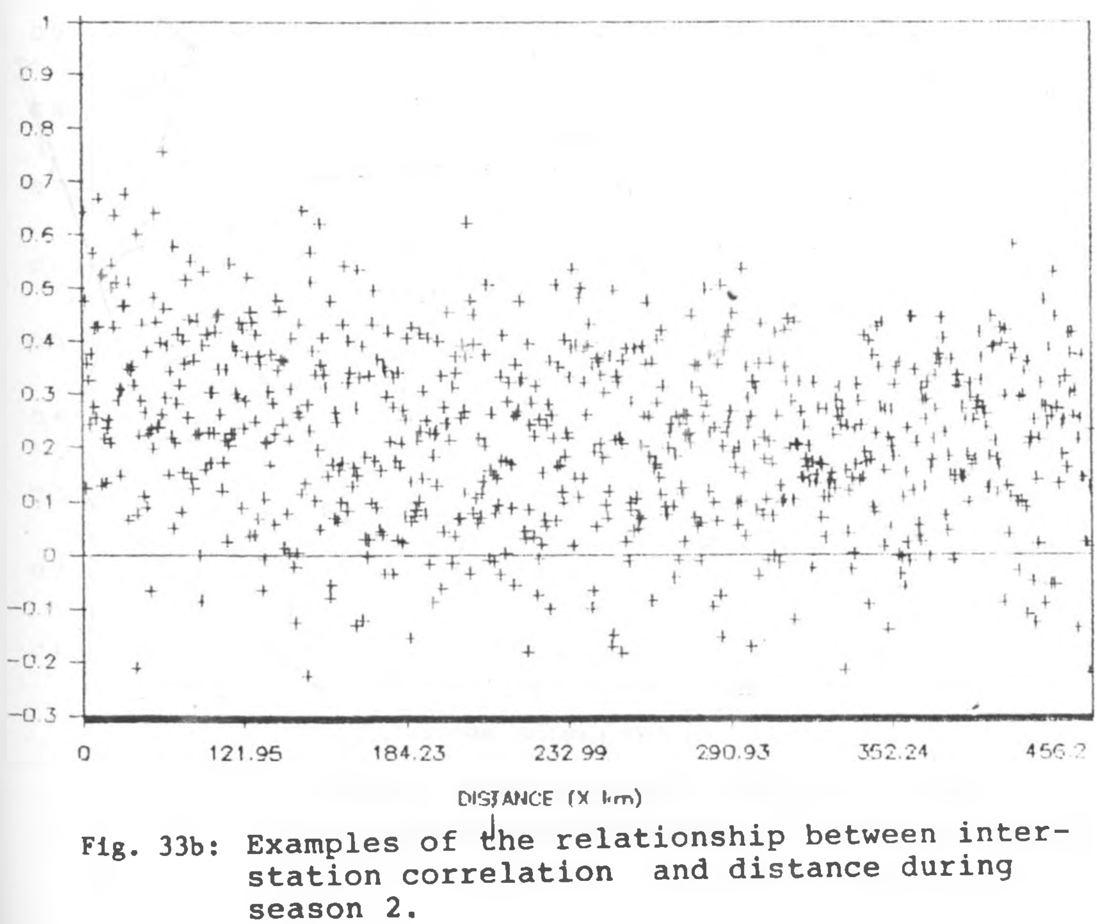


Fig. 33b: Examples of the relationship between inter-station correlation and distance during season 2.

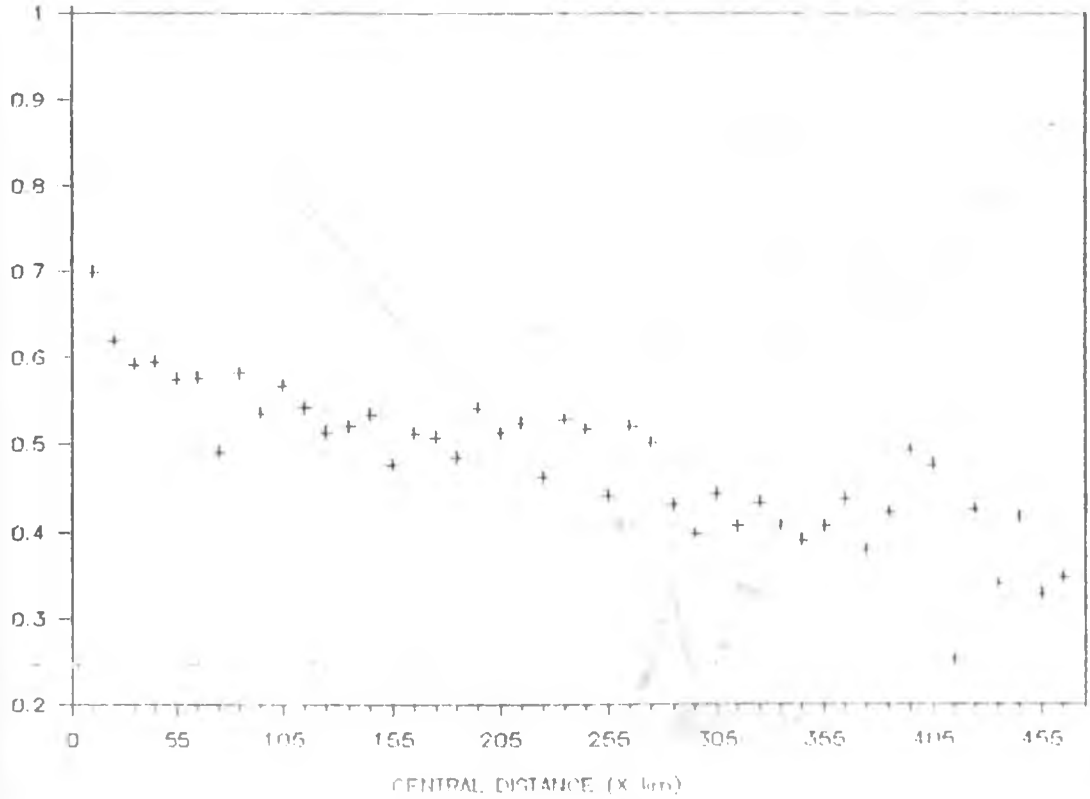


Fig.34a: Averaged relationship between interstation correlations and distances during season 1.

