SECULAR VARIATION

IN

SOUTH AFRICAN RAINFALL

BY

THOMAS GEORGE JEFFREY DYER, M.Sc. (BIRM)

Thesis submitted for the degree of Doctor of Philosophy in the Faculty of Science at the University of the Witwatersrand.

Johannesburg, 1975.

ABSTRACT

Annual rainfall totals over South Africa during the period 1910 - 1972, and where availability of data permitted for the period 1880 - 1972, are examined for evidence of secular variation. It is shown that no evidence exists to support the view that the sub-continent is undergoing progressive desiccation. With the aid of a linear model station records, grouped using principal component analysis, are shown in some cases to possess trend on a micro-scale.

Strong evidence is provided to support the hypothesis that rainfall is oscillatory in nature. Oscillations in the ranges 16-20, 10-12, and 2 - 3 years are shown by different methods to be spatially dependent. A 3 - 4 year oscillation is ubiquitous.

The plausibility of using stochastic models to forecast annual rainfall totals is investigated and tentative forecasts made regarding the general pattern of rainfall behaviour over some regions of South Africa.

This is to declare that this thesis is entirely my own work and has not been previously submitted as a thesis for any degree in any other University.

19th August 1975

PREFATORY NOTE

World climatic change is a field of scientific research which is receiving ever increasing attention from not only climatologists and meteorologists but also from hydrologists, physicists, ecologists, and decision makers in general. It is therefore surprising that so little work has been done in exploring sub-continental scale regional variations, their spatial gradients and links with the general circulation of the atmosphere. In work on climatic change the emphasi has usually been on the analyses of data for discrete stations rather than on the spatial extrapolation of common effects experienced over large areas. This is reflected in the dearth of literature dealing with an approach to climatic variation that considers the whole spectrum of spatial and temporal gradients of climate. This situation is true for South Africa in particular, and for the world as a whole.

Whether South Africa is drying up has been debated for the best part of a century. Even so, definite and reliable conclusions to this effect have yet to be formulated. This is surprising when it is borne in mind that South Africa's continued economic advancement is dependent upon adequate water resources. For a meaningful evaluation of these resources, it is necessary for research to be carried out on a regional rather than an individual station scale into the essential composition of climatic variation.

In this thesis an attempt will be made to provide answers to both the problem of climatic variation in South Africa and the broader issue of how such changes can be delimited simultaneously in both the time and space domains. This having been done, some attempts will be made to use the results obtained to forecast future changes.

More specifically, the aims of this thesis are:

(i) to determine the extent to which South African rainfall has shown a systematic and progressive increase or decrease during the period of meteorological record. If neither of these have occurred,

(ii) to determine exactly what changes have taken place, In so doing,

(iii) to develop methods for the simultaneous delimitation of regional and temporal changes of climate, the spatial gradients thereof and the regional occurrence of the changes. Finally, using the results of this analysis of South African rainfall,

(iv) to attempt, in a preliminary manner only, to forecast both qualitatively and quantitatively what near-future changes in rainfall are likely to be.

In order to meet these objectives, the thesis will be devoted, on the one hand, to the analysis of what has happened, and on the other, to the discussion of this analysis with the view to the possible prediction of future events. This dichotomy is implicit in the arrangement of material.

Part I presents an introduction, and gives an historical background; pertinent aspects of the meteorology of South Africa; and the analytical techniques used in the analysis of South African rainfall data. In Part II, the results of analyses carried out on a macro, and meso scale are presented. A discussion of the results obtained in Part II is given in Part III. Part IV considers the plausibility of using stochastic models to forecast annual rainfall totals. In Part V conclusions are given and an overall summary of the results obtained in the thesis is presented.

During the course of the research for this thesis much of the material presented herein has been published. Substantive parts of Chapters 4 and 6 have been presented in a paper in the *Quarterly Journal of the Royal Meteorological Society* (1975), Vol. 101. A further paper based on Chapters 5 and 7 appears in the same volume of that Journal. Those aspects of the material in Chapters 5 and 7 not contained in the second paper, have been accepted for publication in the *Quarterly Journal of the Royal Meteorological Society* (1976), Vol. 102. A portion of the material contained in Chapter 8 has been published in the *South African Geographical Journal* (1975), Vol. 57.

* * * * *

I should like to acknowledge the assistance given by Miss J. Rouse in the final presentation of the diagrams. The typing of this thesis was carried out by Mrs. J. Youdale. My wife, Teresa, gave invaluable assistance with proof reading. To both these ladies, I extend my appreciation. Thanks must be given to Adrian and Lucy Dyer for the time they devoted to punching computer cards.

Dr. E.T. Stringer, of the Geography Department, University of Birmingham, must be acknowledged as the person who first stimulated my interest in statistical climatology. Throughout the period of this research, my mentors, Professor of Statistics, Dr. D.M. Hawkins, and Professor P.D. Tyson, Head of the Department of Geography and Environmental Sciences at the University of the Witwatersrand, have always been readily available for consultation. They have both, in their respective fields, provided advice, assistance, and stimulating discussion when required, and to them I gladly extend my especial appreciation.

The research was funded in part by a grant from the National Environmental Sciences Programme of the C.S.I.R.

* * * *

CONTENTS

PART I - INTRODUCTION

PART II - THE ANALYSIS

PART I

INTRODUCTION

CHAPTER 1

HISTORICAL BACKGROUND

1.1 RAINFALL VARIATION OVER SOUTH AFRICA: SOME CONTRADICTORY VIEWS

Variation in South Africa's rainfall receipts, and the related consequences thereof, has long been subjected to discussion and investigation. Early reports indicate the one time presence of lakes and rivers in what are now semi-arid parts of the country (Moffat, 1842); (Livingstone, 1857). A declared view that the Kalahari had become drier was made by Wilson (1865). In a general review, Barber (1910) suggested that South Africa as a whole was drying up; this view was widely held by both lay and informed members of the public (Agric. Journ. Un. S. Afr., 1913). The findings of the Union Parliament Senate Select Committee on Drought, Rainfall and Soil Erosion were contrary to popular beliefs (Union of S.A., 1914). In 1918 Schwartz wrote the first of a series of articles dealing with the desiccation of South Africa (Schwartz, 1918, 1919, 1920a, 1920b, 1923, 1925, 1926, 1927), which were all similar in content. In 1935, Wallis concluded, on the basis of an analysis of ten rainfall stations in Bechuanaland, that the Kalahari had become drier. The views and plans of Schwartz were criticised by Cox (1925), Kanthack (1930), Schumann and Thompson (1934), Thompson (1936), and Kokot (1948). Paulsmeier (1946) criticised Schwartz for not having acknowledged that the views he put forward had been given by Paulsmeier to the press some years earlier.

1.2 A CYCLICAL BEHAVIOUR PATTERN

Claims that the rainfall variation was cyclical in nature extend almost as far back in time. Hutchins (1888, 1897, 1898, 1905), and Tripp (1888, 1890) appear to have been the first to publish results leading to this view. Hutchins found evidence to support Meldrum's cycle (Meldrum, 1873) which originated from work carried out in Mauritius correlating sunspot activity with occurrence of cyclones. Furthermore, Hutchins came to the conclusion that the 12.6 year rather than the 11.11 year sunspot cycle had the greatest effect on rainfall in Maritzburg, Durban and the Cape. He also made predictions based on the now well known Bruckner cycle of 35 years (Landsburg, 1960). Tripp was of the opinion that cycles existed, but found no evidence to support the view that they were related to the sunspot cycle. Working on data at the Cape of Good Hope Observatory, Morrison (1900) found evidence which supported the existence of two cycles of length 9 and 10 years respectively. Hutchins (1907) continuing his work on cycles, found good agreement between South African rainfall, Indian rainfall, and sunspot cycles. Evidence for an 18 year periodical fluctuation in Natal's rainfall extending back to the middle of the nineteenth century was suggested by Nevill (1908). Van Reenen (1923, 1925) from knowledge gained whilst Chairman of a Drought Commission in South-West Africa decided that the coincidence between the rainfall period and the period of the Bruckner Cycle was "noteworthy". Cox (1925) found some evidence for a periodicity in rainfall at Cape Town of length 14 years from a consideration of data extending over the period 1838-1920. Rainfall data for the period 1909-1929 at Lourenco Marques was analysed by Peres (1930) who found it possible to suggest an oscillation with a period of 20 years. Levinkind (1941) divided South Africa into four climatic regions and considered the rainfall over each one separately and in combination. He was able to form the opinion that rainfall return periods varied from one area of the country to another. In 1946, an historical survey covering a period of 284 years was carried out by Tylden (1946). Although he speaks of the cyclical nature of rainfall, actual periods were not deduced.

In later years the Report of the Desert Encroachment Committee (Union Government, 1951) stated that allegations of climatic changes tended to arise because "human recollections are apt to be of a gloomy character when they relate to past climatic conditions". Glover and Robinson (1954) discussed the periodical nature of rainfall in terms of return periods, whilst Vorster (1957) decided that a general decline in rainfall could be detected between 1881 and 1950 in the southwestern Cape Province.

Van Rooy's (1966) work suggested that the country is getting wetter, but that this arises from periodical variation with increasing amplitude. Fabricius (1964) formed the opinion that the rainfall over the country as a whole was oscillatory in nature, having no fixed period. Brook and Mametse (1969) confirmed the conclusions arrived at by Hofmeyr and Schulze (1963) regarding a spatial dependence of rainfall variation, both papers employing methods based on simple linear regressions. Recent work (Tyson 1971, 1972; Keen 1971) suggests the reality of an 18-20 year oscillation which shows a distinctive regional distribution in its occurrence.

1.3 A BRIEF CRITIQUE

Past investigations into climatic changes over South Africa have used, with the exception of Tyson (1971, 1972) and Keen (1971), unsophisticated statistical techniques on the one hand, confused thinking and human recollections on the other.

Investigations into the presence of cycles are mostly based on the use of harmonic analysis. The model is fitted by least squares. This assumes that there is a random component superimposed upon a periodic movement. The elements of this random component must have zero mean and constant variance, and they must be mutually independent (Tintner, 1952). It is extremely doubtful whether this last assumption is ever fulfilled with climatological data. An even more serious defect is that the spectral estimates obtained have a standard error which is of the same order as the variance of the original data (Bendat and Piersal, 1968). Kendall (1945) illustrates very clearly that the variability of the spectral estimates does not decrease as the sample size increases. This technique should only be used on deterministic data, that is, for investigating the harmonics of a fixed identifiable frequency $2\pi/n$ under the assumption that the series is a genuine periodic function (Jenkins, 1961). In other cases, for example Nevill (1908), no indication is given as to how the claimed cycle was determined.

Simple linear regression models tend to be used for investigating the presence of trend. Apart from the general relaxation of the necessary conditions that this model requires, it is obviou (Brook and M metse, 1969; Fig. 2, p.137) that many of the regression coefficients have no physical significance as far as secular variation is concerned, even if they are statistically significant. In many cases, for example, Peres (1930), the amount of data available was too little to justify the claims made.

Confusion over physical issues can be gleaned from reading the Schwartz articles mentioned earlier. It would appear that he was inclined to confuse evaporation from the soil with that from an open water surface, and further, to treat potential evaporation as if it were the same as evaporation actually permitted by soil moisture conditions.

Tyson (1971) in smoothing the rainfall series reduced the variance of a transformed series from the original data, and hence also the standard errors of the spectral estimates. However, he was criticised by Currie, (1972) on the grounds that the apparent periodicities were spurious, having been introduced by the smoothing technique. Furthermore, Currie was concerned about this work because trend had not been removed from the raw data. Tyson later corrected for trend and was able to show that its presence had had no serious effects on his conclusions.

The foundation laid by Tyson provides a starting point for this thesis. However, because the thesis provides an introduction to future work into the much broader issues involved with climatic variation, Currie's criticisms must be assessed independently. To establish beyond all reasonable doubt that Currie's concern is without foundation the pertinent parts of Tyson's work will be re-examined independently using more sophisticated techniques that are available for time series analyses.

* * * * *

Sufficient evidence has been supplied to indicate strongly that South Africa's annual rainfall budget is unreliable. Over the course of time, diametrically opposed views have been postulated. Without exception, no one has concluded his findings by attempting to derive forecasting models.

In this thesis an awareness of the relatively short rainfall series available for analysis is reflected in the diversity of relevant investigatory methods applied to each aspect of climatic change. This will increase the reliability of conclusions made in the thesis by providing corroborative evidence and minimising the possibility that a single method may give rise to an incorrect or ambiguous result. That is, the inability to choose an obviously most powerful technique is guarded against; this is an inherent difficulty encountered in time series analysis (Kendall and Stuart, 1966).

Variation in South African rainfall, and the behaviour of such variation over time and space will be investigated. In particular, attention will be focussed on past claims that South African rainfall has undergone either a progressive increase or decrease during the period of meteorological record. If neither claim is substantiated, then the nature of any other variations that could have occurred will be diagnosed. In this event, methods will be developed that describe the simultaneous delimitation of regional and temporal changes in climate, their spatial gradients and regional occurrence.
This aspect of the research will provide the groundwork for a tentative exploration of the possibility of forecasting future climatic variations.

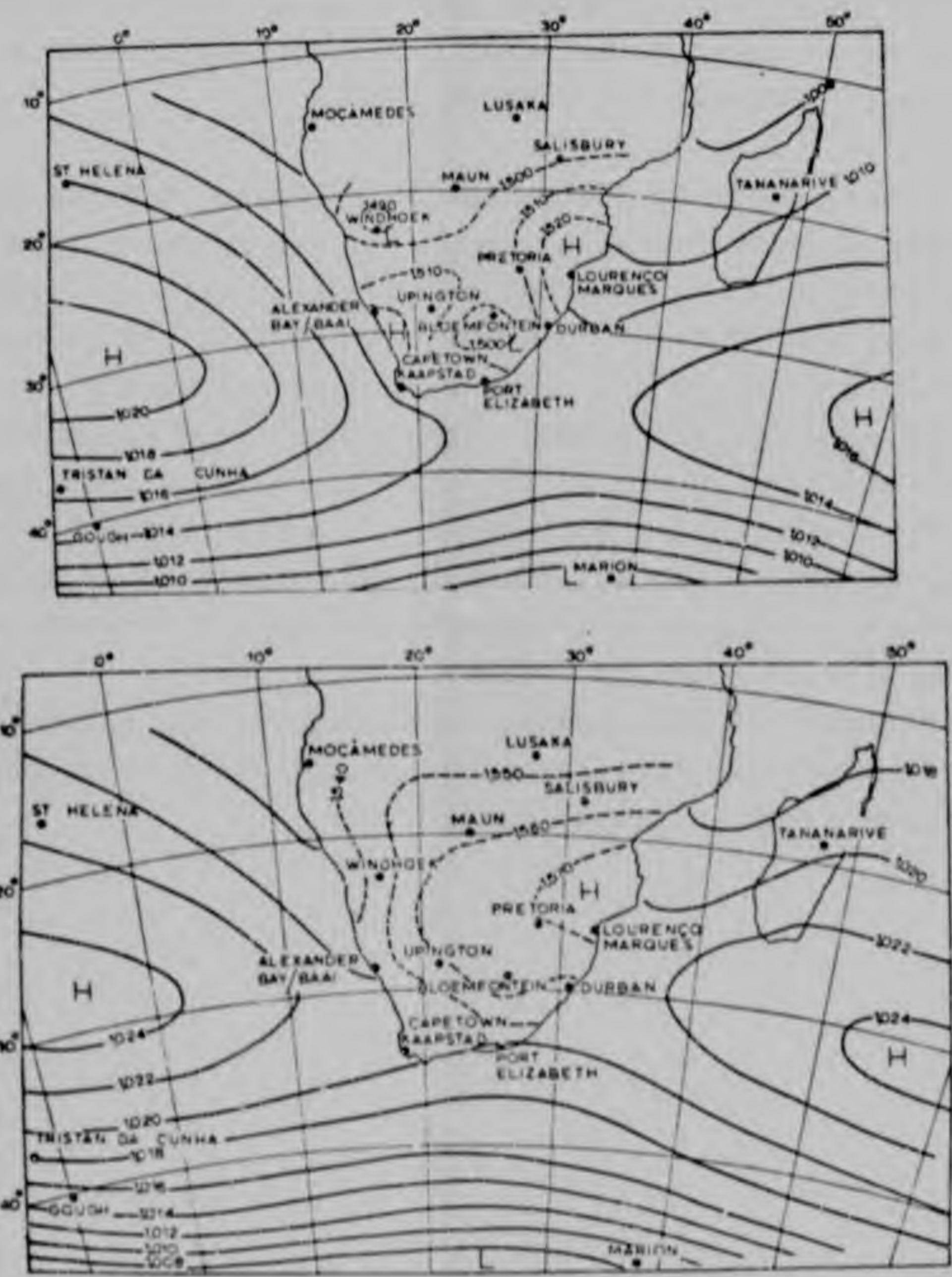


Fig. 2.1: Mean sea-level isobars (full lines) and mean contours of the 850-mbar pressure surface (broken lines) at 14hoo South African Standard Time (12hoo G.M.T.), (18 year average). Upper, January; lower, July (after Griffiths, 1972).

CHAPTER 2

SOUTH AFRICAN METEOROLOGY: AN OUTLINE

2.1 MEAN CIRCULATION PATTERNS

South Africa is situated between latitudes 22 and 35 degrees south, thus being almost completely contained in the high pressure belt of the Southern Hemisphere. This high pressure belt undergoes a seasonal latitudinal displacement of about 4 degrees (Schulze, 1972). Its centre is furthest north during July to August, and furthest south in February. Usually, a temperature inversion exists, at about 3000m., above a relatively shallow layer of maritime air in the high pressure belt; this is due to subsidence in the upper atmosphere (Taljaard and Schumann, 1940; Taljaard, 1955). This inversion is strongest at the west coat, being about 7 degrees Celsius, due to the influence of the cold Benguela Current. At the east coast the inversion is less pronounced and is situated at about 2000m. in summer, changing to 1000m. in winter.