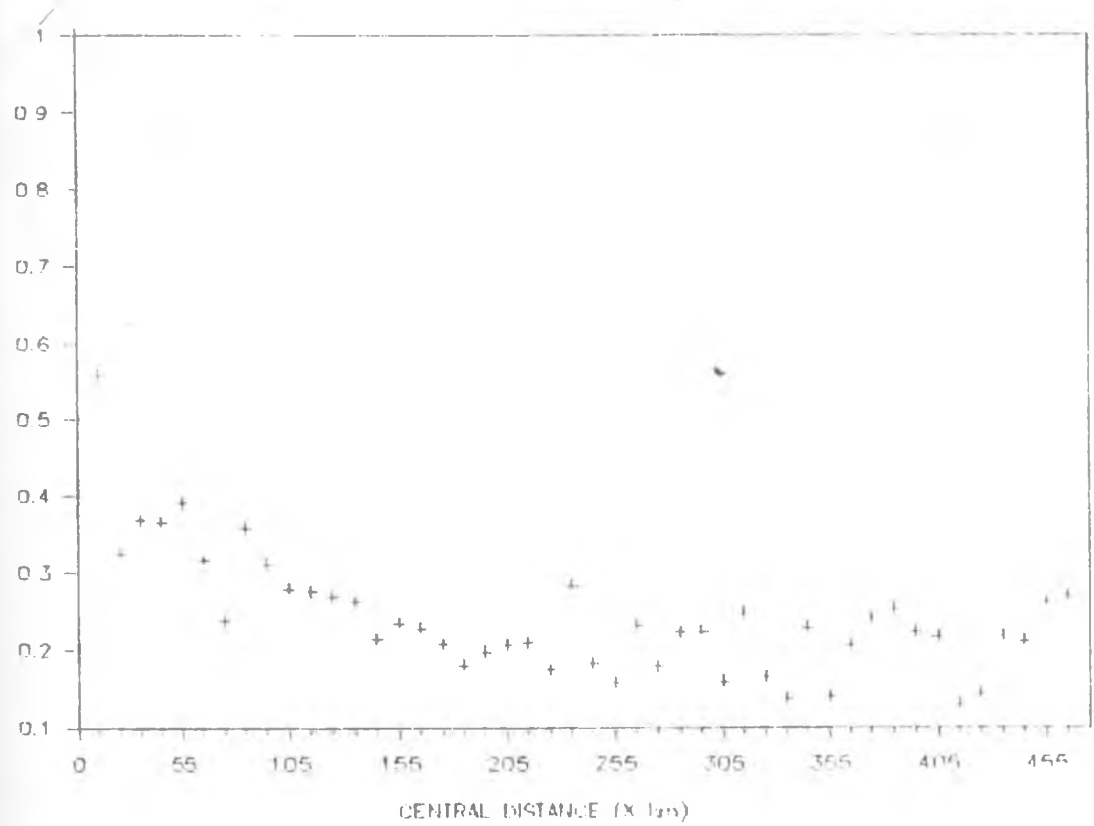


Fig.34b: Averaged relationship between interstation correlations and distances during season 2.

Typical examples of the scatter diagrams which were obtained when the averaged correlation estimates ($\hat{r}(x)$) were plotted against central distances (x) are given in Figures 34a and 34b. The decrease of correlation with distance is quite evident. However, it can be observed from these diagrams that it will be difficult to determine unique free hand curves through the scattergrams. Thus, regression methods discussed under sections 2.4.3.6.2 and 2.4.3.6.3 were used for further analyses. The results from these methods are presented in the next section.

3.4.2.4.2 REGRESSION METHODS RESULTS

The estimates of the parameters for the linear and exponential models which were obtained for the averaged correlation estimates ($\hat{r}(x)$) versus central distances (x) are given in Table 16. The random error of observation, instrument, microclimatological, etc. (η) are also indicated in the table.

Table 16: Parameter Estimates for the Linear and Exponential Regression models.

Season/ Period	Model	Constant (a)	Regression Coefficient (b)	Explained variance (r^2)	Computed $F_{1,7}$ value	Random errors η
Season 1	linear	0.9	- 0.0043	0.82	30.9	0.10
(Dec.-Feb)	Exponential	- 0.11	- 0.0057	0.86	42.1	0.10
Season 2	linear	0.8	- 0.0052	0.71	16.9	0.20
(Mar.-May)	Exponential	- 0.22	- 0.008	0.80	28.0	0.20
Season 3	linear	0.8	- 0.0048	0.60	10.4	0.20
(Jun.-Aug.)	Exponential	- 0.27	- 0.0075	0.68	14.8	0.24
Season 3	linear	0.9	- 0.0043	0.79	27.1	0.10
(Sept.-Nov.)	Exponential	- 0.13	- 0.0058	0.85	38.9	0.12
Mean Annual	Linear	0.8	- 0.004	0.65	13.3	0.20
(Jan.-Dec.)	Exponential	- 0.2	- 0.0054	0.73	18.55	0.18

Note: All constants (a) and regression coefficients (b) values were significant above 95% confidence levels.

The results given in Table 16 indicated that during the four standard seasons, the magnitudes of the random errors (η) are relatively larger during seasons 2 and 3. This may be associated with the complexity of the rainfall generating systems during these two seasons. It was observed earlier from the PCA results discussed under section 3.4.1.1 that relatively more eigenvectors were required to explain the spatial rainfall variance during seasons 2 and 3 than during seasons 1 and 4.

The results further indicated that all values of the slope coefficient (b) were negative. This indicated a decrease of interstation correlation coefficients with the increase in distance between stations.

All the parameter estimates of a and of b given in the table were significant above 95% confidence levels when the students t test is used. Analysis of variance techniques, however, indicated that the exponential model gave better relationships between the averaged correlation estimates ($\hat{r}(x)$) and the central distance (x). In all cases, the variance of the averaged correlation coefficient explained by the central distance were higher with the exponential model than with the linear model. The exponential model was therefore used to estimate the correlation versus distance relationship from which the optimum probability weighting estimates (p) and εopt (equation 44) were obtained.

Table 17 gives the values of εopt which were obtained with the exponential regression model parameter estimates in Table 16. The values presented are those of seasons 2 and 4, the two major rainy seasons. The mean annual values are also given in the table. The values of εopt generated from the exponential parameter estimates of seasons 3 and 1 were relatively similar in magnitude to those obtained for seasons 2 and 4 respectively and have therefore not been included in the table.

Table 17: Sample estimates of εopt values at various distances (x)

Distance (x km)	VALUES OF εopt		
	Season 2 Mar.-May	Season 4 Sept.-Nov.	Mean Annual Jan.-Dec.
0	0.36	0.23	0.33
10	0.39	0.25	0.35
20	0.42	0.28	0.37
30	0.45	0.30	0.39
40	0.47	0.33	0.41
50	0.50	0.35	0.43
60	0.53	0.37	0.45
70	0.56	0.40	0.47
80	0.58	0.42	0.49

It can be observed from Table 17 that season 2 records gives the largest values of εopt for any given distance (xkm) between stations. Thus, the largest gauge densities for reducing εopt to given levels are required during season 2. These observations may be due to the spatial and temporal rainfall characteristics which have already been given in the text. A raingauge network design for season 2 would therefore fulfil the raingauge requirements of all rainy seasons. Season 2 was therefore used to design the optimum network with the error of intepolation method.

Examples of the optimum probability weighting (p) values obtained from the study for season 2 are given in Table 18 at various distance (x).

Table 18: Sample Estimates of the optimum probability weighting (p) at various distances (x).

Distance (xkm)	0.	5	10	15	20	25	30	35	40	50	60	70	80
Weigh-ting (p)	0.40	0.40	0.40	0.40	0.39	0.39	0.39	0.39	0.38	0.38	0.37	0.37	0.36

It can be observed from Table 18 that $p < 0.5$ for all values of x including $x = 0$. This is due to the random errors (η) being greater than zero in all cases.

Once p values are known, equation (51) can be used to obtain the optimum rainfall estimates ($\hat{R}_{ij}$) for the various mid-points of observations stations.