The high pressure belt is separated into two cells. This is a direct effect of the land mass of the subcontinent intruding into a large area of water, the mass being heated differently in summer and winter. These two cells constitute the Atlantic and Indian Ocean highs, (Fig. 2.1). Much of the country's climate is therefore dominated by anticyclonic circulation patterns and attendent subsidence. This is particularly so in winter when the frequency of anticyclonic systems exceeds 70 per cent (Vowinckel, 1955).

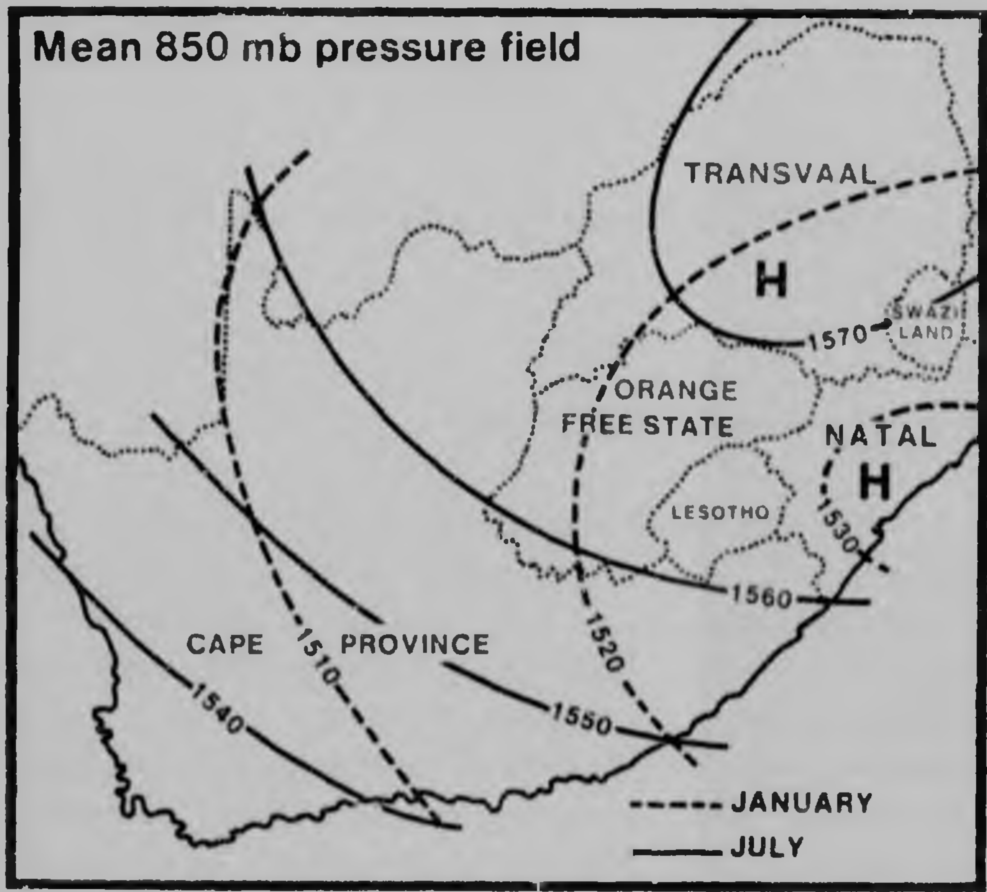


Fig. 2.2: Distribution of the mean 850-mbar pressure fields in January and July (after Jackson, 1961).

Until 1952 it was generally considered that Southern Africa was under the major influence of cyclonic conditions during summer and anticyclonic conditions during winter. At that time it was shown that the general circulation patterns over the plateau undergo no basic change through the passage of a year (Jackson, 1952). The only variation is in the strength of the continental anticyclone which weakens in summer, it being forced to move southward (Fig. 2.1).

The intensification of the high over land in winter inhibits the passage of maritime air into the interior, thus giving dry winters (Fig. 2.2). During this period when cyclonic activity is most active south of the subcontinent, cyclogenesis may occur in subtropical latitudes (Taljaard, Schmidt and van Loon, 1961; Taljaard and van Loon, 1962). These cyclonic disturbances provide the southwestern Cape with winter rainfall. The South Atlantic Ocean between latitudes 30 and 40 degrees is the general area of origin of these cyclones, (Fig. 2.3). Their centres travel east south eastwards and are followed by anticyclones travelling eastwards and often, especially in summer, moving northeastwards around the South African coast, (Fig. 2.4).

In summer, rainfall occurs over northeastern South Africa. Anticyclonic circulation transfers maritime, potentially unstable, air from the region of the Limpopo mouth and further north onto the plateau. The formation of shallow lows assists the passage of moist air into the interior, (Fig. 2.1). Rainfall is probably a result of convergence between the plateau anticyclone and the adjacent oceanic anticyclone (de Bliy, 1964). During this season little rain is received in the southwestern Cape. Only along the southern Cape coastal belt is rainfall received throughout the year.

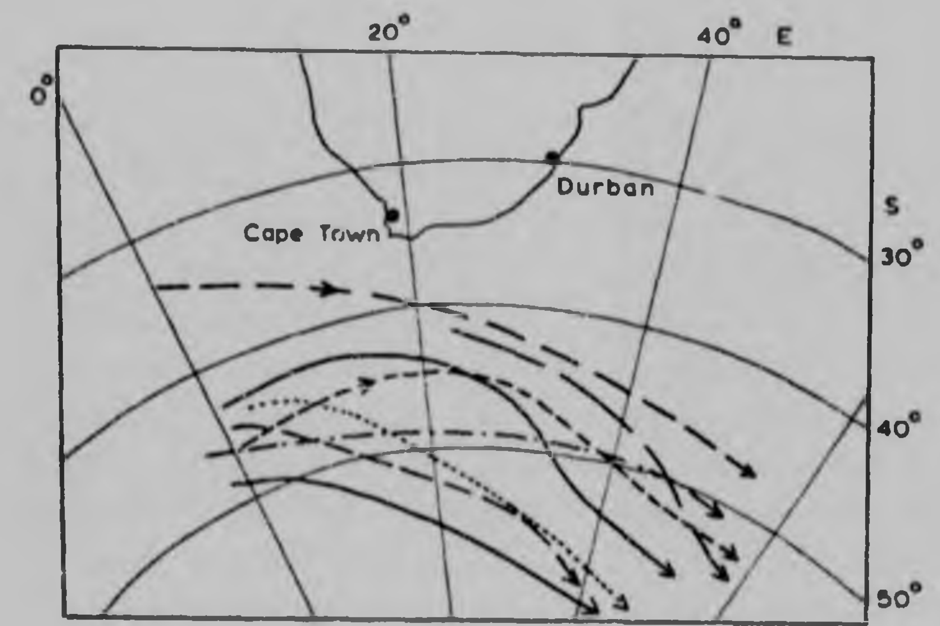


Fig. 2.3: Tracks of cyclone centres, April, 1958 (after Taljaard and van Loon, 1962, 1963).

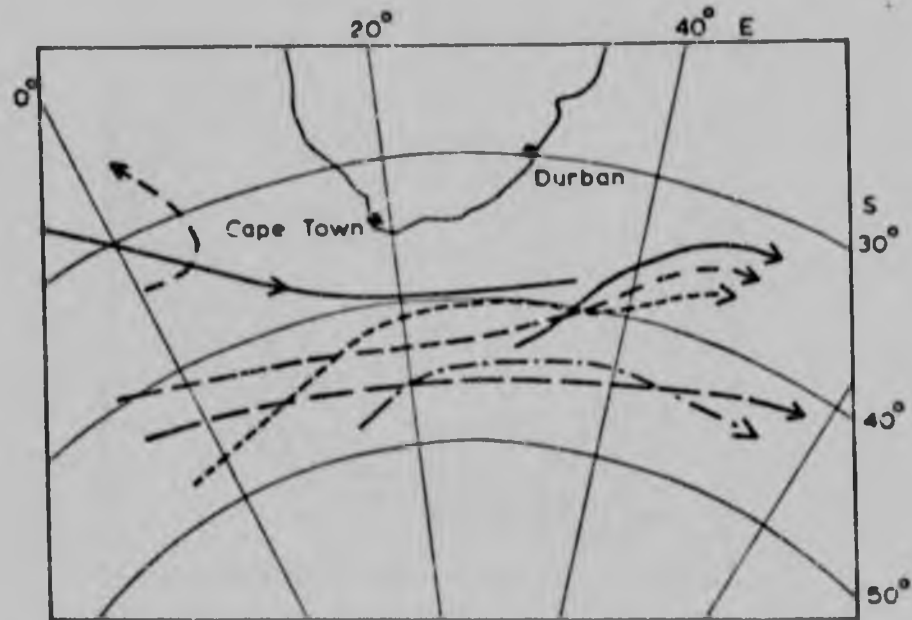


Fig. 2.4: Tracks of anti-cyclone centres, July, 1957 (after Taljaa 1, 1964, 1965).

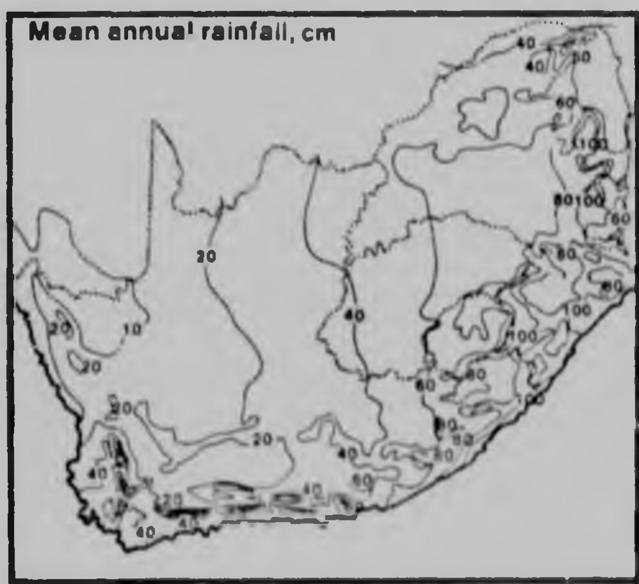


Fig. 2.5: Mean annual rainfall (after Jackson, 1961).

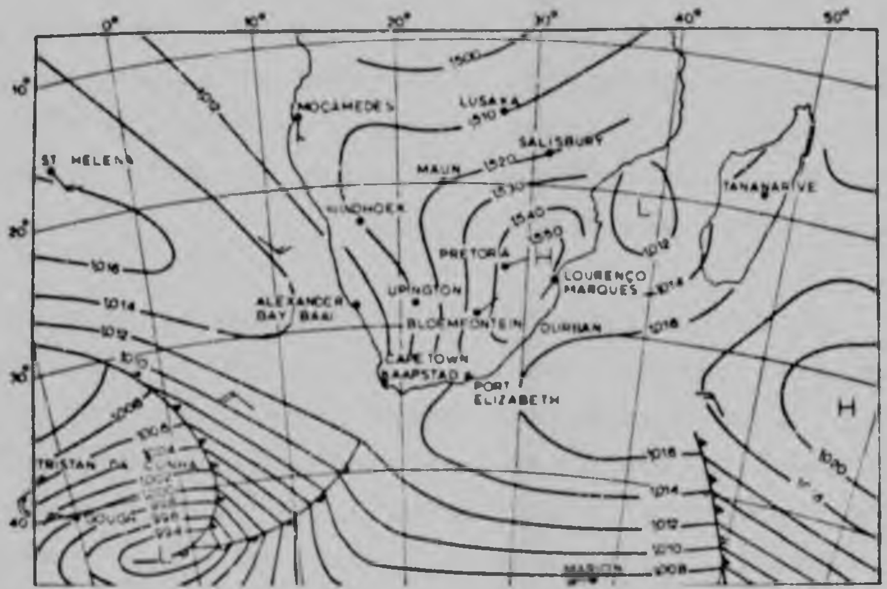


Fig. 2.6: Surface synoptic chart, January 10, 1961 - summer drought (after Griffiths, 1972).

The net result of the seasonal and daily circulation patterns is to produce a mean annual distribution of rainfall possessing a gradient from the wetter east coat to the arid west coast, (Fig. 2.5).

2.2 SYNOPTIC SITUATIONS

Weather situations over South Africa follow general patterns which are modified by numerous, but minor, variations. The frequency of occurrence of these patterns has not yet been fully investigated. Vowinckel (1955), presents the frequency distributions of some synoptic situations experienced on the high plateau.

One of the most common situations is that of widespread anticyclonic activity, which may continue for periods of up to fourteen days, (Fig. 2.6). On the surface chart the isolines on the plateau are the contours of the 700-mbar pressure level, (Fig. 2.7). Although this situation is most common in winter, it does occur sometimes in summer with accompanying disastrous droughts. This high pressure phenomenon, when present in summer, often coincides with a tropical cyclone situated in the Mozambique Channel. In winter, the southern cyclonic activity reaches furthest north. Mainly orographic rain occurs during northwesterly winds against the western, southwestern and southern escarpment as the cold front approaches land. This condition is taken as normal in winter and explains the Mediterranean climate of the southwestern Cape. The presence of a minor coastal low on the west coast, which moves around the south coast, is quite common. Under these conditions hot Berg winds blow which are approximately perpendicular to the coastline. They can persist for some days, terminating when the cold front reaches the land. Inland, on the plateau, clear skies persist. In general, this situation is described closely by Vowinckel's (1955) type 1. He gives the frequency of occurrence of this type as being between 60-70

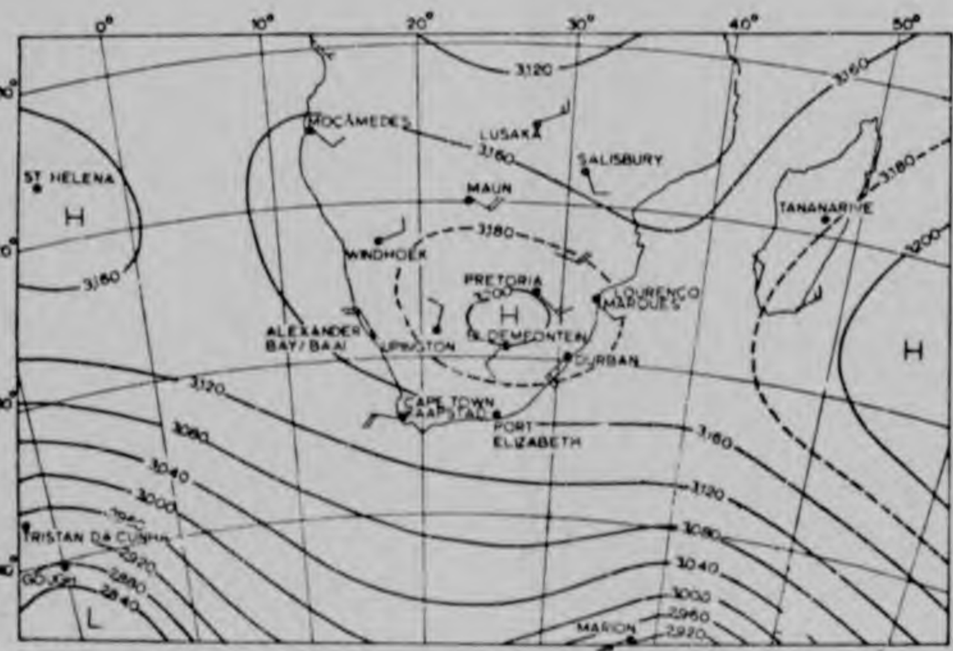


Fig. 2.7: Contours of the 700-mbar level (m), January 10, 1961 (after Griffiths, 1972).

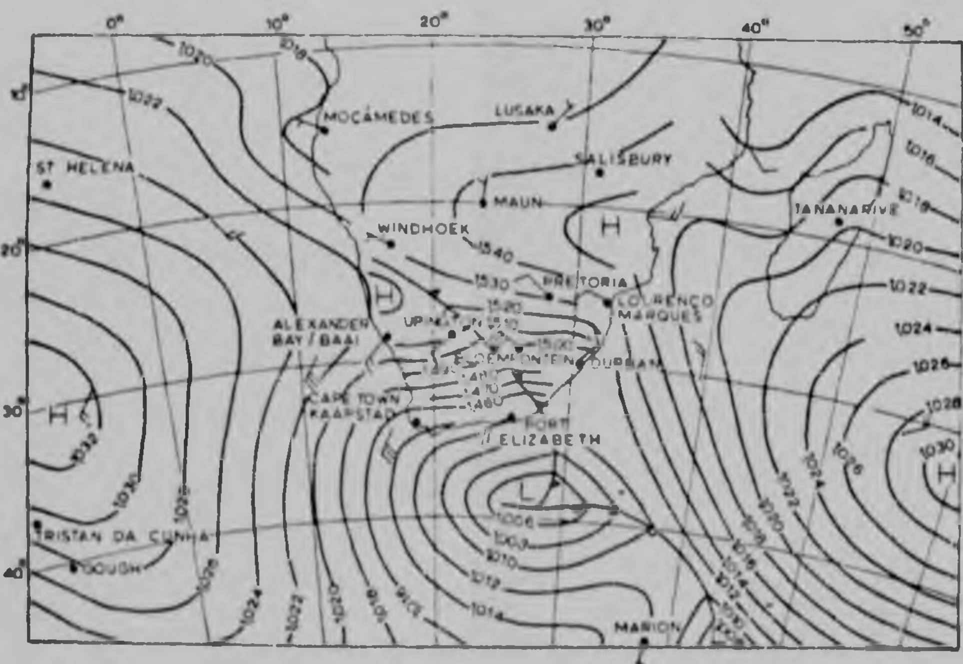


Fig. 2.8: Surface synoptic chart, August 11, 1964 - cold front (after Griffiths, 1972).

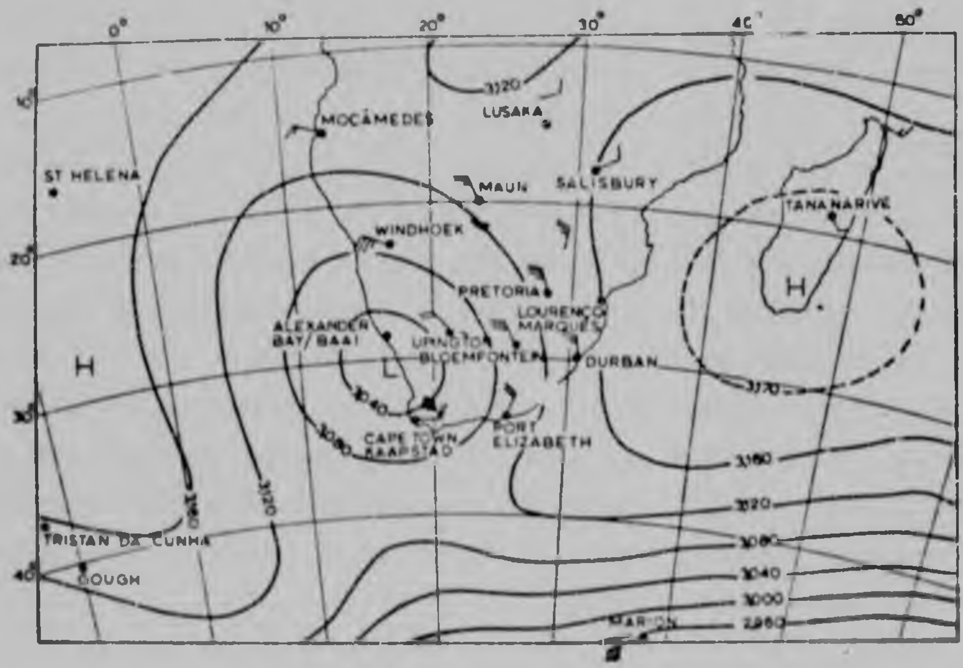


Fig. 2.9: Contours of 700-mbar level (m), August 11, 1964 (after Griffiths, 1972).

per cent in mid-winter and about 8 per cent in mid-summer.

Often in late winter and spring, short cold spells occur by virtue of an influx of cold air over the whole country, (Figs. 2.8, 2.9). Generally bad weather, with gales along the southwestern, southern, and southeastern coasts can occur when this synoptic condition prevails. The severity of the weather is dictated by the intensity of the low pressure system and the proximity of its centre to the coast. Passage of the cold front across the coast brings about associated pre-frontal rain over the southern Cape Province, whilst post-frontal showers tend to clear quickly from the west. This cold air mass can result in snow on the high mountain regions in the south and southeast (Lesotho). The high pressure system, which follows in the wake of the low one, moves around the coast causing orographic cloud and drizzle on the southeastern and eastern escarpments. During this process the interior experiences clear skies.

2.3 SEASONALITY

Most of South Africa experiences seasonal rainfall (S.A.W.B., 1972). Of the area that is so affected, the largest portion receives its rain during the summer season which is taken to be from October to March. A decidedly smaller area in the western and southwestern Cape Province has winter rainfall. All season rainfall, but having a bi-modal distribution, is experienced by southern coastal areas, (Fig. 2.10). Rainfall over the country experiences a marked change from being seasonally winter in the south to seasonally summer in the north, and this is clearly indicated by the isolines of summer seasonal rainfall expressed as a percentage of the normal annual rainfall totals, (Fig. 2.11). An interior transitional zone exists between these two parts of the country. In this zone a six-month wave (equinoctial maximum) is experienced.

Fig. 2.10: Histograms to show spatial variation of mean monthly rainfall over South Africa (after Keen, 1971).

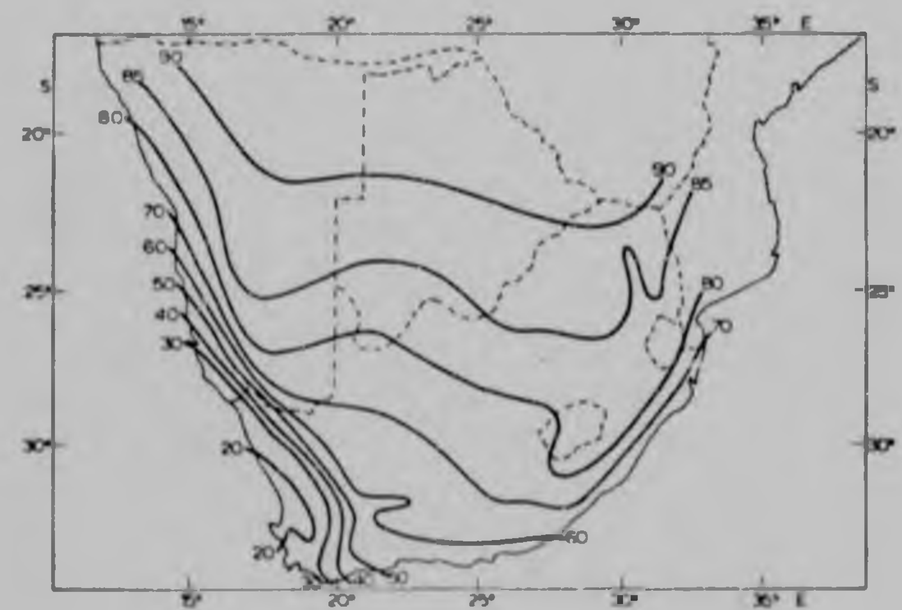


Fig. 2.11: Seasonal rainfall (October-March) as a percentage of the normal annual rainfall (after S.A.W.B., 1965).

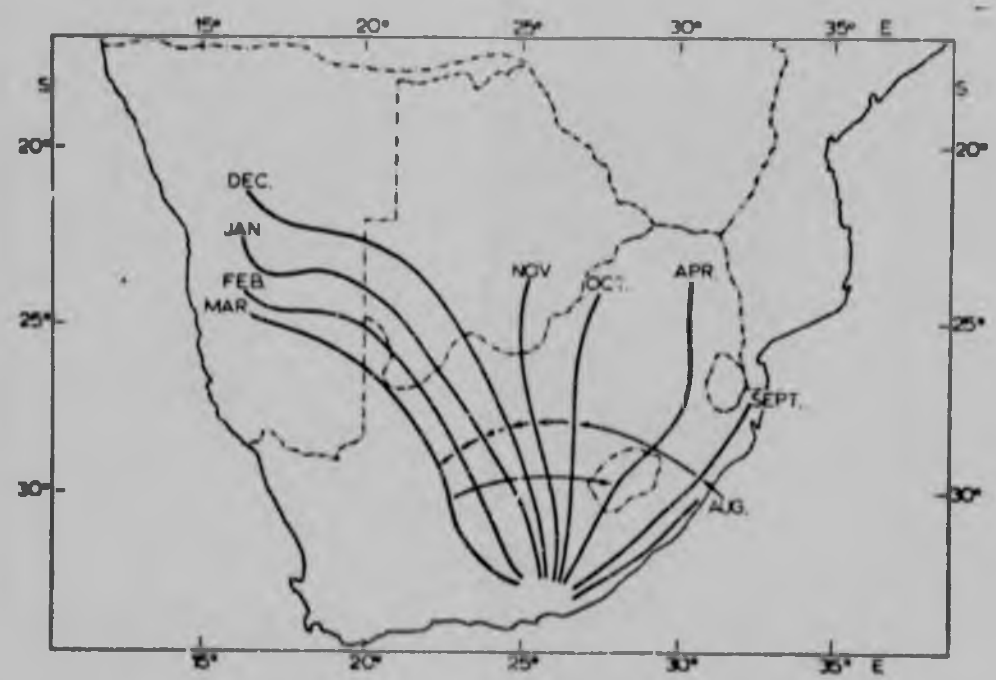


Fig. 2.12: Temporal and spatial variation of the 50-mm isohyet over the summer rainfall region (after S.A.W.B., 1965).

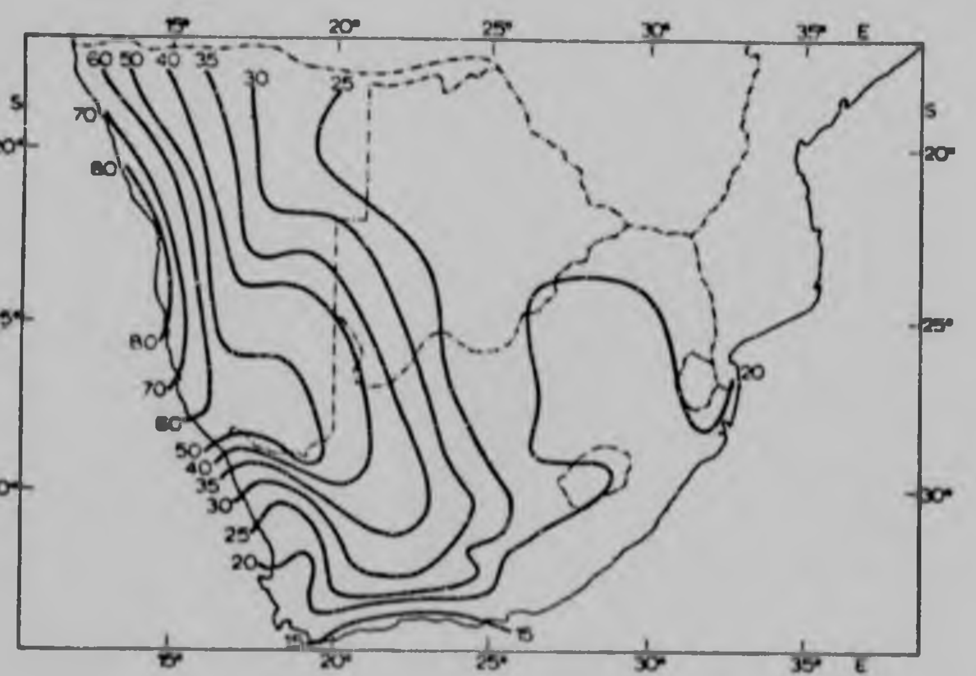


Fig. 2.13: The relative variability of annual rainfall in percent (after S.A.W.B., 1965).

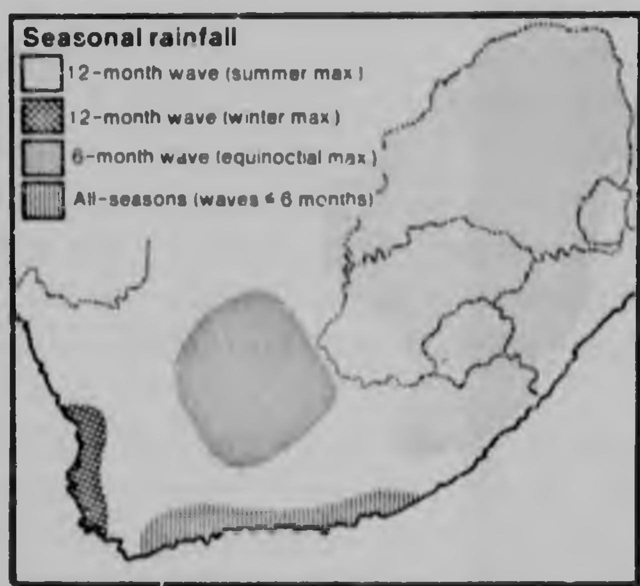


Fig. 2.14: Seasonal rainfall regimes (after Keen and Tyson, 1973).

A different illustration of seasonality is obtained from a consideration of the temporal and spatial oscillatory behaviour of individual isohyets. The 50-mm. one, viewed over the summer rainfall region of the country, is at the southeastern coast in winter. During September it commences a westward passage across the country and attains its maximum distance from the coast in March, (Fig. 2.12). Its return passage eastward is more sudden, reaching the eastern escarpment in April. This isohyet has a definite axis of rotation which is centred to the north of Port Elizabeth, (Fig. 2.12). The behaviour of this isohyet provides a marked feature of the summer rainfall regime over South Africa (Taljaard, 1955).

The relative variability, defined here as mean absolute deviation, divided by the mean rainfall expressed as a percentage, shows marked change over the different seasonal rainfall areas of the country. It is least over the south and southwestern coastal regions and over much of the summer rainfall area. Over the western interior its gradient is severe, reaching a maximum on the arid coast of South-West Africa, (Fig. 2.13).

* * * * *

Southern Africa's rainfall is closely associated with anticyclonic activity, resulting from the subcontinent's virtual complete containment in the southern hemisphere's high pressure belt. This rainfall is highly seasonal, progressing from winter and all seasons in the south to summer over a larger portion of the northern part of the subcontinent. The interior transitional zone is clearly defined, (Fig. 2.14).

CHAPTER 3

ANALYTICAL TECHNIQUES

3.1 INTRODUCTION

The statistical techniques used in the subsequent analyses of rainfall data are described. Theoretical aspects of the analyses have not been dealt with in detail unless new theory is developed; standard techniques have various interpretations; the results would be ambiguous without elucidation of the techniques used; or new applications of a standard technique are used or developed.

3.2 TREND INVESTIGATION

A detailed analysis of a time series requires knowledge of its degree of stationarity. In general a series should be at least weakly stationary (Anderson, 1971). A commonly accepted method for deciding on the presence or absence of trend is to inspect the graph of the raw data (Chatfield and Pepper, 1971). This technique has the weakness of permitting only subjective conclusions. Of the many objective methods, the most common is that of considering the serial correlation for a time lag of unity. An inherent weakness with this test is that the underlying assumptions are seldom satisfied by climatological data. Statistical conclusions can therefore be unreliable. Furthermore, the first serial correlation is insensitive to non-linear trend (Kendall and Stuart, 1966). When dealing with rainfall data, apart from the type of trend of interest, a non-parametric test statistic should be used (Dyer, 1974). Kendall's *Tau* has been used in this thesis (Kendall, 1962). It is non-parametric; sensitive to both linear and non-linear trend; and has a power efficiency of 91 per cent where the basic assumptions necessary for a parametric test are satisfied.

Tau is defined as

$$\tau = \frac{4P}{n(n-1)} - 1 \qquad (3.1)$$

where P is the total number of pairs for which $u(j)>u(i)$ for $j > i$ and $1 \leq i \leq (n-1)$, in a time series defined as the sequence of

$$u(t) \text{ where } 1 \leq t \leq n \qquad (3.2)$$

n being the number of terms in the series.

In the null case

$$E(\tau) = 0 \qquad (3.3)$$

with $$Var(\tau) = \frac{2(2n+5)}{9n(n-1)} \qquad (3.4)$$

The distribution of *Tau* tends rapidly to normality for $n > 10$ (Kendall and Stuart, 1961). An alternative hypothesis of trend (two-sided) is taken against a null one of no trend. A five per cent level of significance has been used for the rejection of the null hypothesis. Where significant trend has been removed, it was done by fitting least squares polynomials (Yule and Kendall, 1958).

3.3 FILTERING

Where filtering of the rainfall data $u(t)$ for $1 \leq t \leq n$ has been applied, it has been done by employing a 5-term binomial low-pass filter (Mitchell et al, 1966). The elements of the smoothed series are given by

$$S(t) = 0.06\ u(t-2) + 0.25\ u(t-1) + 0.38\ u(t) + 0.25\ u(t+1) + 0.06\ u(t+2) \quad (3.5)$$

for $3 \leq t \leq (n-2)$

The first and last two terms have been estimated using

$$S(1) = 0.54\ n(1) + 0.46\ u(2) \quad (3.6)$$

$$S(2) = 0.25\ u(1) + 0.50\ u(2) + 0.25\ u(3) \quad (3.7)$$

$$S(n-1) = 0.25\ u(n-2) + 0.50\ u(n-1) + 0.25\ u(n) \quad (3.8)$$

$$S(n) = 0.54\ u(n) + 0.46\ u(n-1) \quad (3.9)$$

Equations 3.6 to 3.9 inclusive are an adaption of the Hanning smoothing technique (Bendat and Piersol, 1968).

The frequency response function for the smoothed series $S(t)$ for $3 \leq t \leq (n-2)$ is given by

$$R\ (f) = \cos^{p} \Pi\ f\ \Delta t \quad (3.10)$$

where f is the relevant frequency; $p = 5$ the order of the binomial expansion used; and $\Delta\ t = 1$ the sampling interval

(Stringer, 1972).

This response function gives ratios of smoothed to unsmoothed amplitudes of 0.952, 0.925 and 0.813 for periods of 20, 16 and 10 years. Therefore, unlike ordinary simple moving averages, the tendency to convert peaks into troughs does not exist. Furthermore, binomial weighting coefficients produce no phase shift in the data (Stringer, 1972).

The order of smoothing has been kept small purposely. This reduces the tendency of smoothing to generate spurious oscillations because of the Slutzky-Yule effect described by Moran (1949); (Slutzky, 1937; Yule, 1926).

Taking moving averages of order k, over a random series, introduces serial correlations in the smoothed series (Kendall and Stuart, 1966). The s-*th* serial correlation, $r(s)$, for the smoothed series is given by

$$r(s) = \frac{\sum_{j=1}^{k-s} a_j a_{j+s}}{\sum_{1}^{k} a_j^2}, \quad |s| < k \qquad (3.11)$$

$$= 0 \qquad \text{elsewhere}$$

where the a_j are the weighting factors. For a simple moving average, the weights are equal and equation 3.11 reduces to

$$r(s) = 1 - \frac{|s|}{k}, \quad |s| < k \qquad (3.12)$$

$$= 0 \qquad \text{elsewhere}$$

These equations indicate that $r(s)$ resulting from binomial weights is less than the corresponding value when simple moving averages are used (Table 3.1). Hence the tendency to generate spurious oscillations is reduced for the binomial system.