It was noted under section 2.4.2.6 that εopt values should be standardized with $\hat{R}_{ij}$ as indicated in equation (49) to enable comparisons of the errors of interpolation within the various locations. The standardization requires the standard deviations (S) to be known at the various points. Figure 35 gives the spatial patterns of the standard deviation for season 2 mean rainfall totals which were used in the the study. The corresponding spatial distribution of the existing errors of interpolation expressed as percentages (γ) are given in Figure 36.

[illegible] be observed from Figure 36 that relatively high values of γ were concentrated near large water bodies and over dry areas. These areas were also relatively sparsely gauged. Low γ values concide with well gauged areas. The value of γ ranged from 20% to 45%.

The results shown in Figure 36 can be used to identify areas where more gauges could be added in order to reduce the existing percentage error of interpolation (γ). Thus, more gauges would be [illegible] the dry areas [illegible] around water bodies. The percentage errors of interpolation [illegible] in the well gauged

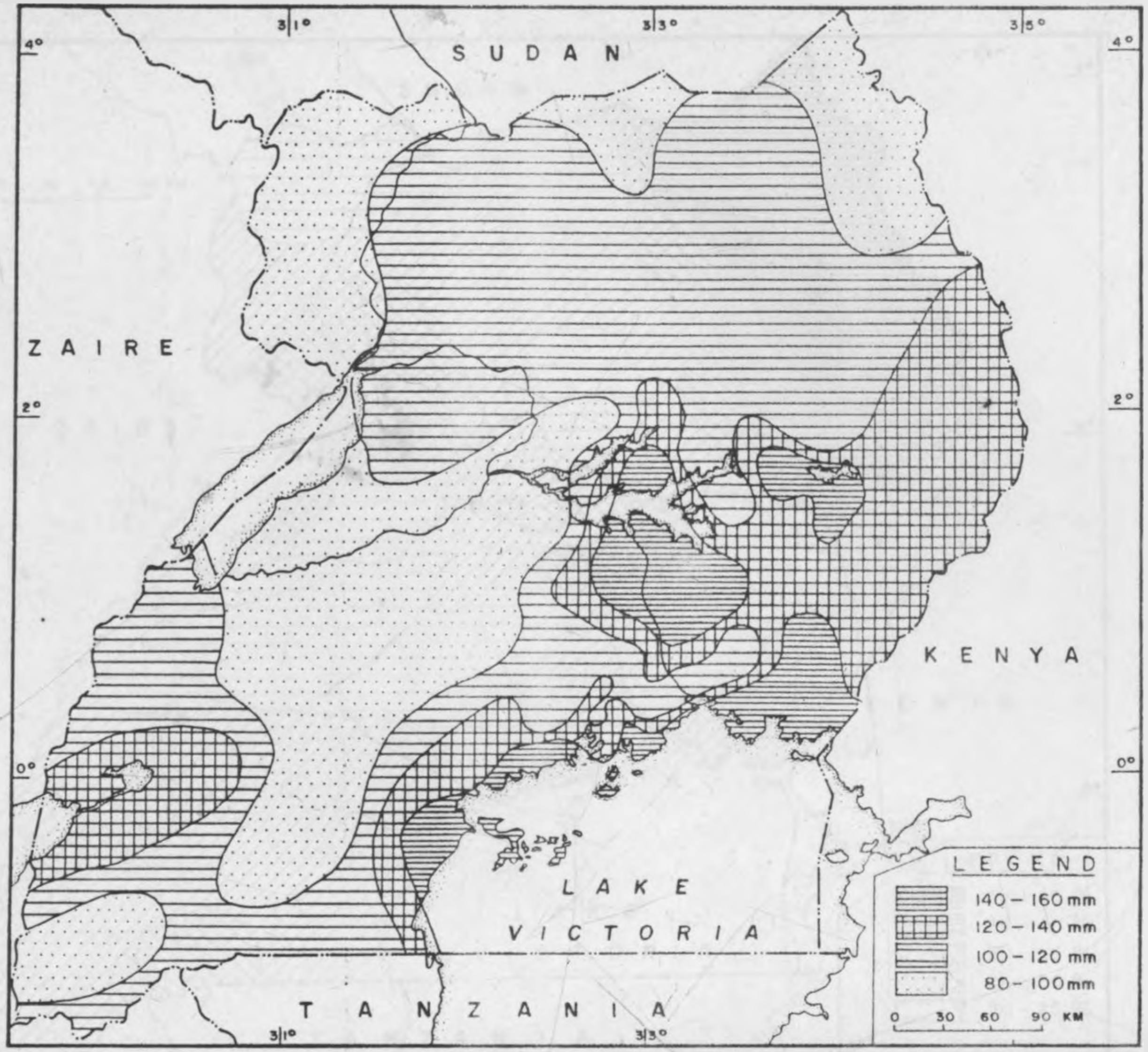


Fig. 35 : The standard deviations of season 2

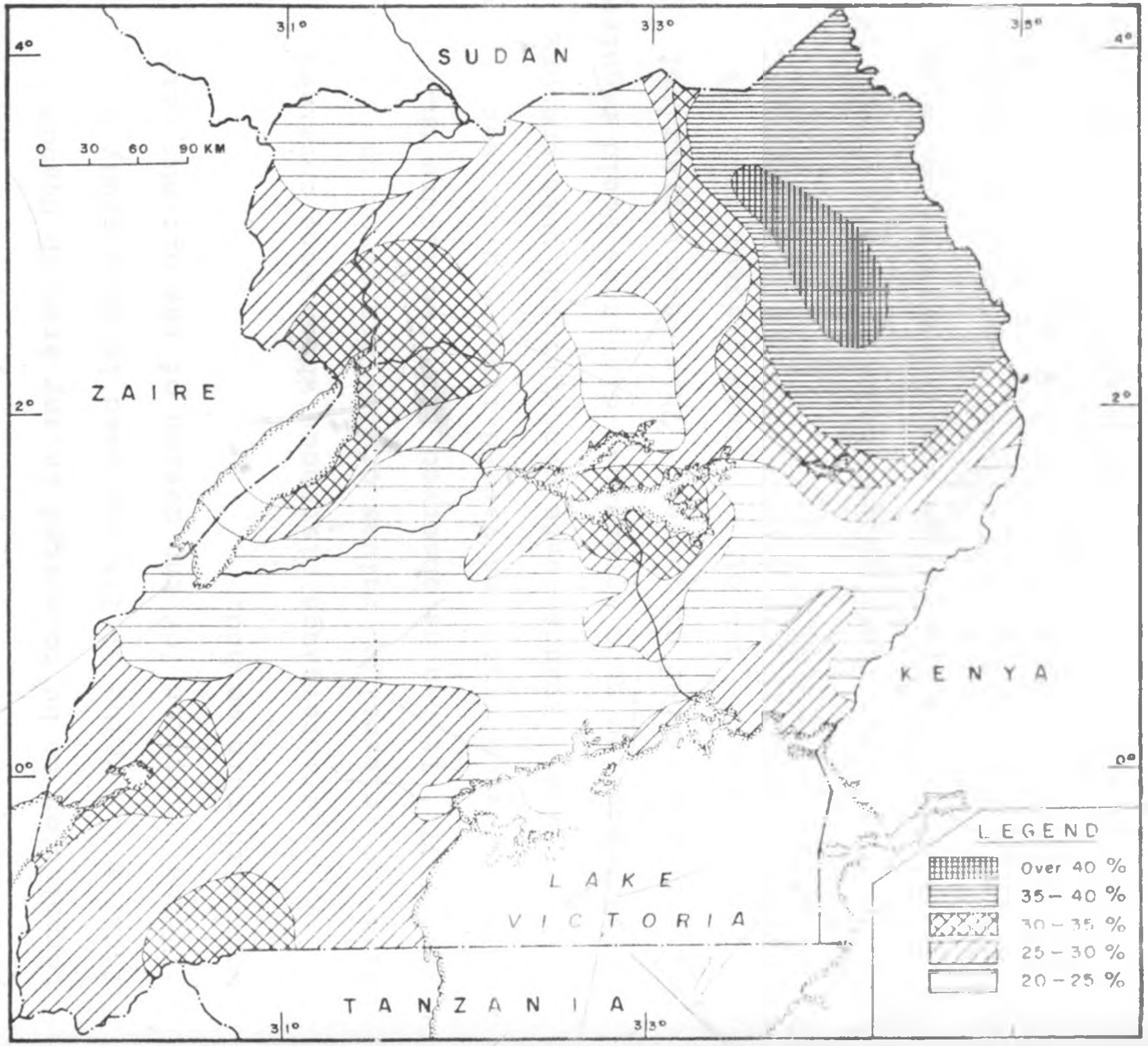


Fig. 36: Existing percentage errors of interpolation for season 2.

areas ranged from 20% to 25%. The upper limit of γ = 25% can therefore be used as a realistic expectation of the maximum percentage error of interpolation to be tolerated in any area in Uganda. Thus, a γ value of 25% was used in this study to fix the criterion for the design of the optimum network with this method.

The intergauge distances which were obtained with the critical γ values of 25% are given in Figure 37. It can be observed from the figure that the intergauge distances varied from zero to over [illegible] km. Zero distances were concentrated over the dry areas. Meeting zero intergauge distances would require an infinite number of gauges. This is not practical and suggests the use of γ values greater than 25% in these areas.