TABLE 3.1 - SERIAL CORRELATIONS INTRODUCED BY SMOOTHING DATA

s	Weighting system	
	simple	binomial
1	0.80	0.79
2	0.60	0.39
3	0.40	0.11
4	0.20	0.01

3.4 SERIAL CORRELATIONS

Serial correlations have been calculated from

$$r(k) = \frac{\sum_{t=1}^{n-k} u(t)\ u(t+k)}{\sum_{t=1}^{n} \{u(t)\}^2}, \quad -(n-1) \leq k \leq (n-1) \qquad (3.13)$$

which has been chosen in preference to

$$r'(k) = \frac{\frac{1}{n-k}\sum_{t=1}^{n-k} u(t)\, u(t+k)}{\frac{1}{n}\sum_{t=1}^{n} \{u(t)\}^2}, \quad -(n-1) \leq k \leq (n-1) \qquad (3.14)$$

because although it provides biased estimates, the variances are less than those obtained from 3.14 (Otnes et al, 1972). In each of 3.13 and 3.14 $u(t)$ and $u(t + k)$, for $o \leq k \leq (n-1)$, are terms from a time series having zero expectation (if necessary the series must be coded to accomplish this expectation). Because of the symmetrical properties of $r(k)$, its values need only be considered in the range $o \leq k \leq (n-1)$.

3.5 FOURIER ANALYSIS

The general Fourier equation can be written as

$$Y_i = \sum_{h=1}^{\infty} (A_h C_h + B_h S_h) \qquad (3.15)$$

where

$$C_h = \cos \frac{2h\pi X_i}{\lambda} \quad \text{and} \quad S_h = \sin \frac{2h\pi X_i}{\lambda}$$

The subscript h is the harmonic number, λ the fundamental wavelength or first harmonic, Y_i the observations taken at equally spaced intervals X_i. For a finite series containing n equally spaced observations

$$A_h = \frac{2}{n} \sum_{t=1}^{n} Y_i \cos \frac{2h\pi X_i}{\lambda} \qquad (3.16)$$

$$B_h = \frac{2}{n} \sum_{t=1}^{n} Y_i \sin \frac{2h\pi X_i}{\lambda} \qquad (3.17)$$

The coefficient B_o is zero, and $A_o = \Sigma Y_i/n$. The amplitude of the h-th harmonic is

$$am_h = \frac{(A_h^2 + B_h^2)^{\frac{1}{2}}}{2} \qquad (3.18)$$

and its phase angle is

$$\phi_h = \tan^{-1} \left[\frac{-B_h}{A_h} \right] \qquad (3.19)$$

The power or variance association with the h-th harmonic is

$$s_h^2 = A_h^2 + B_h^2 \qquad (3.20$$

and if s_y^2 is the total variance of Y in the series, the percentage of the contribution of the h-th harmonic is simply

$$\frac{s_h^2}{2s_y^2} \cdot 100\% = \text{percentage of contribution} \qquad (3.21)$$

3.6 UNIVARIATE SPECTRAL ANALYSIS

Sample smoothed spectral estimates have been obtained from

$$\bar{R}(f) = 2\Delta\{1+2\sum_{k=1}^{m-1} \lambda(k)r(k)\cos 2\pi fk\Delta\}, \quad o\leq f\leq \frac{1}{2\Delta} \qquad (3.22)$$

in preference to

$$\bar{C}(f) = 2\Delta\{C(o)+2\sum_{k=1}^{m-1} \lambda(k)c(k)\cos 2\pi fk\Delta\}, \quad o\leq f\leq \frac{1}{2\Delta} \qquad (3.23)$$

since the former provides standardised estimates. In each of 3.22 and 3.23 Δ is the sample rate; $\lambda(k)$ the smoothing function or lag window; f the frequency at which spectral estimates are obtained; m the maximum value of the time lags, k. In 3.22 $\bar{R}(f)$ are the sample smoothed spectral density estimates; $r(k)$ are the serial correlations for $1 \leq k \leq (m-1)$. In 3.23 $\bar{C}(f)$ is the sample smoothed spectral power estimate; $c(o)$ the sample variance; and $c(k)$ the serial covariances for $1 \leq k \leq (m-1)$. Single sided spectra have been obtained in this thesis, permitted by the symmetrical properties of the spectral estimator; this accounts for the factor two appearing in both 3.22 and 3.23.

A confusingly large number of lag windows $\lambda(k)$ are available. Differences in their mathematical properties are minimal (Hannan, 1967). Some windows give less bias and better resolution than others but this is usually associated with increased variance of the spectral estimates, and vice versa. The choice of window appears to be almost arbitrary (Jenkins and Watts, 1968). The window used most commonly is that due to Blackman and Tukey, (1958), but for no better reason than that it became known early in the development of spectral analysis and is thus used in many statistical packages. A window not dealt with by Hannan is Parzen's (1957, 1958, 1961). This one never produces negative ordinates. The correlation between non-adjacent estimates is less than Tukey's as is the variance of the estimates. In this thesis Parzen's window has been used.

This window is defined as

$$\lambda(k) = 1 - 6\left[\frac{k}{m}\right]^2 + 6\left[\frac{|k|}{m}\right]^3, \quad |k| \leq \frac{m}{2} \qquad (3.24)$$

$$= 2\left[1 - \frac{|k|}{m}\right]^3, \quad \frac{m}{2} < |k| \leq m$$

$$= 0, \qquad \text{elsewhere}$$

It follows that the unsmoothed spectrum is a special case of the smoothed one. A periodogram can be obtained for the case when $\lambda(k)$ is equal to unity for all k, $o \leq k \leq (m-1)$, and m is put equal to n, the number of terms in the series.

No standard procedure has yet been developed for obtaining m, the cut-off or maximum lag, k. Granger and Hatanaka (1966) suggest $m < n/3$ for small n and $n/5<m<n/6$

for large n. The value of n used in this thesis is 63, which is considered to be small. However, reliable spectral estimates can be obtained from this sized sample (Granger and Hughes, 1968). Small m gives large bias with correspondingly smaller variance of the spectral estimates. As m is increased the bias is decreased, but the variance increases. A technique known as window closing has been used in this thesis as a means of choosing the optimum value of m (Jenkins and Watts, 1968). Three values of m are chosen equal to, say, $n/3$, $n/4$ and $n/5$. A spectrum is obtained for each m, and the m that best balances bias against variance, judged by the commencement of instability in the spectrum, is taken as the optimum value. The method is by no means foolproof. A more meaningful spectrum can sometimes be obtained by using different m values for different sections of a spectrum. When more than one m is used it often entails having one value for the high frequency end, and another for the lower frequency end of the spectrum.

The behaviour of the correlograms should also be noted, and used as an additional aid in the determination of m. This entails noting the value of k beyond which all further $r(k)$ are zero, within sampling limits. Also, the correlogram is oscillatory if the time series is periodic or nearly so; this provides a further reason for using serial correlations rather than covariances as an intermediate step in obtaining the spectral estimates. In the present work, m took on values 10, 20 and 30 years; m equal to 20 was found satisfactory.

Commonly, $(m + 1)$ values of $\bar{R}(f)$ are obtained by letting f take the values 0, $1/m$, $\frac{1}{2}m$,....$\frac{1}{2}$. Jenkins and Watts (1968) are inclined to the view that this spacing, $1/m$ is too large. It is possible that peaks in the spectrum will be missed. However, too small an interval can give rise to correlation problems between adjacent estimates (Bendat and Piersol, 1968). In this thesis, values of $\bar{R}(f)$ have been obtained for a frequency spacing of $1/2F$ where $F = 2m$. The final computation equation for $\bar{R}(f)$ with $\Delta = 1$ year is given by

$$\bar{R}(f) = 2\left[1 + 2\sum_{k=1}^{m/2} r(k)\left(1 - 6\left(\frac{k}{m}\right)^2\left\{1 - \frac{k}{m}\right\}\right)\cos\frac{\pi l k}{F}\right.$$

$$\left. + 4\sum_{k=\frac{m}{2}+1}^{m-1} r(k)\left(1 - \frac{k}{m}\right)^3 \cos\frac{\pi l k}{F}\right] \qquad (3.25)$$

$for\ l = 0, 1, \ldots\ldots, F.$

The significance of spectral ordinates are often tested against a null hypothesis of a first order Markov generating process for the time series (Mitchell et al, 1966). Such a model has been used by Erat (1973a,b), Jaganathan and Parthasarathy (1973), and Parthasarathy and Dhar (1974). The rainfall data for South Africa did not satisfy the conditions for this model. Within sampling limits $r(2) = r^2(1)$, $r(3) = r^3(1)$, etc., when the Markov model is the underlying process. A number of other null hypot ses may be postulated and identified by methods presented by Box and Jenkins (1970), Nelson (1974). In the work being described the null hypothesis is that the generating process of the time series is random. Statistical significance of peaks in the rainfall spectra have been tested using standard methods (Jenkins and Watts. 1968; Granger and Hatanaka, 1966).

Spectral analyses have been performed on raw data. This policy is justified because no series had severe trend as judged by inspecting graphs of the data that have significant trend statistics. This follows the policy advocated by Tukey (1961), and removes the necessity of pre-whitening the data and the subsequent post-darkening. Pre-whitening can have adverse effects on the final spectral estimates (Granger and Hughes, 1971). Removal of trend by taking first differences or using least-squares straight lines is often done as a matter of course. However, the resulting phase shift is just as commonly ignored. This phase movement can be quite significant at the low end of the spectra, the very section of the spectra where clarification is often sought. A further guide for the presence of severe trend is excessive power at the low frequency end of the spectrum (Chatfield and Pepper, 1971). The correlograms of trend-possessing data tend to remain positive, or negative up to quite large k values. It must be borne in mind that it is often difficult to distinguish between trend and an oscillation of long period.

3.7 MULTIVARIATE SPECTRAL ANALYSIS

The sample smoothed coherency spectrum between two stochastic processes can be defined as

$$\overline{Coh}\,(f) = \frac{\overline{co}^{2}(f) + \bar{q}^{2}(f)}{\bar{R}u_{1(t)}(f) + \bar{R}u_{2(t)}(f)} \qquad (3.26)$$

and the smoothed phase spectrum as

$$\overline{ph}(f) = tan^{-1}(\bar{q}(f)/\overline{co}(f)) \qquad (3.27)$$

where $\overline{co}(f)$ is the smoothed co-spectrum defined as

$$\overline{co}(f) = 2\left[A(o) + 2 \sum_{k=1}^{m} \lambda(k).A(k)\cos\frac{\pi k l}{F} \right] \tag{3.28}$$

for $o \leq l \leq F$, $F = 2m$

$\overline{q}(f)$ is the smoothed quadrature spectrum defined as

$$\overline{q}(f) = 4 \sum_{k=1}^{m-1} \lambda(k).\ B(k)\sin\frac{\pi k l}{F} \tag{3.29}$$

for $o \leq l \leq F$, $F = 2m$

$$A(k) = \frac{1}{(n-k)} \sum_{j=1}^{n-k} \{u_1(j).u_2(j-k) + u_2(j).u_1(j-k)\} \tag{3.30}$$

$$B(k) = \frac{1}{(n-k)} \sum_{j=1}^{n-k} \{u_1(j).\ u_2(j-k) - u_2(j).u_1(j-k)\} \tag{3.31}$$

$\overline{R}_{u_1(t)}(f)$, $\overline{R}_{u_2(t)}(f)$ are the smoothed spectral estimates for any two time series $u_1(t)$ and $u_2(t)$ respectively. Parzen's lag window $\lambda(k)$ is used in the present study. Estimation problems are similar to the univariate case; a maximum lag of $m = 20$ years has been used. Other forms of the expressions given in 3.26 to 3.31 inclusive are given in Granger and Hatanaka (1966), and Gudmendsson (1971).

3.8 PRINCIPAL COMPONENT ANALYSIS

Over recent years principal component analysis has proved a useful technique in meteorology and climatology

(Barry and Perry, 1973; Colgate, 1975; Craddock, 1965; Craddock and Flintoff, 1970; Craddock and Flood, 1969; Grimmer, 1963; Kutzbach, 1967; Stidd, 1967; Trenberth, 1975). It often enables the data matrix to be represented by a smaller set of new orthogonal variables which account for most of the variance in the original matrix. The analysis has the advantage, unlike many multivariate techniques, of not relying on the distributional properties of the sampled population (Gower, 1966). In this thesis the application of the technique has been broadened.

Eigenvalues have been obtained for the correlation matrix of variables (rainfall stations); this reduces all variables to equal importance as measured by scale (Cattell, 1965; Kendall and Stuart, 1966). The cut-off value for eigenvalues contributing a significant reduction in the total variance has been taken as unity (Jeffers, 1967; Jolliffe, 1972; Kaiser, 1961).

Rotation of the eigenvectors has been performed using the varimax method (Child, 1970; Kaiser, 1958). For this operation the I.B.M. Scientific Subroutine VARMX has been used. The advantages of rotation become evident in the discussion of results which follows later in the thesis. The modes of analysis available are defined in Table 3.2. In this thesis S mode is used; rainfall stations are the variables (entities) and years are cases (occasions), (Rummel, 1970).

Examples of entities are rainfall recording stations, countries; of characteristics, rainfall, temperature; and of occasions, usually recordings over space or time.

TABLE 3.2 - DISTINGUISHING FEATURES OF THE MODES USED IN FACTOR ANALYSIS

Mode	Variable (column)	Cases (rows)
R	characteristics	entities
Q	entities	characteristics
O	occasions	characteristics
P	characteristics	occasions
T	occasions	entities
S	entities	occasions

3.9 THE COMPONENTS OF ANNUAL RAINFALL TOTALS

A linear fixed two-way unreplicated model has been fitted to sets of time series obtained from the principal component analysis, making it possible to investigate, individually, the behaviour of the four components of which annual rainfall totals are assumed to consist. The model may be defined as

$$X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, \quad i = 1,\ldots,n; \quad j = 1,\ldots,s \tag{3.32}$$

where X_{ij} is the annual rainfall total for the i-*th* year and the j-*th* station; μ is the overall mean annual rainfall for a region containing a set of s stations with records of annual totals extending over n years; α_i is the i-*th* level of the *regional temporal effect (A)*; β_j is the j-*th* level of the *regional spatial effect (B)*; and ε_{ij} is the $N(o,\sigma^2)$ error variable. The interaction $(\alpha_i\beta_j)$ between the main effects A and B is assumed to be zero (Afifi and Azen, 1972).

To obtain unique least-squares estimators for the parameters, the side conditions

$$\sum_{i=1}^{n} \alpha_i = 0 \, , \qquad \sum_{j=1}^{s} \beta_j = 0 \tag{3.33}$$

are imposed, and thus

$$\hat{\mu} = \frac{1}{ns} \sum_{i=1}^{n} \sum_{j=1}^{s} x_{ij} \tag{3.34}$$

$$\hat{\alpha} = \frac{1}{s} \sum_{j=1}^{s} x_{ij} - \hat{\mu} \tag{3.35}$$

$$\hat{\beta}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij} - \hat{\mu} \tag{3.36}$$

and

$$\hat{\varepsilon}_{ij} = x_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j \tag{3.37}$$

• • • • •

The techniques outlined above provide a versatile approach to the analysis of South African rainfall when 157 recording station records of annual totals are considered as a set of time series. The results obtained from these analyses are next presented, to be followed in later chapters by a discussion of their significance.

PART II

THE ANALYSIS

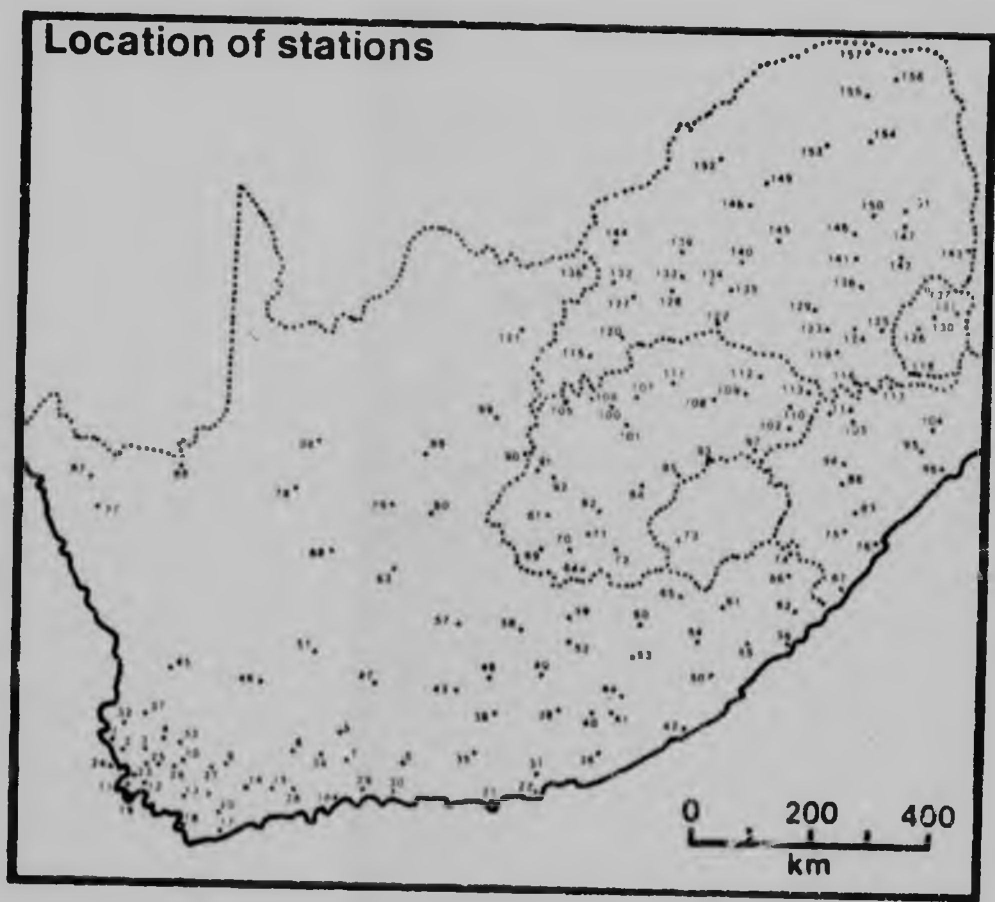


Fig. 4.1: Location of rainfall recording stations used in the present study.

CHAPTER 4

A MACRO-SCALE ANALYSIS OF SOUTH AFRICAN RAINFALL

4.1 INTRODUCTION

The rainfall station records analysed in the following chapters have been screened both for spatial representation and continuity in previous work, (Tyson, 1970). The data matrix of annual rainfall totals is large, having order (157 x 63). If the results from this thesis show spatially dependent oscillatory patterns in rainfall, then future research using smaller sampling rates will be justified.

In this chapter, results of a trend analysis of 157 rainfall series will be considered for the period 1910-1972, (Fig. 4.1). A subset of these series, with records extending back to 1880 has also been analysed for trend behaviour. These series, having been detrended or adjusted to zero means, were submitted to a 5-term binomial filtering. New series were derived from the smoothed series by taking 5 year and 9 year cumulative totals. Maxima and minima of these totals are mapped to show that South Africa's rainfall possesses regional spatial and temporal variation. For purposes of comparison, an identical analysis is performed on randomly generated simulated series.

Univariate spectral analysis is carried out on the unsmoothed series. Isolines of spectral density are used

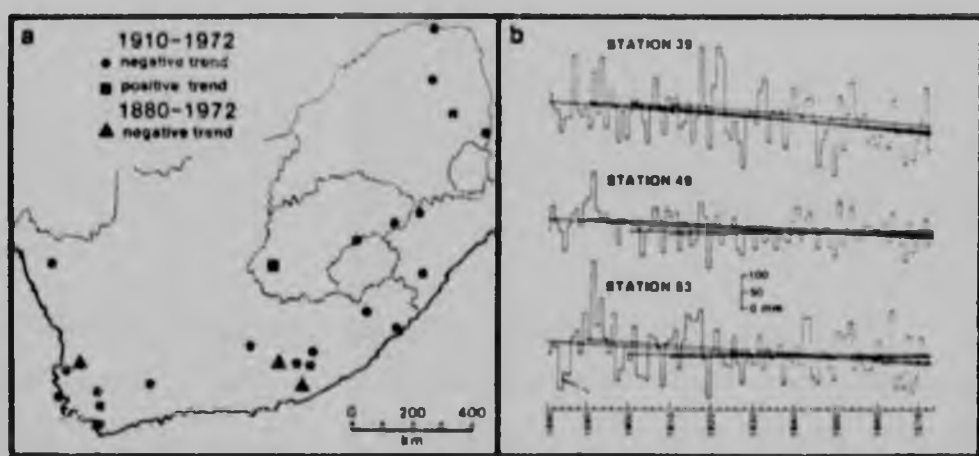


Fig. 4.2: (a) Distribution of stations showing significant trend (5 per cent, two tail) during the periods 1910-1972.

(b) Some examples showing the time dependence of trend in annual rainfal totals.

to investigate further the regional behaviour of rainfall patterns over the country.

4.2 THE TREND ANALYSIS

Of the 157 rainfall series tested for trend, 26 have significant Tau values. The distribution of stations with series showing significant trend is given in Fig. 4.2a. Twenty four records show significant negative Tau values, the other two being positive, (Table 4.1). The rejection, or critical, values of Tau are -0.1692 and 0.1692 at the five per cent level (two-tailed).

TABLE 4.1 - SIGNIFICANT TAU VALUES FOR 26 FROM 157 RAINFALL RECORDS ANALYSED FOR TREND

Station No.	Tau	Station No.	Tau
157	-0.1715	44	-0.2904
154	-0.1914	41	-0.1787
151	-0.3507	40	-0.1695
143	-0.2350	39	-0.2145
114	-0.2432	37	-0.3200
102	-0 1787	36	-0.2002
93	-0.1705	32	-0.2197
81	+0.1818	24	-0.1900
77	-0.2524	19	+0.3180
75	-0.1828	18	-0.2873
61	-0.1767	13	-0.1695
56	-0.2565	10	-0.1982
48	-0.1951	8	-0.2043

On an assumption that the 157 rainfall records are independent, the expected results under the null hypothesis for significant Tau values is four each positive

TABLE 4.2 - THE TIME DEPENDENCE OF TREND IN SOUTH AFRICAN RAINFALL

Number of stations for given observation periods	Number of stations showing significant trend over different intervals during the observation period			
	1880-1972	1890-1972	1900-1972	1910-1972
1880-1972 : 30	3 (10%)	4 (13%)	5 (17%)	5 (17%)
1890-1972 : 55		8 (15%)	9 (16%)	10 (18%)
1900-1972 : 76			11 (15%)	15 (27%)
1910-1972 : 157				26 (17%)

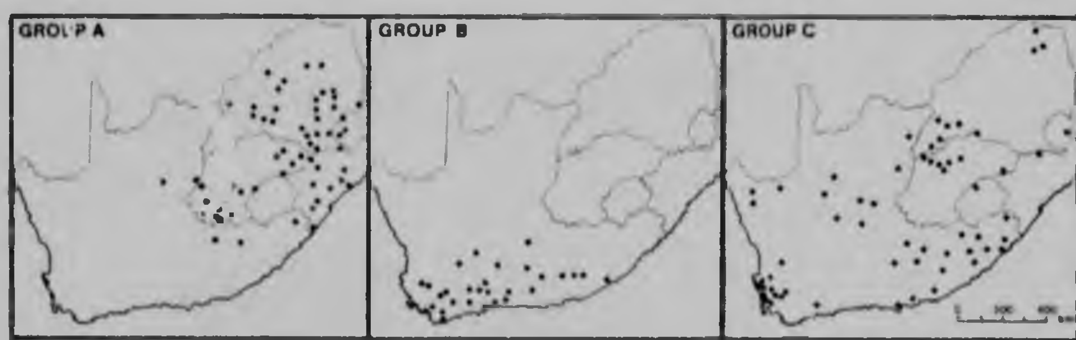


Fig. 4.3: Spatial distribution of the Group A, Group B, and Group C stations.

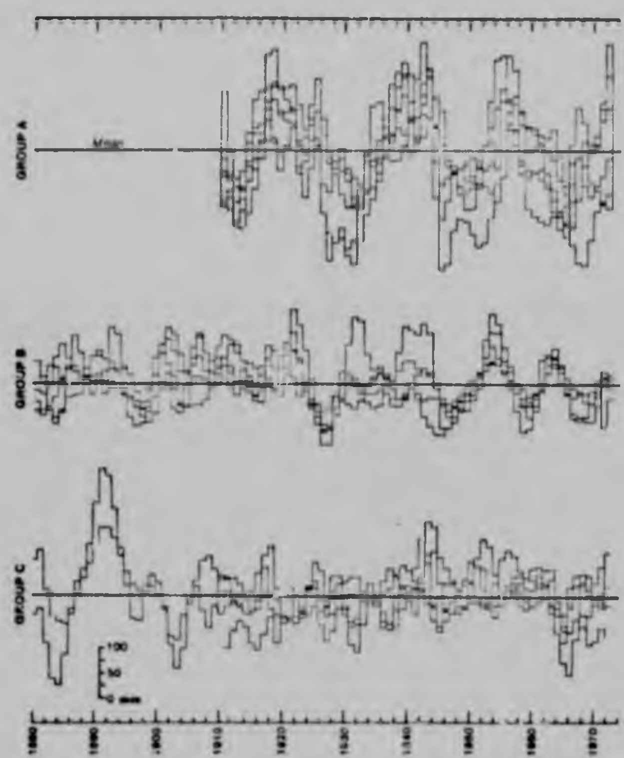


Fig. 4.4: Superimposed deviations from mean annual rainfall smoothed by taking 5-year weighted averages for stations in northern and eastern South Africa (Group A), in the southern Cape Province (Group B) and in remaining areas (Group C). The Group A stations are 85, 86, 103, 130, 139 and 48. Those given in Group B are 27, 34, 39, 46 and 47, and those in Group C are 53, 68, 79, 120 and 121.

and negative at the five per cent level. A one-sample Chi-square test gives the probability of the observed values of Tau occurring under the null-hypothesis of no trend as less than 0.001, with two degrees of freedom ($\chi^2 = 103$). However, although statistical tests can be readily implemented, they are difficult to interpret in this case because of the doubtful validity of the underlying assumption of independence. These results can only be used as an indication of the true situation regarding the presence or absence of trend.

Some station records extend further back in time. When these are submitted to trend analysis, the dependence of trend on length of record is clear, (Table 4.2). In general it appears that as the length of record increases so the tendency for the occurrence of trend decreases, thus showing that the trend is artefactual. The trend is linear and examples displaying its behaviour over time are shown (Fig. 4.2b). Further elaboration of this section is provided in a later chapter dealing with the discussion.

4.3 OSCILLATORY BEHAVIOUR OF THE SMOOTHED SERIES

Smoothing the rainfall series, to remove the influence of noise, shows the emergence of oscillatory patterns which are spatially dependent. Group A stations comprising those from the north and northeastern part of the country exhibit an oscillation with a period of the order 20 years (Fig. 4.3). Stations in the southern Cape Province form Group B, in which rainfall series have an oscillation of the order 10 years. The behaviour of the rainfall of the remaining stations, those in Group C, and notably those from the central area, show no organised patterns for oscillations with periods greater than the order of smoothing, (Fig. 4.4). The Group A station records indicate peaks over the intervals 1918-1922, 1938-1942, and 1958-1962. Corresponding

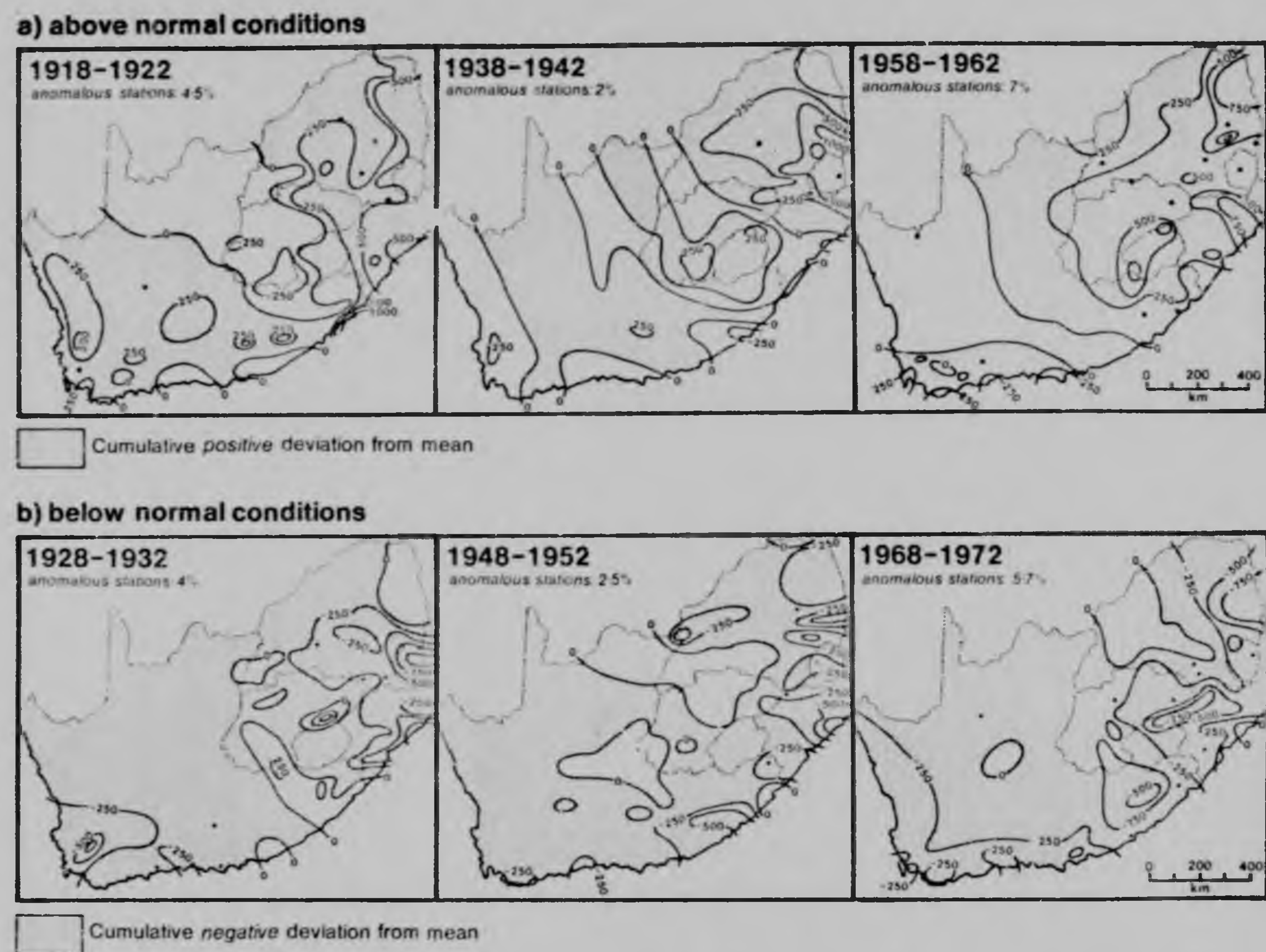


Fig. 4.5: Cumulative filtered positive and negative deviations from mean annual rainfall (in mm) to show decades of above and below normal rainfall conditions. Stations not conforming to the generalised contours are indicated by means of dots.

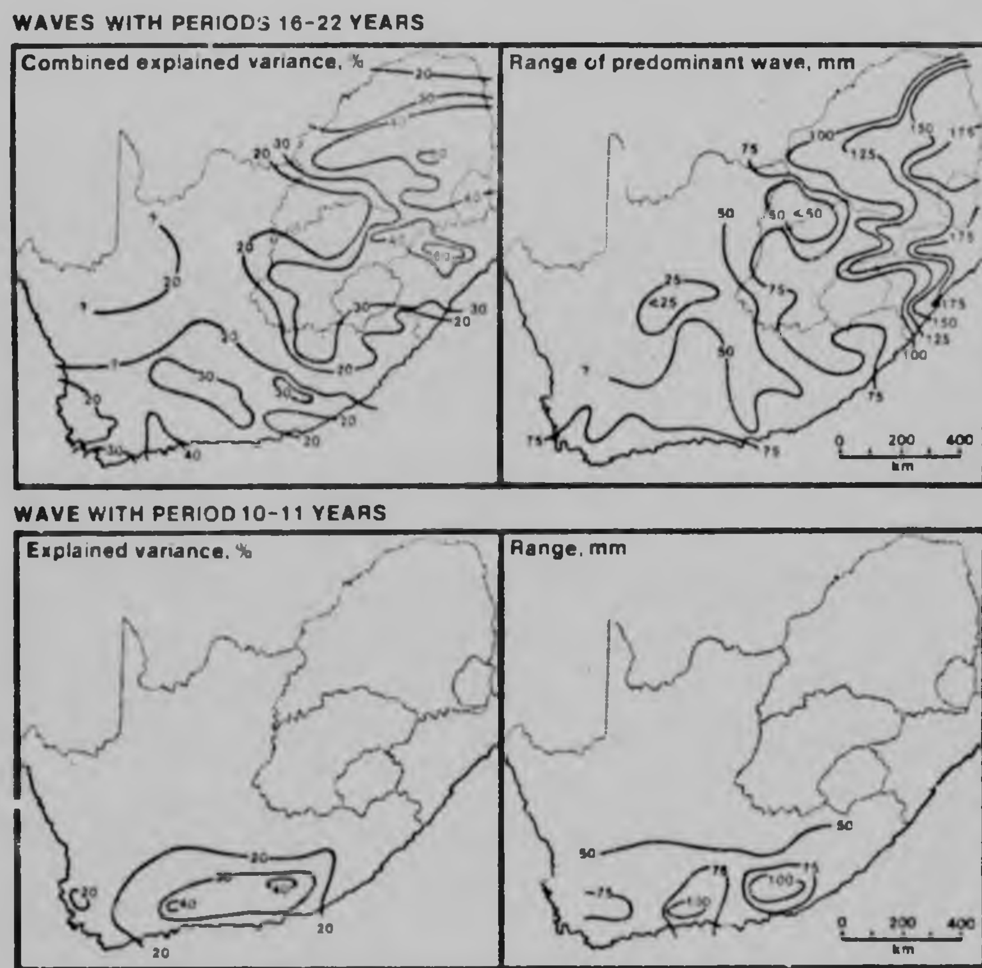


Fig. 4.6: Variance and range associated with Fourier waves having periods in the intervals 16-22 and 10-12 years.