However, the results shown in Figure 37 can be used to derive actual gauge requirements by choosing some realistic distancing between gauges. For example, 120 stations would be required over Uganda with this method if interstation distance is fixed to 50 km.

One of the disadvantages of this method of design is the need to describe the functional relationship between interstation correlations and the distances between the stations. This relationship is quite complex in Uganda like many other tropical areas due to the regional and local factors discussed

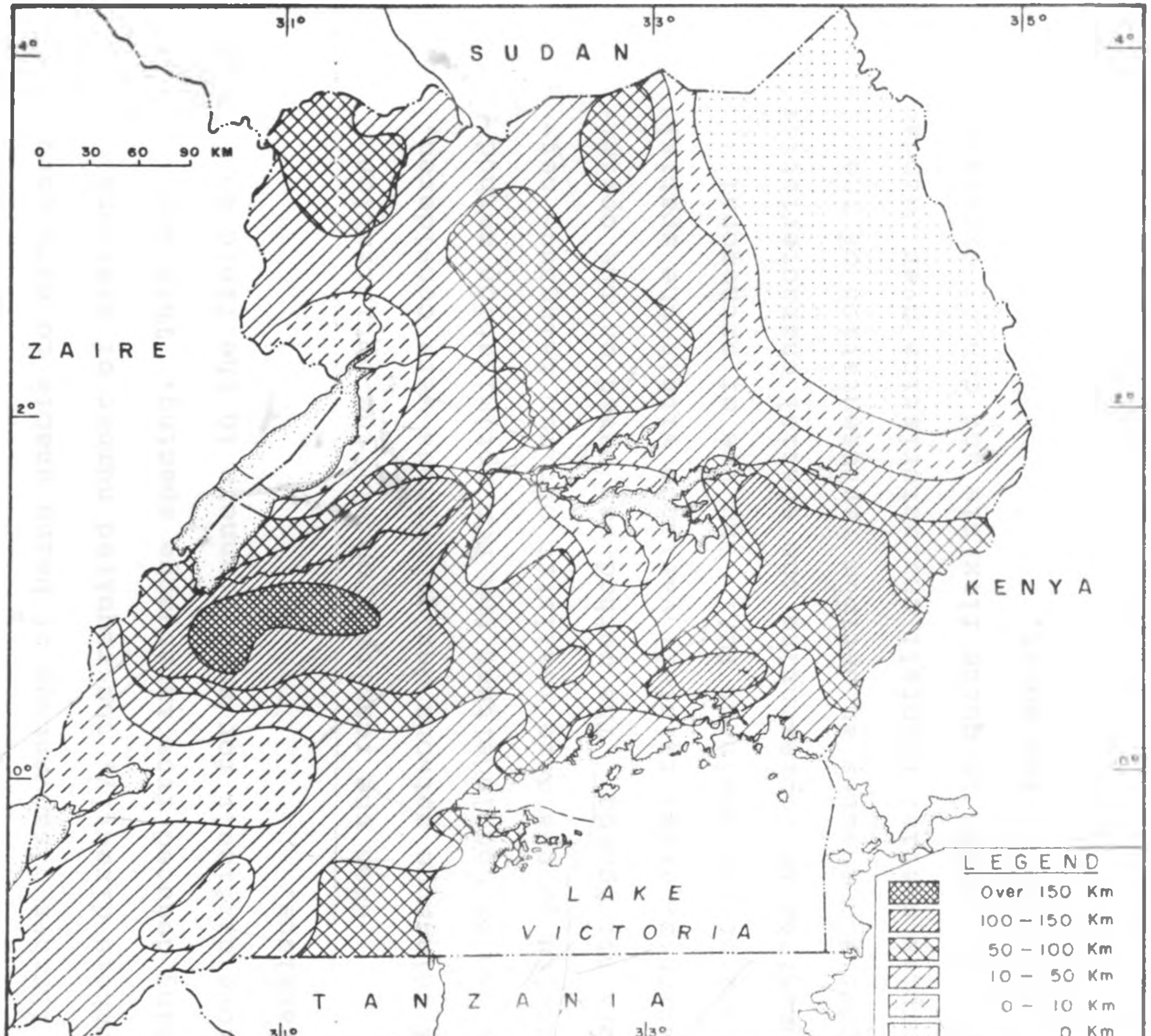


Fig. 37: Intergauge distance based on γ = 25%
during season 2

earlier in the text. Simplifications of the relationship with averaged correlation values cannot be expected to give a good representation of the relationship for all locations of Uganda. The method has a further disadvantage of being unable to give gauge numbers directly, the required number of stations being deduced from intergauge spacing. This may prove difficult for some planners in the field to translate.

When the results from all the optimization of raingauge network design methods were considered, it can be concluded that the PCA optimization method results gave a more realistic optimum raingauge network for Uganda. The method was based on the optimization of areal rainfall variance within the spatial modes of the dominant PCA. The use of orthogonal functions in representing the spatial characteristics under PCA ensures an optimum representation of the complex spatial rainfall characteristics over Uganda. The PCA method is quite flexible and can be extended to any part of the world.