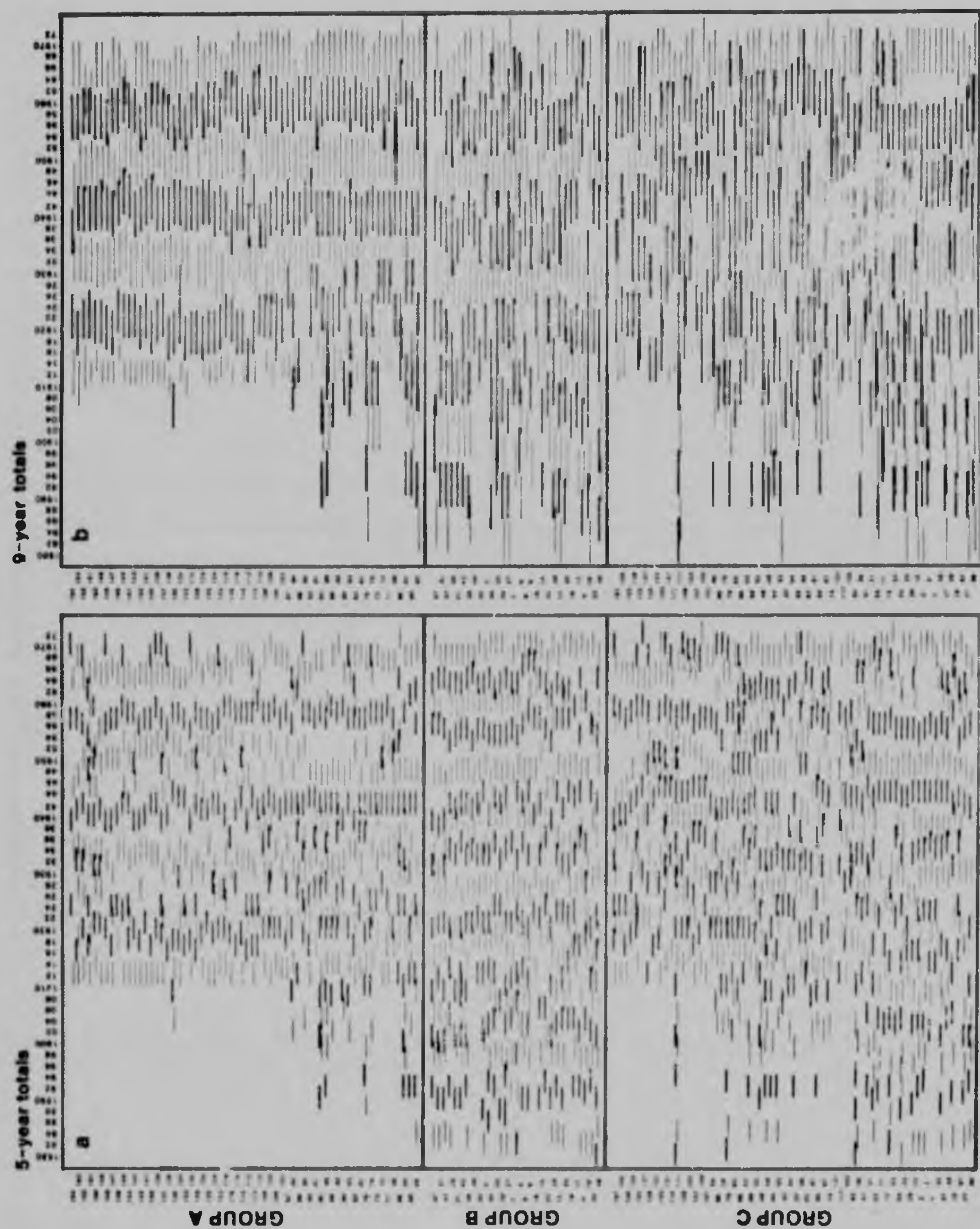


Fig. 4.7: The occurrence of 5-year and 9-year rainfall maxima and minima at 157 stations over South Africa. Heavy lines indicate the periods of maxima and light lines periods of minima. Broken lines indicate weak maxima.

troughs are present during the intervals 1928-1932, 1948-1952, and 1968-1972.

Mapping cumulative positive and cumulative negative deviations for these quinquennial periods reveals the areal extent to which the country as a whole experiences these peaks and troughs in its rainfall budget, (Fig. 4.5). The number of anomalous stations, i.e., those not conforming to the contours of the cumulative deviations, is less than eight per cent of the 157 analysed.

Fitting Fourier series to the filtered data indicates the extent to which periodicities account for the observed fluctuations in the rainfall series, (Panofsky and Brier, 1965). Over the interval 1910-1972 the combined variance associated with waves having periods in the range 16-22 years delimits those parts of the country experiencing a rainfall fluctuation of about 20 years. Both the range of the fluctuation, and the percentage variance accounted for by it, increase towards the north-east. In this region, the range of the fluctuation exceeds 150mm over a wide area. The quasi-10 year oscillation suggested in the Group B stations, is confined to the southern Cape. In few cases does its range exceed 75mm, (Fig. 4.6).

4.4 MAXIMUM AND MINIMUM OF CUMULATIVE DEVIATIONS

A further demonstration of the oscillatory behaviour of the rainfall series can be obtained from a consideration of the occurrence of maximum values of cumulative deviations from the mean, or trend line of the filtered series. When trend was absent in a series, deviations were taken from the mean of the series; when it was present the trend line was taken as the reference level. Cumulative deviations of order 5, and order 9 are taken. The maxima and minima of these series for all 157 stations are mapped, (Fig. 4.7). Stations whose records show

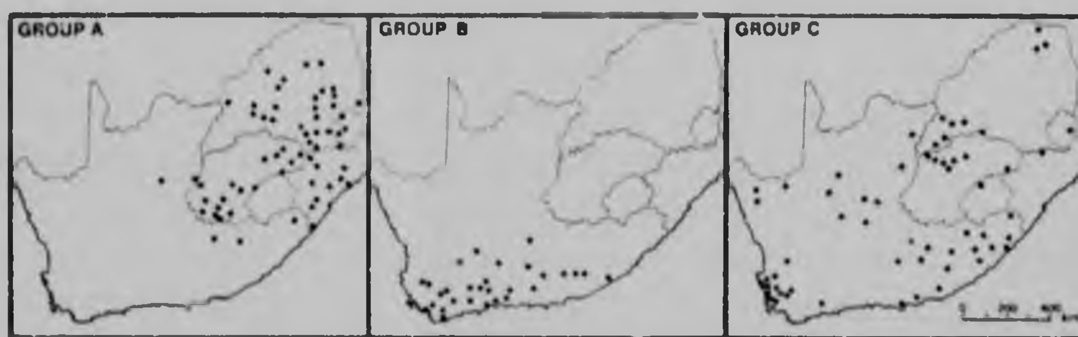


Fig. 4.8: Spatial distribution of the Group A, Group B, and Group C stations shown in Fig. 4.7.

similar time intervals in which maximum values and minimum values occur are grouped together, (Fig. 4.7). The spatial distribution of the stations forming these groups is shown, (Fig. 4.8).

A 16-20 year fluctuation is evident from the spacing of the 9 yearly maxima and minima for Group A stations, numbers 94 - 157. The maxima are centred about the intervals 1920-22, 1940-42, and 1958-60. For the Group B stations the pattern is still discernible, though with less clarity than in Group A. No recognisable overall pattern is obvious for Group C.

The peaks and troughs based on 5-year cumulative deviations for the Group B stations show the extent to which a quasi 10-year oscillation is present. The line of demarcation between Group A and Group B is to a large extent geographical, as is that between Group B and Group C. The peaks based on 5- year cumulative totals for Group A divide near the boundary with Group B. There is a tendency for them to regroup in the region of the boundary between Group B and Group C. This is particularly noticeable for the maxima centred on 1940-42, and 1958-60.

4.5 A SIMULATION STUDY

For purposes of comparison with the results obtained in section 4.4, 157 rainfall records were simulated and subjected to an identical analysis. Pseudo random numbers following the rectangular distribution over (0.1) were generated by means of the Tansworthe generator (Whittlesey, 1968). Each number was then used to produce a normal (0.1) random variable using the inverse method, (Tocher, 1963). The transformation was made using a table lookup and linear interpolation in the range (0.001, 0.999). Outside this range the Newton-Raphson algorithm was used (Hildebrand, 1956).

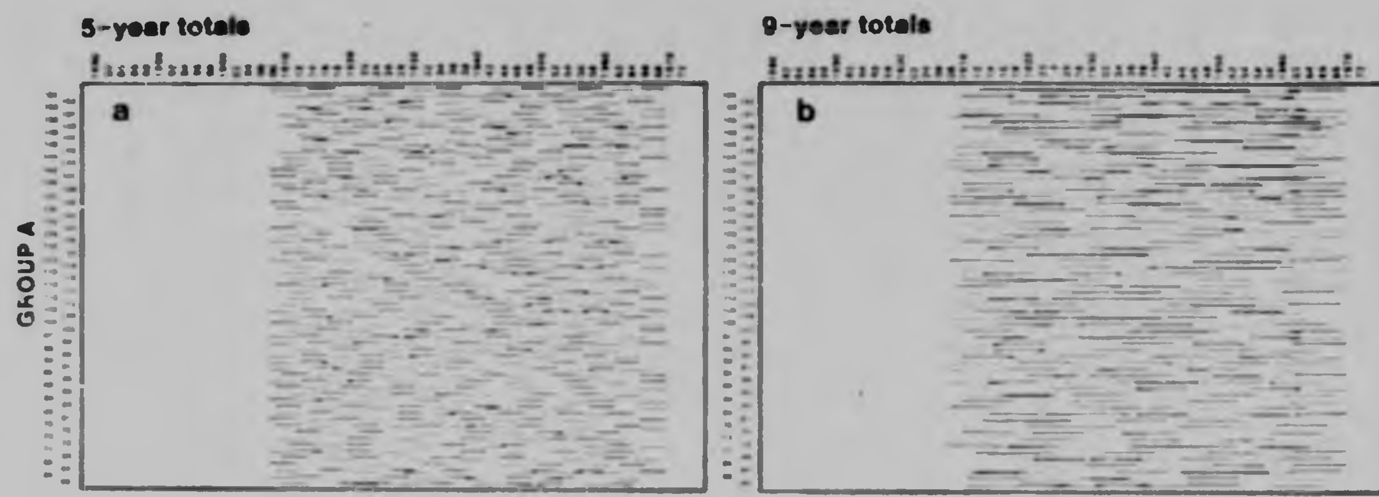


Fig. 4.9: The occurrence of 5-year and 9-year simulated rainfall series for Group A stations.

The mappings of these simulated series is completely devoid of organised patterns. Those for the Group A stations illustrate this clearly, (Fig. 4.9).

4.6 CORRELATIONS BETWEEN ABOVE AND BELOW NORMAL RAINFALL

By determining the frequency with which the rainfall maxima and rainfall minima shown in Fig.4.7 occur within specific time intervals, the inverse correlation between periods of above and below normal rainfall in Group A and Group B stations can be demonstrated, (Fig. 4.10). Similarly, the lack of correlation in the case of the Group C stations and the simulated series is at once evident. Pearson's correlation coefficient between 9-yearly maxima and minima, for Group A stations with 1910-1972 data, is -0.97. That between the 5-yearly maxima and minima for Group B stations is -0.93. For the Group C stations the coefficients are 0.03 and -0.29 for the 9- and 5-yearly totals respectively. No stations in Group A have records extending back to 1880. Group B stations with records for the period 1880-1972 show a correlation coefficient between 5-yearly maxima and minima of -0.84. Even though the basic conditions for the proper use of Pearson's correlation coefficient are not satisfied, the results obtained are nevertheless useful. The coefficients provide a measure, or indication, of the degree of linear agreement between the variables considered (rainfall series).

4.7 SPECTRAL ANALYSIS

Oscillatory patterns in South African rainfall having been established for smoothed data, attention is redirected to the raw data. A theoretically more sophisticated approach than those adopted so far for stochastic processes is spectral analysis. It provides a means of

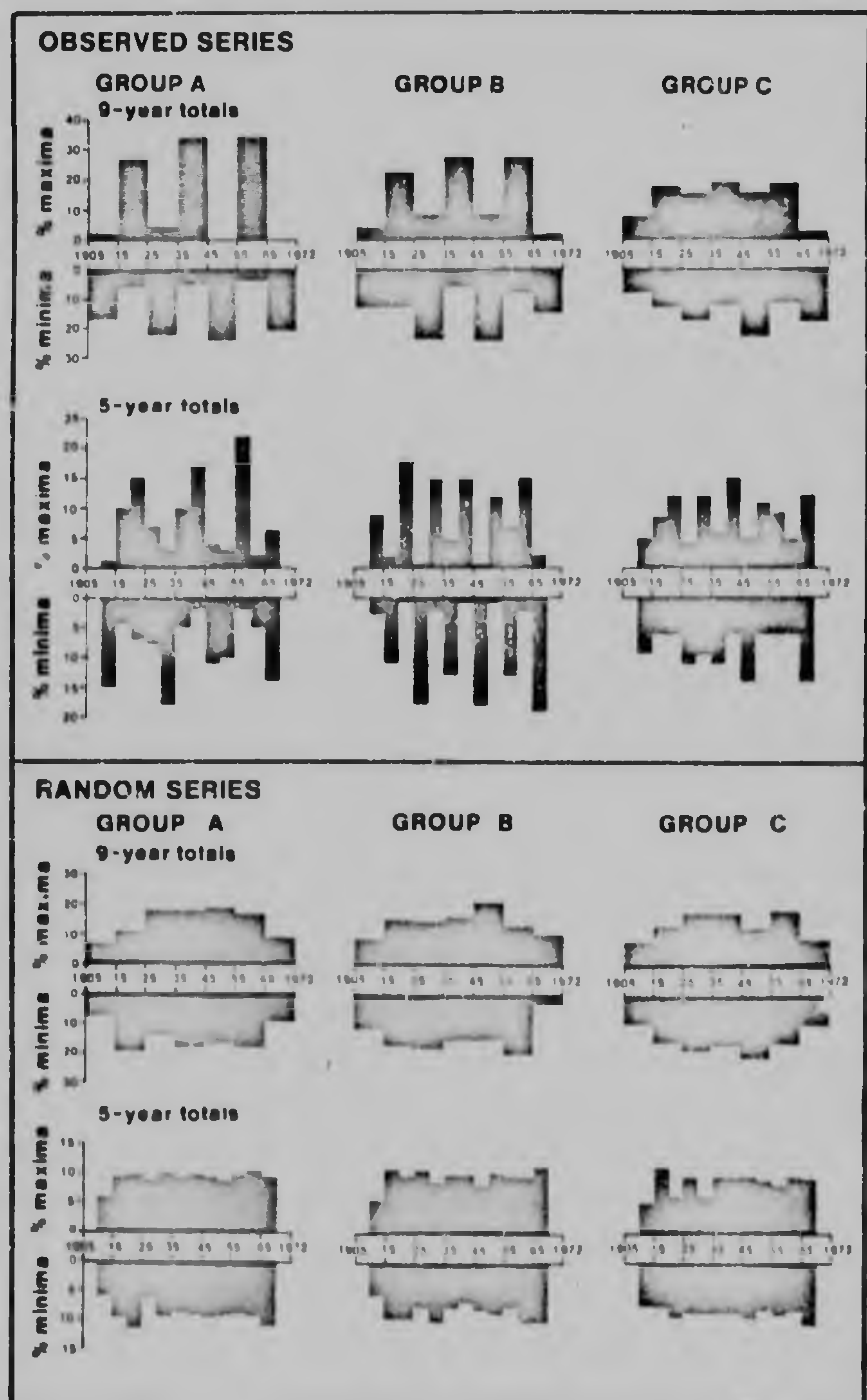


Fig. 4.10: Frequency analysis of the periods of maximum and minimum rainfall shown in Fig. 4.7.

Fig. 4.11: Correlograms for Station 148 (an example of the Group A type), for Station 27 (a Group B type) and Station 79 (a Group C type).

investigating the physical mechanism generating a time series; determines the behaviour of a dynamic linear system in response to random excitations; and aids in the simulation of time series for the purpose of building forecasting models (Parzen, 1961).

Correlograms for Group A stations are typified by that for station 148; for Group B by station 27; and for Group C by station 79, (Fig. 4.11). Regular fluctuations are clearly indicated for Group A and Group B stations. Superimposed on a wave with a period of approximately 18 years, for Group A, is a ripple of order less than five years. For the Group B stations a wave with period ten years is evident, together with a ripple of shorter period. The correlograms for Group C stations reflect the absence of waves having periods of the order found in Group A and Group B. There is, however, evidence to suggest regular fluctuations in the higher frequency waveband.

Each group of stations possesses individual spectral patterns, (Fig. 4.12). The spectra present, in an alternative form, the information given by the correlograms. For the Group A stations, the spectra have peaks in the 16-20 year interval, and the 2-3 and 3-4 year bands. The spectral ordinates, at 20 years, are significant at the five per cent level for stations 103, 116 and 148; significance at the ten per cent level applies to station 130.

The spectra belonging to the Group B stations peak in bands containing 10-year and 3-4 year oscillations. For the peaks centred at the wavelength of 10 years, significance is obtained at the five per cent level for station 27, and at the ten per cent level for stations 34, 39 and 46.

Lack of significance of the spectral ordinates does not necessarily imply lack of reality, (Jenkins, 1961;

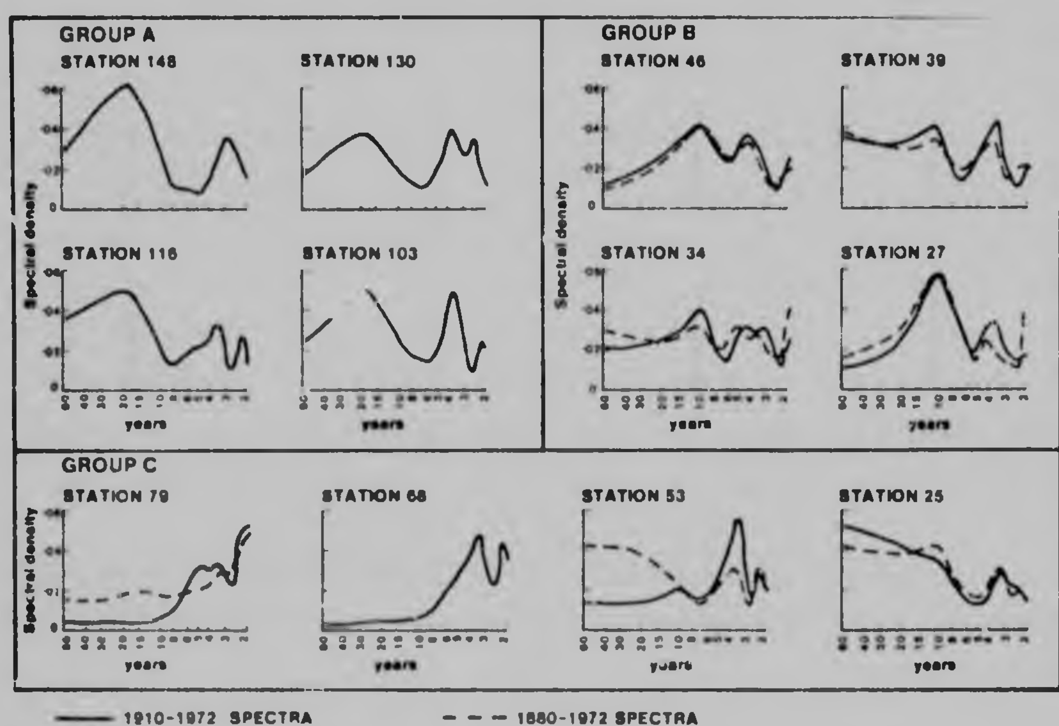


Fig. 4.12: Rainfall spectra for some Group A, Group B and Group C stations over the periods 1910-1972 and 1880-1972.

Granger and Hatanaka, 1966). The reality of the peaks is strongly suggested by the similarity in shape of large numbers of spectra. That this view is correct, is strongly suggested by the previous analysis of rainfall maxima and minima, (Fig. 4.5, 4.7). The stability of the 10 year ordinate is evident from the comparison of the 1910-72 and 1880-1972 spectra.

The spectra for Group C stations do not all possess similar characteristics, (Fig. 4.12). Large subsets of the group do, however, possess similar spectra. Typical spectra for stations in the south-western,north-eastern and northern (interior) Cape Province are those for stations 25, 53, 68 and 79 respectively, (Figs. 4.1, 4.12). In some spectra, maximum power is concentrated at the high frequencies, whilst in others at low frequencies.

In almost all spectra, irrespective of whether from Group A, B or C stations, peaks occur in the 2-3 and 3-4 year ranges.

In general, spectra are independent of their time histories. Those for the interval 1910-72 are substantially the same as those obtained for the larger interval 1880-1972. This indicates that the series are of sufficient length to give a true overall picture of the distribution of variance over frequency. The increase in power for the 1880-1972 spectra, at the low frequency end, is a result of a generally wet decade from 1880 to 1890, (Table 4.2, Fig. 4.2b).

Mapping spectral densities of integrated frequency bands enables an assessment of the regional organisation and spatial gradients of specific spectral components to be made, (Fig. 4.13). Contours associated with rainfall oscillations having periods greater than 20 years are fairly flat over the country as a whole. The exception

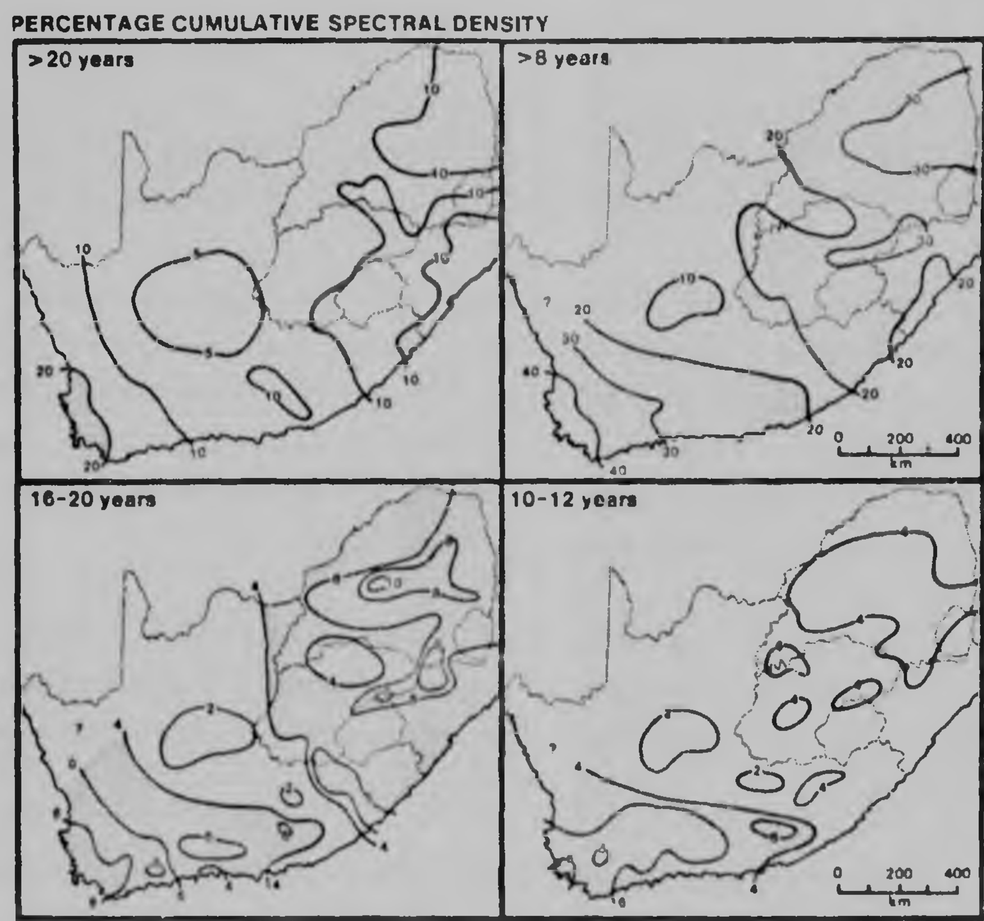


Fig. 4.13: Percentage cumulative spectral density associated with rainfall oscillations greater than 20 years, greater than 8 years, in the range 16-20 years and the range 10-12 years.

is the south-western Cape (a subset of Group C) where the contour values are highest, accounting for twenty per cent of the total variance. In the interior, the contour values for this oscillation are at their lowest. Oscillations with periods greater than eight years are indicated over two major portions of the country. The gradients are negative from the south-western Cape Province toward the interior, and from the Transvaal toward the interior.

Taking a specific interval of 16-20 years, most spectral power associated with this band lies over the central Transvaal. The spectral power for this band is of almost equal intensity over the south-western Cape, being virtually non-existent in the interior and associated corridor to the eastern coast, (Fig. 4.13).

The 10-12 year oscillation is noticeably weaker and is confined to the southern Cape Province. Over the remainder of the country the contours for this wave show no relief.

Oscillations having periods in the 3-4 year interval are ubiquitous, showing little regional organisation; the associated spectral gradients are weak, (Fig. 4.14). The quasi-biennial oscillation shows a highly distinctive spatial organisation. Maximum power associated with this wave occurs over the interior and southern Transvaal, (Fig. 4.14). Furthermore, the percentage variance attributable to this wave is, in general, stronger over all areas of the country than any other wave.

* * * * *

About one rainfall record in six shows significant trend, 24 being negative and 2 being positive in a total of 157 series analysed. Lack of independence between

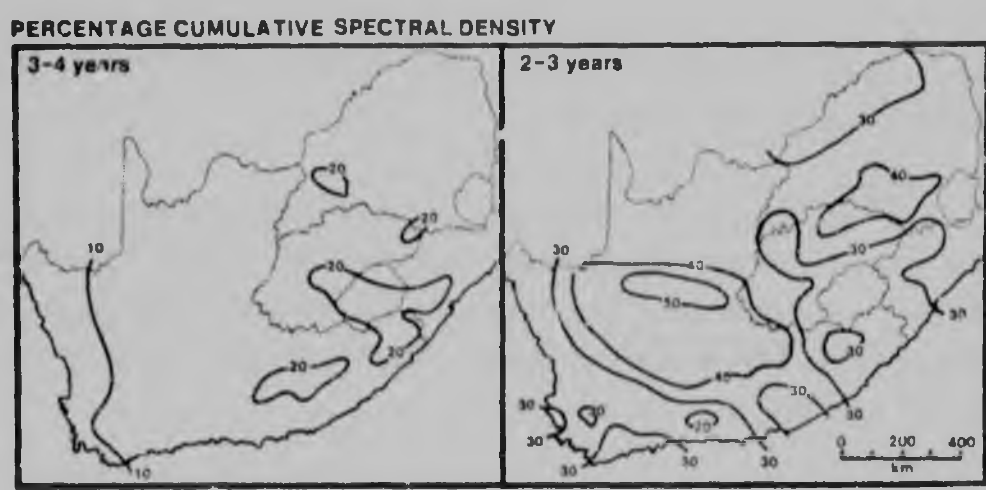


Fig. 4.14: Percentage cumulative spectral density associated with rainfall oscillations having periods in the 3-4 year and 2-3 year ranges.

rainfall records has been established, thus rendering statistical tests inconclusive. In many cases trend is shown to be time dependent. Various methods provide evidence in favour of relatively long wave variation in rainfall (greater than 20, 16-20 and 10 years) that have distinctive regional patterns. A 3-4 year wave is ubiquitous over the country. Although the quasi-biennial wave shows strong regional patterns, it is in evidence over all parts of the subcontinent.

It is unlikely that any of the waves mentioned above are purely local phenomena, because of their recorded presence over other regions of the earth's surface (Chatfield and Pepper, 1971; Lamb, 1972; Stringer, 1972; Barry and Perry, 1973).

CHAPTER 5

A MESO-SCALE ANALYSIS OF SOUTH AFRICAN RAINFALL

5.1 INTRODUCTION

Principal component analysis has been used here to replace the inter-dependent rainfall series with a set of orthogonal variables accounting for most of the variance in the original rainfall series. It provides an objective means of delimiting the country into independent rainfall regions.

Fitting linear two-way unreplicated models to the results obtained from the principal component analysis enables large regions of the country to be represented by single independent time series. The models also decompose these individual series into three component effects namely, a regional temporal, a regional spatial and a residual micro-scale effect. Spectral analysis on the regional temporal series enables the oscillatory behaviour of rainfall on a regional scale to be further demonstrated.

5.2 PRINCIPAL COMPONENT ANALYSIS

In the first instance a straight forward principal components analysis was carried out on the 157 rainfall

TABLE 5.1 - VARIANCE EXTRACTED BY EACH OF THE FIRST 29 EIGENVALUES TOGETHER WITH THE ASSOCIATED CUMULATIVE PERCENTAGE OF THE TOTAL VARIANCE

Eigenvalue No.	Variance extracted	Cumulative percentage of total variance
1	43.88	27.95
2	17.12	38.86
3	13.15	47.23
4	8.95	52.94
5	6.53	57.10
6	5.15	60.38
7	3.95	62.90
8	3.77	65.30
9	3.12	67.29
10	2.95	69.17
11	2.73	70.91
12	2.49	72.50
13	2.33	73.98
14	2.21	75.39
15	2.07	76.71
16	1.95	77.95
17	1.78	79.09
18	1.67	80.15
19	1.65	81.21
20	1.53	82.18
21	1.50	83.13
22	1.42	84.04
23	1.35	84.90
24	1.32	85.74
25	1.26	86.54
26	1.20	87.30
27	1.10	88.00
28	1.06	88.68
29	1.04	89.35

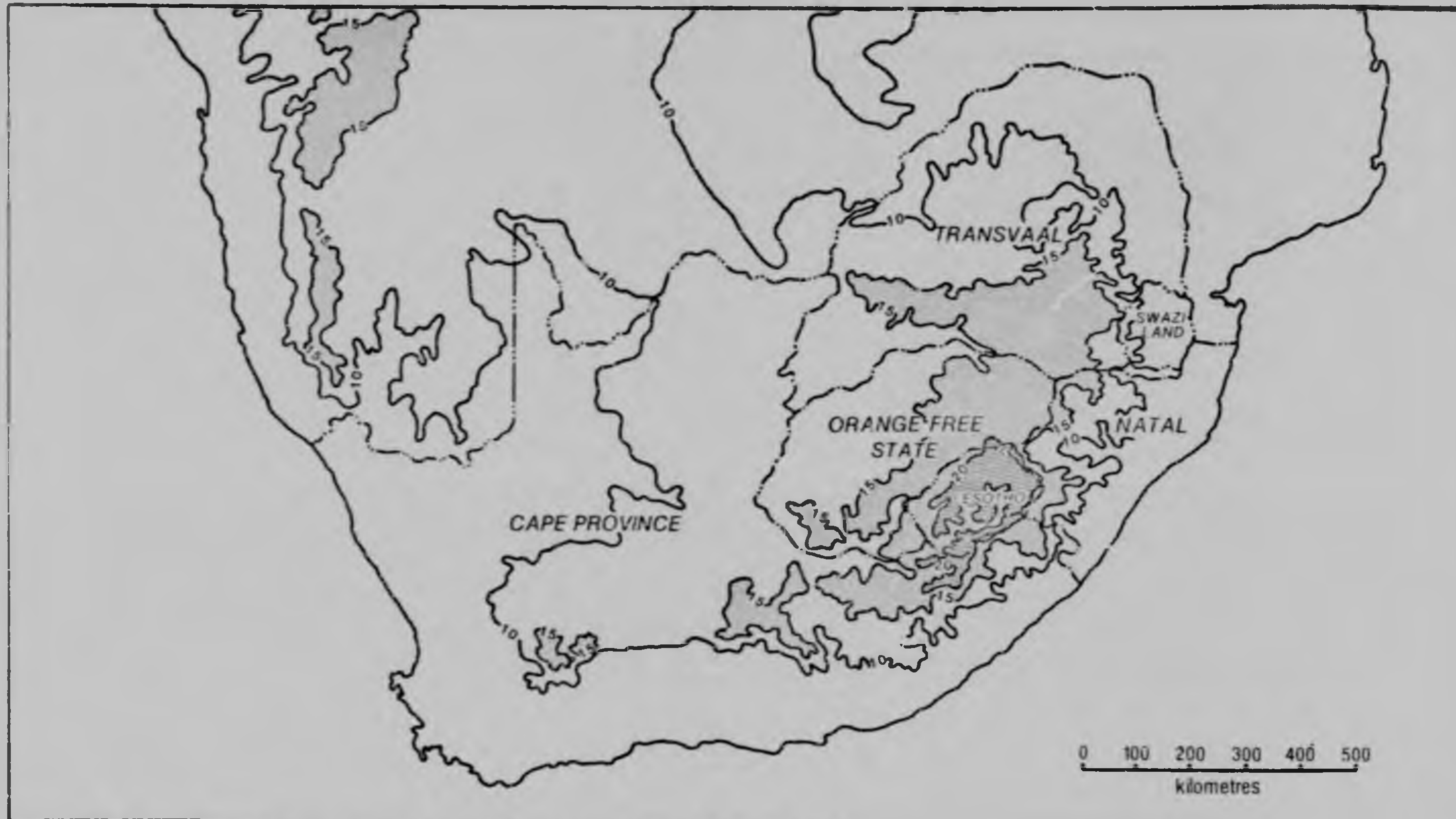


Fig. 5.1: Showing a simplified topography of South Africa.

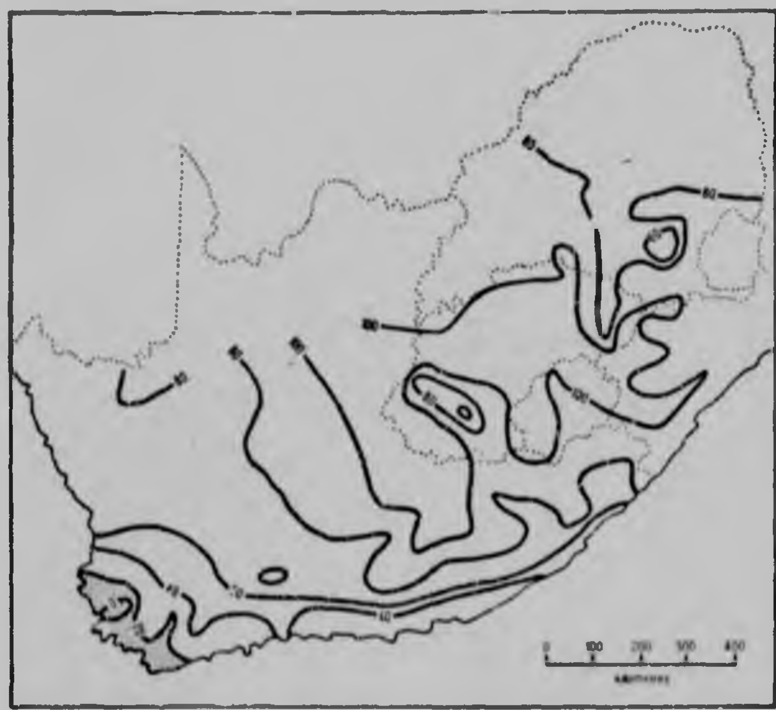


Fig. 5.2: Contours of the first eigenvector show little relief over most of South Africa. But for a small group of stations in the south-western Cape Province , this vector shows high association with rainfall records throughout the country.