CHAPTER 4

4.0 SUMMARY AND CONCLUSION

In this chapter, an attempt is made to summarize the results derived from this study. The conclusions that may be drawn from the results are also included. Finally, some future, possible extensions of the research are suggested.

4.1 SUMMARY OF THE RESULTS

The results from the study indicate that most of the records were homogeneous signifying the quality of the estimated missing records which were not declared heterogeneous.

The results from principal component analysis (PCA) delineated Uganda into 14 homogeneous rainfall regions. The number of significant factors extracted by PCA varied significantly from season to season. The number of significant factors ranged from four (4) for the mean climatological records to sixteen (16) for the month of September explaining 65% and 85% of the rainfall variances respectively. The high seasonal variability in the proportions of variance accounted for by the significant factors has been attributed to the high degree of spatial variability in the characteristics of the regional synoptic scale systems that control the rainfall.

The minimum network design based on the PCA spatial modes consisted of fourteen (14) stations. The fourteen (14) stations would form the basic nextwork for the synoptic stations in Uganda. The fourteen stations which were included in the minimum network design were: Buwunga, Buvuma, Kivulu, Mbarara, Vukula, Serere, Mukongoro, Kotido, Kitgum, Gulu, Yumbe, Nebbi, Kijura and Kisomoro. Only two of these stations, namely: Mbarara and Gulu are included in the current network of synoptic stations of Uganda. These have been started at national administrative centres or towns. Thus, the current network of synoptic stations may not give a good representation of the climatology of Uganda.

Analysis of variance techniques indicated that the variance of areal rainfall explained by the minimum network design stations at the individual homogeneous regions varied significantly from season to season. This has been attributed to the high temporal variability of the rainfall over Uganda.

The generated optimum raingauge network designs gave 89 stations for the PCA method; 1517 stations for the physiographic effects method, 2483, 374 and 1477 stations for the three coefficient of variation methods and 120 stations for the optimum error of interpolation method. The largest networks were obtained with the physiographic effects, the three coefficient of variation and optimum error of interpolation methods but most of these stations were located in the dry areas and near water bodies. More stations were also required during the dry seasons with these methods.

The network generated from the PCA method of optimization was considered as the most appropriate optimum network design for Uganda due to the fundamental principles of PCA which enable most of the spatial characteristics of the rainfall to be considered in the network design. The method can be extended to any part of the world due to the flexibility of the characteristics of empirical orthogonal functions.

4.2 CONCLUSIONS

It may be concluded from the study that Uganda can be delineated into fourteen (14) homogeneous climatological regions. Given that agriculture, the backbone of Uganda's economy and other rain water dependent activities have to be planned according to the rainfall characteristics in order to realize maximum benefits, the information derived from the study will be quite useful in their planning and management processes.

The study has also provided a climatologically representative minimum network design consisting of fourteen (14) stations based on at least one gauge at each homogeneous region. This network may be used to form the minimum synoptic network. Many more stations, however, may be required to give optimum rainfall informations. The suggested optimum network for Uganda consisted of 89 stations.

Finally, a new method of optimization of a raingauge network design based on the principles of the maximization of the areal rainfall variance explained within a PCA homogeneous climatological region is being recommended. This method can be extended to any region due to the flexibility in the characteristics of the empirical orthogonal functions.

4.3 SUGGESTED FUTURE WORK

An optimum raingauge network should provide reliable areal rainfall estimates which may be used in obtaining effective rainfall estimates. Effective rainfall, estimated from the ratio of rainfall to potential evapotranspiration, can be used in the development of regional drought indices. These drought indices can then be used as an early warning of drought using autoregressive models.

Attempts can also be made to develop some crop weather models using the characteristics of effective rainfall as the input and crop production as the output within homogeneous climatological regions.

The regional areal rainfall estimates may also be used in hydrological studies to determine available surface water resources. The rainfall can be used as the input and surface runoff estimates as the output in these models.

In areas of extreme climates, use could be made of satellite derived rainfall estimates to avoid difficulties associated with gauge establishment and their maintenance. The reliable rainfall estimates in the gauged areas can further be used to calibrate the satellite data.

ACKNOWLEDGEMENTS

I wish to extend my sincerest thanks to my supervisors, Prof. L.J. Ogallo and Dr. F.M. Mutua for their guidance, discussions and invaluable suggestions throughout the various phases of this work My thanks also go to those members of the Meteorology and other Departments both within and outside the University who assisted me in one way or another especially Prof. J.K. Ng'ang'a, Prof. J.W. Odhiambo of Mathematics Department and my friend Mr. G.O. Ouma for his kind understanding.

I also wish to thank Mrs. G.N. Macharia for her patience in typing this work and Mr. S. Okumu for drafting the diagrams.

Finally, and most importantly, I am grateful and thankful to my dear wife Flora, without whose patience, kind understanding and the burdens of care for the little ones, this work would not have been possible.

This study was financed wholly by a generous grant from the German Academic Exchange Service (DAAD) to whom I am greatly indebted. I am also indebted to my employers, Makerere University, Kampala, for granting me the study leave throughout the duration of the study.

REFERENCES

Ahuja, P.R. (1959): Planning of a Precipitation Network for Water Resources Development in India. Hydrologic Networks and Methods. Transaction of the Inter-Regional Seminar on Hydrologic Networks and Methods, Bangkok, Thailand. WMO/ECAFE, 106-112.

Akin, J.E. (1971): Calculation of Mean Areal Depth of Precipitation. J. Hydrl. 12, 363-376.

Alaka, M.A. (1970): Theoretical and Practical Considerations for Network Design. Met. Monogr. No. 33, 1970. 20-27.

[illegible]a, E.K. (1983): On the Monthly Mean Lower Tropospheric Circulation and the Anomalous Circulation During the 1961/62 Floods in East Africa. M.Sc. Thesis, University of Nairobi, 1983.

Arkin, P.A. and Meisner, B.N. (1987): Relationships between Large Scale Convective Rainfall and Cold Cloud over Western Hemisphere During 1982-84. Mon. Wea. Rev., 115, 51-74.

Asnani, G.C. and Kinuthia, J.K. (1979): Diurnal Variations of Precipitation in East Africa. Kenya Met. Dept. Res. Report 8/79, 58 pp.

Balek, J. and Holecek, J. (1965): The Influence of Density of Ombrographic Network on the Accuracy of Hydrologic Computation, Symposium on Design of Hydrologic Networks WMO/IASH, Quebec 15-6-22-6-1965.

Barnston, A.G. and Livezey, R.E. (1978): Classification, Seasonality and Persistence of low Frequency Atmospheric Patterns. Mon.Wea.Rev.VOL. 115, 1083-1126.

Barret, E.C. (1970): The Estimation of Monthly Rainfall From Satellite Data. Mon.Wea.Rev., 98, 322-327.

Barret, E.C., D'souza, G. and Power, C.H. (1986): BIAS: Interactive Satellite-Aided Rainfall Monitoring Method. Journal of British Interplanetary Society, 39, 517-326.

Bärring, L. (1987): Spatial Patterns of Daily Rainfall in Central Kenya: Application of Principal Component Analysis, Common Factor Analysis and Spatial Correlation. J. of Climat. VOL. 7, 267-289.

Bärring, L. (1988): Regionalization of Daily Rainfall in Kenya by Means of Common Factor Analysis. J. of Climat. VOL. 8, 371-389.

Barry, R.G. and Chorley, R.J. (1976): Atmosphere, Weather and Climate. LBS, 432 pp.

Basalirwa, C.P.K. (1979): Estimation of Areal Rainfall in some Catchments of the Upper Tana River. M.Sc. Thesis, University of Nairobi, 1979.

Basalirwa, C.P.K. (1984): Raingauge Network Design for a Catchment Using Annual Rainfall Totals. M.Sc. (Hydrology) Thesis, University College Galway, Ireland.

Battan, L.J. (1976): Z-R Relation. J. of Appl. Meteorology (1976): 1120-1121.

Blasing, R.J. and Lofgren, G.R. (1980): Seasonal Climatic Anomaly Types for the North Pacific Sector and Western North America. Mon.Wea.Rev., 108, 700-719.

Bleasdale, A. (1965): Raingauge Network Development and Design with Special Reference to the United Kingdom. Symposium on Design of Hydrological Networks. WMO/IASH, Quebec. 15-6-22-6, 1965. 146-157.

Bras, R.L. and Rodriguez-Iturbe, I. (1976): Network Design for Estimation of Areal Mean Rainfall Events. Water Resources Research. VOL.12, No.6, 1185-1195.

Buishand, T.A. (1982): The Analysis of Homogeneity of Long Term Rainfall Records in Netherlands. Wetenschappelijk, Rapport 81-87 De Bilt.

Chadwick, A.F., Dugdale, G. and Milford, J.R. (1986): Operational Rainfall Mapping over the Sahel. EUMETSAT, 6th Meteosat. Scientific Users Meeting, Amsterdam, NOv., 1986.

Child, D. (1970): Factor Analysis. Holt, Rinehert and Winston, Ltd., 1970. 107 pp.

Chidley, T.R.E. and Keys, K.M. (1970): A Rapid Method of computing Areal Rainfall. J. Hydrol. 12, 15-24.

Chua, S. and Bras, R.L. (1982): Optimal Estimators of Mean Areal Precipitation in Regions of Orographic Influence. J. of Hydrol. 57 (1982): 23-48.

Cislerova, M. and Hutchnison, P. (1974): The Redesign of the Raingauge Network of Zambia. Hydro. Sciences Bull. XIX, 4, 12/1974, 423-434.

Cohen, S.J. (1983): Classification fo 500mb Height Anomalies Using obliquely Rotated Principal Components. J. of Appl. Meteor. VOL. 22, No. 12, 1975-1988.

Conrad, V. and Pollack, L.W. (1962): Methods of Analysis In Climatology. Harvard University Press, Cambridge Massachusetts, 1950. 459pp.

Craddock, J.M. and Flood, C.R. (1969): Eigenvectors for representing the 500mb Geopotential Surface over the Northern Hemisphere. Quart. J. Roy. Met. Soc., 95, 576-593.

Daily Nation (1988): Focus on Uganda Nationspapers Ltd. July, 30,1988.

Delfiner, P. and Delhomme, J.P. (1973): Optimum Interpolation by Kriging. in Davis, J.C. and Mc Cullagh (Editors), Display and Analysis of Spatial data. Wiley london, 96-114.

Dhonneur, G. (1974): Nouvelle Approche des Realities Meteorologiques de l' Afrique Occidental et Centrale. ASECNA, Dakar, VOL. 2.

Dugdale, G. and Milford, J.R. (1987): Rainfall Mapping over Southern Africa in 1987. Report to FAO, Remote Sensing Centre, Agr. Division Project GCP/INT/432/NET.

Dymond, J. (1982): Raingauge Network Reduction. J. of Hydrol. 57, 81-91

EAMD (1962): Climatic Seasons of East Africa. E.A. Met. Dept. Report NO.8, 4pp.

Ebdon, D. (1977): Statistics in Geography. Basil Blackwell Billing and Sons Ltd., Worcester, 1977, 232 pp.

Edwards, K.A. (1972): Estimating Areal Rainfall by Fitting Surfaces to Irregularly Spaced Data. in Distribution of Precipitation in Mountainous Areas WMO/OMM NO. 326; 565-587.

Findlater, J. (1968): The Month to Month Variation of Winds at Low Level over Eastern Africa. EAMD Tech. Memo. No. 12, 27 pp.

Findlater, J. (1969): A Major Low Level Air Current Near the Indian Ocean During the Northern Summer. Quart. J. Roy. Meteor. Soc., 95, 362-380.

Flemming, D. (1970): Notes on Easterly Disturbances Affecting East Africa 5-7 Sept. 1967. E.A. Met. Dept. Tech. Memo. 13, 13pp.

Flitcroft, I.D, McDougall, V.D., Milford, J.R. and Dugdale, G. (1986): The Calibration and Interpretation of Meteosat Based Estimates of Sahelian Rainfall. 6th Meteosat Sci. Users Meeting, Amsterdam, Nov., 1986.

Flitcroft, I.D., Milford, J.R. and Dugdale, G. (1989): Relating Point to Area Average Rainfall in Semi-Arid West Africa and the Implications for Rainfall Estimates Derived From Satellite Data. American Meteor. Society, J. of Appl. Meteor. 252-256.

Gandin, L.S. (1970): The Planning of Meteorological Networks. Tech. Note No. III, WMO, Geneva, 1970, 35 pp.

Ganguli, M.K. Rangarajan, R. and Pachang, G.M. (1951): Accuracy of Mean Rainfall Estimates. Data of Damodar Catchment. Irrigation and Power Journ. VOL. 8, 278-284.

Glasspoole, J. (1922): A comparison of the Fluctuations of the Annual Rainfall over the British Isles. British Rainfall, 62, 260-266.

Glasspoole, J. (1928): The Rainfall of Norfolk. British Rainfall 1928, 271-272.

Glover, J., Robinson, P. and Henderson, J.P. (1954): Provisional Maps of Reliability of Rainfall in East Africa. Quart. J. Roy. Meteor. Soc. 80, 602-609.

Gregory, S. (1954): Annual Rainfall over Areas of Southern England. Quart. J. Roy. Meteor. Soc., VOL. 80, 610-618.

Gregory, S. (1963): Statistical Methods and the Geographer. Longman Group Ltd. The Chaucer Press, Bungay, Suffolk, 1963. 240 pp.

Gregory, S. (1965): Rainfall over Sierra Leone. Dept. of Geogr. University of Liverpool, Research Paper NO. 2, 58 pp.

Gregory, S. (1975): On the Delimitation of Regional Patterns of Recent Climatic Fluctuations. Weather 30: 276-287.

Griffiths, J.F. (1958): Climatic Zones of East Africa. East African Agric. J. 23, No.3.

Griffiths, J.F. (1969): East Africa: Its People and Resources. Edited by W. Morgan: Oxford University Press, 1969.

Griffiths, J.F. (1972): Eastern Africa. In Climates of Africa, H.E. Landsberg, Editor in Chief. World Survey of Climatology. VOL. 10 Elsevier, 604 pp.

Haan, C.T. (1977): Statistical Methods in Hydrology. The Iowa State University Press/Ames. 1982, 378 pp.

Hanna, L.W. (1975): Potential Water Deficits in Uganda: An Assessment of Wet and Dry Seasons." Trans. of Inst. of Brit. Geogr. VOL.1 No. 21976, 190-202.

Harman, H.H. (1976): Modern Factor Analysis. The University of Chicago Press. Chicago and London, 1976, 487 pp.

Harnett, D.L. and Murphy, J.L. (1980): Introductory Statistical Analysis. Addison-Wesley Publishing Co. London, 1980, 594 pp.

Hartigan, J.A. (1975): Clustering Algorithms. John Wiley and Sons. New York, 1975, 351 pp.

Helvey, J.D. and Patrick, J.H. (1965): Design Criteria for Interception Studies. Symposium on Designs of Hydrological Networks. VOL. 1 Quebec 15-6-22-6, 1965, WMO/IASH, Pub. No.67, 72-79.

Hielkema, J.U., Prince, S.D. and Astle, W.L. (1986): Rainfall and Vegetation Monitoring in the Savanna Zone of the Democratic Republic of Sudan Using NOAA Advanced Very High Resolution Radiometer. Int. J. Remote Sensing, 1986, VOL. 7, No. 11, 1499-1513.

Hutchinson, P.(1974): The Redesign of the Climatological Network of Zambia. Int. Trop. Meteorology Meeting, American Meteor. Soc./E.A. Met. Dept./WMO Jan. 31- Feb. 7, 1974, 21-26.

Ininda, J.D.M. (1985): Spatial and Temporal Characteristics of Drought in Eastern and Southern Africa. M.Sc Thesis, University of Nairobi, 1985.

Jackson, I. (1969): The Persistence of Rainfall Gradients over Small Areas of Uniform Relief. The East African Geogr. Review NO.7, April, 1969, 37-43.

Jackson, I. (1977): Climate, Water and Agriculture in the Tropics. Longman, 1977, 248 pp.

Janowiak, J.E., Krueger, A.F. and Arkin, P.A. (1985): NOAA Atlas NO.6, Atlas of Outgoing Longwave Radiation Derived from NOAA Satellite Data. U.S. Dept. of Commerce, 44 pp.

Johnson, D.H. and Mörth, H.T. (1961a). Some Trans-Equatorial Contour Analyses. Memoirs, VOL. III No.8 Nairobi. E.A. Met. Dept. 1961.

Johnson, D.H. and Mörth, H.T. (1961b). Forecasting Research in East Africa. Memoirs, VOL. III No.9, Nairobi. E.A. Met. Dept. 1961.

Johnson, D.H. (1962): Rainfall in East Africa. Quart. J. Roy. Met. Soc., 88, 1-19.

Johnson, D.M. (1980): An Index of Arizona Summer Rainfall Developed through Eigenvector Analysis. J. Appl. Meteorol., 19, 849-856.

Justice, C.O., Holben, B.N., Gwynne, M.D. (1986): Monitoring East African Vegetation Using AVHRR Data. Int. J. Remote Sensing, 1986, vol.7, No.11, 1453-1474.

Kagan, R.L. (1972a): Precipitation-Statistical Principles. Casebook on Hydrological Network Design Practice. WMO-No.324, I-1.1-9.

Kagan, R.L. (1972b): Planning the Spatial Distribution of Hydrometeorological Stations to Meet Error Criterion. WMO No. 324, III - 1.2-1-8.

Kagenda-Atwoki (1975): A factor Analytic Approach to Delimitation of Rainfall Regions of Uganda. The East Afr. Geogr. Rev. No.13, 9-36.

Kagenda-Atwoki, (1976): Trend Surface Models for Estimating Rainfall for Successive Ten-Day in Uganda. The East Afr. Geogr. Rev. No.14 June, 1976, 35-60.

Kaiser, H.F. (1959): Computer Program for Varimax Rotation in Factor Analysis. Educ. Pysch. Meas., 19, 413-420.

Kendall, M.G., (1976): Time Series. London, Charles Griffin.

Kohler, M.A. (1958). Design of Hydrological Networks. WMO Tech. Note No.25, WMO-No.82, TP.32.

Kraus, E.B. (1955): Secular Changes of East Coast Rainfall Regimes. Quart. J. Roy. Met. Soc. 81, 430-439.

Krishnamurti, T.N. and Ogallo, L.J. (1989): Recent African Climate (1979-1984). Dept. of Met. Florida, 32306, Nov. 1989, Report No. FSU-89-13.

Langbein, W.B. (1972): Introduction to Casebook on Hydrological Network Design Practice. WMO-NO.324. VOL.I. 1972.

Lethbridge, M. (1967): Precipitation Probability and Satellite Radiation Data. Mon. Wea. Rev., 95, 487-490.

Lifiga, U.B. (1979): Some Mesometeorological Aspects of Nairobi Storms. M.Sc Thesis, University of Nairobi, 1979.

Lund, I.A. (1963): Map-Pattern Classification by Statistical Methods. J. Appl. Meteor. 2, 56-65.

Maronna, R. and Yohai, V.J. (1978): A Bivariate Test for the Detection of a Systematic Change in the Mean. J. Amer. Assoc. 73, p.640.

Mill, H.R. (1908): Map Studies of Rainfall. Quart.J. Roy. Met. Soc. 34, 65-86.

Nahikian, H.M. (1966): Topics in Modern Mathematics. The McMillan Company, London, 1966. 264 pp.

Nicholson, S.E. (1986): The Spatial Coherence of African Rainfall Anomalies: Inter Hemispheric Teleconnections. J. Appl. Met., 25, 1365-1381.

Nicks, A.D. (1965): Field Evaluation of Raingauge Network Design Principles. Symposium on Design of Hydrological Networks. WMO/IASH, Quebec. 15-6-22-6, 1965, 82-93.

Nie, N.H., Bent, D.H. and Hull, C.H. (1970): Statistical Package For Social Sciences (SPSS). Compiled by Glasgow University McGraw-Hill Book Co. (U.K.), 1970.

Nieuwolt, S. (1978): Rainfall Variability and Drought in East Africa. Erdkunde VOL. 28, 81-88.

NOAA, OFDA and AID (1986): An Introduction to NOAA Satellite Assessment Technology for Crops and Range Lands. NOAA, NESDIS, AISC, Nairobi, Kenya, May, 1986.

North, G.R., Bell, T.L. and Cohalan, R.F. (1982): Sampling Errors in Estimation Empirical Orthogonal Functions. Mon. Wea. Rev. VOL. 110, 699-706.

O'Connel, P.E., Beran, M.A., Gurney, R.J., Jones, D.A. and Moore, R.J. (1977). Methods for Evaluating the U.K. Raingauge Network. Natural Environment Research Council, Report NO.40, 1977. Institute of Hydrology, Wallingford Oxon.

Ogallo, L.J. (1980): Regional Classification of East African Rainfall Stations into Homogeneous Groups Using the Method of Principal Component Analysis. Stat. Climatology. Devel. in Atmos. Sci. 13, 255-266.

Ogallo, L.J. (1981): The Nature of Homogeneity in Rainfall Records over East Africa. Res. Report No. 4/81, Kenya Met. Dept. 1981, 18 pp.

Ogallo, L.J. (1983): Temporal Fluctuations of Seasonal Rainfall in East Africa. Mausam, 35, 175-180.

Ogallo, L.J. (1986): Stochastic Modelling of Regional Annual Rainfall Anomalies in East Africa. J. of Appl. Stat. VOL. 13 No.1, 1986, 49-56.

Ogallo, L.J. (1987), climate Data Processing. WMO Regional Training Seminar for the National Intructors for Regions I and IV, 26-10-6-11-1987. Niamey, Niger.

Ogallo, L.J. (1988): Climatology of Rainfall in East Africa. WMO TMPR No.28, 136-142.

Ogallo, L.J. Janowiak, J.E. and Halpert, M.Sc. (1988). Teleconnection Between Seasonal Rainfall over East Africa and Global Sea Surface Temperture Anomalies. Met. Soc. of Japan. 807-822.

Ogallo, L.J. (1989): The Spatial and Temporal Patterns of the East African Seasonal Rainfall Derived from Principal Component Analysis. International J. of Climatology, VOL. 9, 145-167.

Oludhe, C. (1978): Statistical Characteristics of Wind Power in Kenya. M.Sc Thesis. University of Nairobi, 1987.

Ouma, G.O. (1988): Estimation of Areal Rainfall from Satellite Data. M.Sc Thesis, University of Nairobi, 1988.

Paegle, J.N. and Kierulff, L.P. (1974): Synoptic Climatology of 500mb Winter Flow Types. J. Appl. Meteor. 13, 205-212.

Peak, J. Wilson, E.W. Elsberry, R.L. and Chan, J.C.L. (1986): Forecasting Tropical Motion Using Empirical Orthogonal Function Representation of the Environmental Wind Fields. Mon. Wea. Rev., VOL. 114, 2466-2477.

Potts, A. (1971): Application of Harmonic Analysis to the Study of East African Rainfall Data. J. Tropical Geogr. 34, 31-42.

Radok, U. (1966): An Appraisal of TIROS III Radiation Data for South East Asia. Atmospheric Science Paper NO. 102, Dept. of Atmospheric Science, Colorado State University, Fort Collins, Colo., 37 pp.

Reed, W.G. and Kincer, J.B. (1917): The Prepartion of Precipitation Charts. Mon. Wea. Rev. 45, 233-235.

Richman, M.B. (1981): Obliquely Rotated Principal Components: An Improved Meteorological Map typing Technique? J. of Appl. Meteor. VOL. 20, 1145-1159.

Rinne, J. and Järvenoja, S. (1979): Truncation of the EOF Series Representing 500mb Heights. Quart. J. Roy. Meteor. Soc. VOL. 105. 885-897.

Rinne, J. and Karhila, V. (1979): Empirical Orthogonal Functions of 500mb Heights in Northern Hemisphere Determined from a Large Sample of Data. Quart. J. Roy. Meteor. Soc. VOL. 105, 873-884.

Rodda, J.C. (1969): Hydrological Network Design. Reports on WMO/IHD Projects. Report No. 12, 57 pp.

Rodda, J.C. (1972): Precipitation in A casebook on Hydrological Network Design Practice. WMO-NO.324, 1-1.2-15.

Rodriguez-Iturbe, I. and Mejia, J.M. (1974): The Design of Rainfall Networks in Time and Space. Water Resources Research VOL. 10, Aug. 1974. No.4.

Siegel, S. (1956): Non-parametric Statistics for Behavioral Sciences. McGraw-Hill, Kogakusha, Ltd., Tokyo, 1956, 312 pp.

Singh, N. (1986): Spatial Variation of Probability Distribution of Annual Precipitation in India. Mausam, 37, 4, 507-510.

Sneath, P.H.A. (1979): Some Empirical Tests for Significance of Clusters. Int. Symp. on Data Analysis and Informatics. 2d. Versailles, 1978, 491-508.

Snijders, T.A.B. (1985): Interstation Correlations and Climate. J.Appl. Met. VOL. 25, 524-531.

Stein, C. (1945): A Two Sample Test for a Linear Hypothesis Whose Power is Independent of Variance. Ann. Math. Stat. 16, 243-258.

Thiessen, A.N. (1911): Precipitation Averages for Large Areas. Mon. Wea. Rev. 39, 1082-1084.

Thom, H.C.S. (1951): Some Aspects of the Design of Meteorological Experiments. Presented at State College Pa (Meeting of American Met. Soc.) August, 1951.

Thompson, B.W. (1957): Some Reflections of Equations and Tropical Forecasting. E.A. Met. Dept. Tech. Memo. No.7.

Thompson, B.W. (1965): The Climate of Africa. Oxford University Press. 132 pp.

Tomsett, J.E. (1969): Average Monthly and Annual Rainfall of East Africa. E.A. Met. Dept. Tech. Memo. No. 14, Nairobi, Kenya.

Trewartha, G. (1961): The Earths Problem Climates. Methuen and Co. Ltd., 1961, 334 pp.

Tucker, C.J. (1986): Maximum Normalised Difference Vegetation Index Images for Sub-Saharan Africa for 1983-1985. Int. J. Remote Sensing, 1986. VOL. 7, No.11 1383-1384.

U.K. Met. Office (1975): Calculations of Areal Rainfall. U.K. Meteorological Office. Met. 08C, London, Road, Bracknell, Berkshire.

UN(ECAFE)-WMO (1960): Hydrological Networks and Methods. Flood Control Series No. 15, 1960.

Walsh, J.E. and Mostek, A. (1980): A Quantitative Analysis of Meteorological Anomaly Patterns over the United States, 1900-1977). Mon. Wea. Rev. VOL. 108, 615-629.

Whitmore, J.S., Van EEden, F.J. and Harvey, K.J. (1961): Assessment of Average Annual Rainfall over Large Catchments. Int. African Conf. Hydrol. Nariobi. C.C.T.A. Pub. 66, 100-107.

Wilm, H.G., Nelson, A.Z. and Storey, H.C. (1939): An Analysis of Precipitation Measurements on Mountain Watershed. Mon. Wea. Rev. 67, 163-172.

WMO (1962): The Use of Statistics in Climatology. Guide to Climatological Practices. WMO Suppl. NO.3, Sept. 1962.

WMO (1970): Guide to Hydrometeorological Practices. WMO No. 168 Tp 82.

WMO (1974): Guide to Hydrological Practices. Third Edition - WMO 168. Secretariat of the WMO Geneva Switzerland.

WMO (1979): Climatological Series. WMO Tech. Note. No. 79; WMO, 1979.

WMO (1981): Guide to Hydrological Practices. Data Acquisition and Processing. WMO No. 168.

WMO (1986): Guidelines on the Quality Control of Surface Climatological Data. WCP-85, WMO/TD-No. 111, 56 pp.

WMO (1989): Drought and Desertification. WCAP-7, Feb. 1989. WMO/TD-No. 286, 73 pp.

Yamane, T.C. (1973): Statistics. An Introductory Analysis. Harper and Row, Publications, Inc., New York.

Yevjevich, V. (1972): Probability Statistics in Hydrology. Water Resources Publications, Fort Collins Colorado, U.S.A., 1972. Printed by Lithocrafters, Ann Arlor, Michigan, U.S.A., 302 pp.