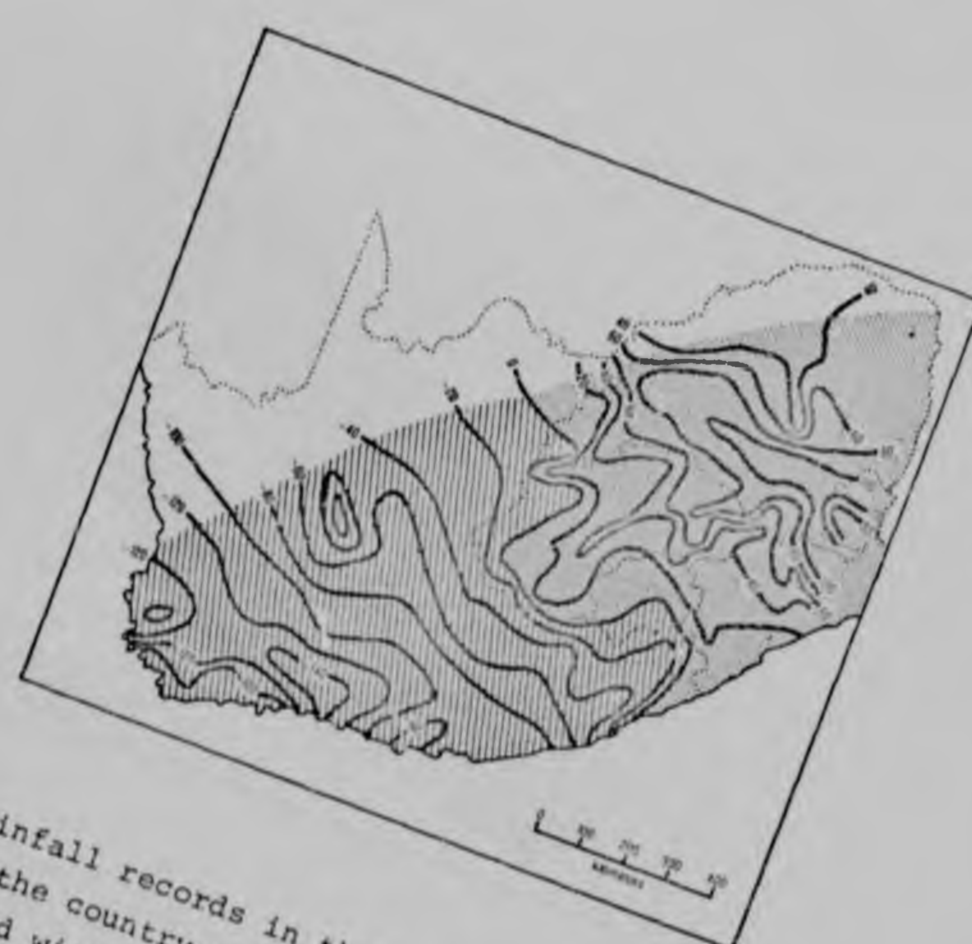


Fig. 5.3: Rainfall records in the north-eastern part of the country ▒▒ are positively correlated with this vector, whilst those in the south-western area |||||||||||| are negatively correlated with it, with strong gradients existing over the furthermost region of the sub-continent.

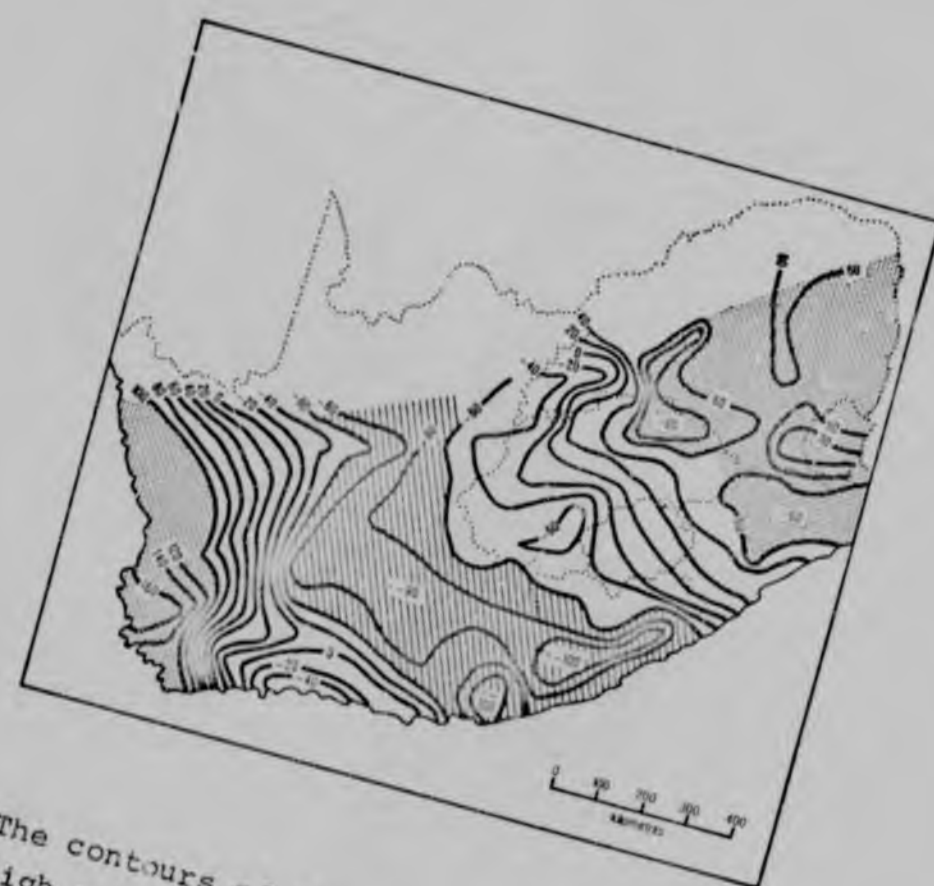


Fig. 5.4: The contours of the third eigenvector show high positive correlation between rainfall records in the north-eastern and south-western parts of South Africa, and this vector. Stations contained in a corridor stretching across the country have records that possess high negative correlation with the third eigenvector .

series. The first three eigenvalues extracted 47 per cent of the total variance in the original rainfall series. Using Kaiser's criterion, which rejects eigenvalues of magnitude less than unity, as a cut-off, the first 29 eigenvalues accounted for 89.35 per cent of the total variance, (Table 5.1).

Mapping coefficients of the eigenvectors for each principal component shows the level of correlation between each station's rainfall and the newly derived orthogonal variables. The contour field for the first eigenvector shows little relief over the major portion of the country. For a small area of the south-western Cape Province, (Fig.5.1), correlation between rainfall and the first principal component is considerably less than elsewhere, (Fig. 5.2). The north-eastern half of the country is positively correlated with the second orthogonal variable, whilst negative correlation exists between it and stations in the south-western half, (Fig. 5.3). Strong gradients are present over the contour field for the third principal component, (Fig. 5.4). The south-western coastal rea is highly correlated with this third orthogonal variable, with a similar situation existing for stations in the Transvaal and northern Free-State. A corridor from the central interior to the east coast, representing the region of transition between predominantly cyclonic activity to the south and anticyclonic to the north, has a large negative correlation with this same variable. The diagrams shown so far, indicate the possible usefulness of rotating the vectors. This follows from the suggested alignment of stations in different parts of the country with specific eigenvectors.

An aim in rotation is to decrease ambiguities that may have arisen in the direct analysis. For example, the first eigenvector has high coefficients over an area of the country known to experience two different oscillations in annual rainfall totals. These are the 16-20 year wave to

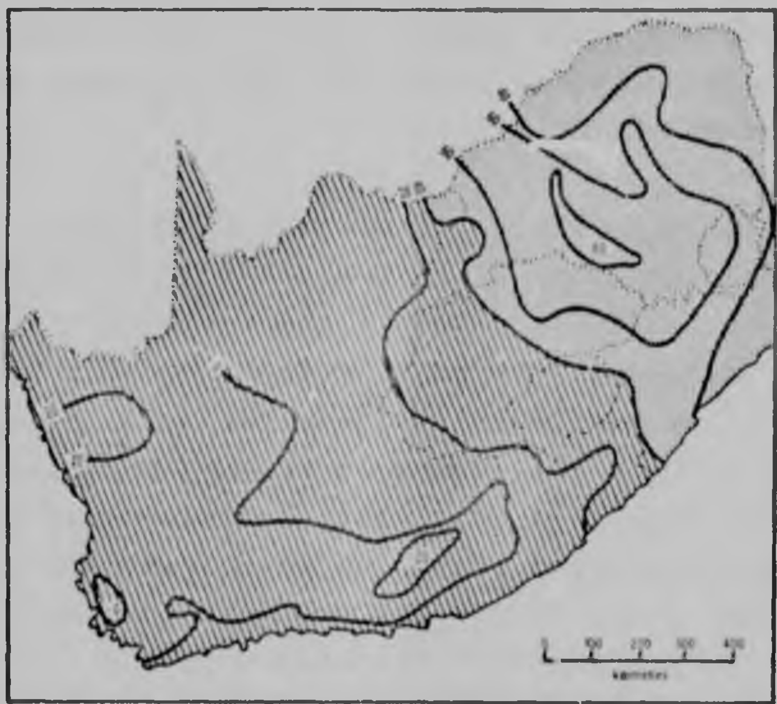


Fig. 5.5: The first rotated factor shows close association only with rainfall records for the north-eastern part of South Africa ▒▒▒. Elsewhere, little correlation exists between this factor and rainfall records |||||||||||||.

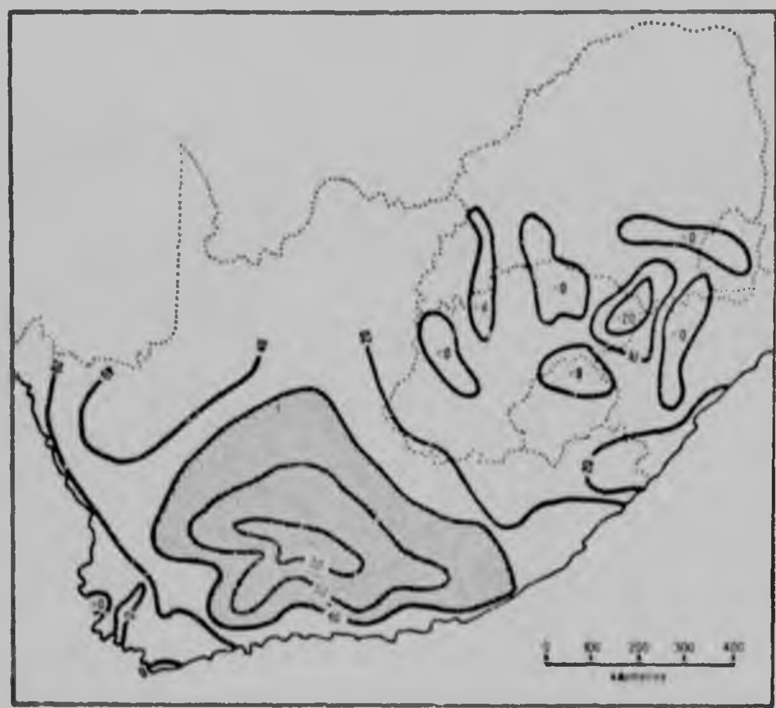


Fig. 5.6: Station rainfall records in the southern and interior Cape Province are highly correlated with the second rotated factor as indicated by the contour pattern for this vector ▒▒▒▒ . In no other area of the country does this vector account for significant variance in the rainfall records.

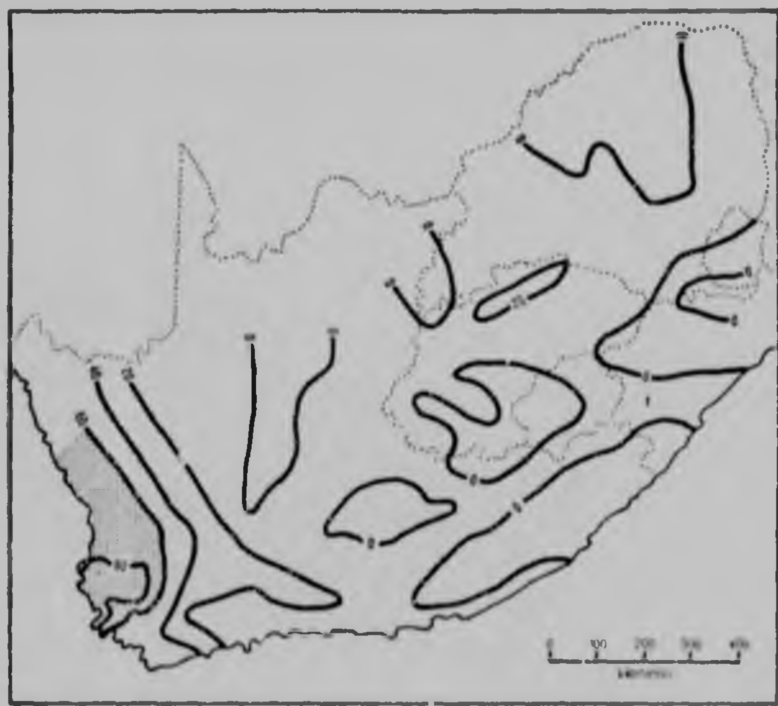


Fig. 5.7: The contour field of the third rotated factor shows that it accounts for significant variance only in those rainfall records belonging to stations situated along the south-western Cape coast .

the north of the country and the 2-3 year one over the interior. This further clarification is accomplished if the rotated factors contain high loadings on some original variables and low loadings on others. Middle value loadings will then have altered and moved into one or other of these categories, namely high or low. The total percentage extracted variance remains unchanged upon rotation; it may be redistributed amongst the factors.

Mapping the loadings of the first rotated factor reveals a high level of correlation between stations in the northern part of the country and this factor,(Fig. 5.5). Stations giving rise to a ridge of high loadings over the central and southern Transvaal give way to stations with uniformly low loadings over the southern Free-State and the remainder of the country, including the northern Transvaal. Loadings on the second rotated factor group stations in the southern interior, (Fig. 5.6). The ridge formed by these loadings falls away quite rapidly in directions on all points of the compass. Mapping the loadings for the third rotated factor shows that only a small group of stations, contained in the south-western coastal strip, possesses high correlation with this factor, (Fig.5.7). Over the remainder of the country the contour field is fairly uniform, and low in value. Along the east coast two groups of stations are seen to have negative correlations with this factor. They are separated by a small area, adjacent to Lesotho, having low but positive values. A similar group of stations appear in the southern Free-State. However, the values are low in absolute value and may have arisen from sampling fluctuations.

Rainfall station groupings are suggested by the results shown in Figs. 5.5 to 5.7 inclusive. Each rainfall series has 29 rotated factor loadings. For each station the factor with the highest loading on that station has been noted. It was found that homogeneous groups of stations were formed.

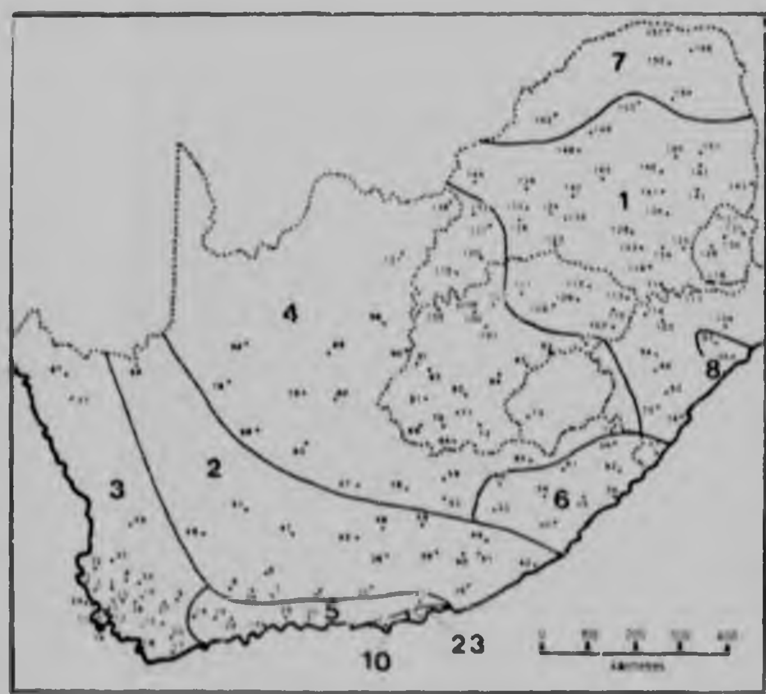


Fig. 5.8: Showing the rainfall recording station groupings according to rotated factors as criteria. The numerals denote dominant factors within each group of stations.

The homogeneity arises from spatial clustering of stations having their highest loadings on the same factor. This analysis results in the formation of ten distinct rainfall regions, (Fig. 5.8).

Factor seven is the dominant one in the group of stations comprising the far northern Transvaal. Stations in the remaining part of the Transvaal, the northern Free-State and much of Natal are most closely correlated with the first factor. Two eastern coastal regions, a small one midway along the Natal coast, and another larger one in the north-eastern Cape Province exhibit the eighth and sixth factors respectively as dominant. Factor two accounts for most of the explained variance of stations forming a belt stretching from the north-western interior to the east coast. The southern coastal stations form a group most closely related to the fifth factor, and contained in this region are two stations represented by factor ten, and twenty three. The south-western Cape Province and west coast are best delimited regionally by factor three.

All the regions shown in Fig. 5.8 will henceforth be referred to by the dominant factor number within the region. In no case was a spatially non-homogeneous array of stations present; each factor describes a unique geographical grouping of stations experiencing like rainfall changes over time.

5.3 PARSIMONY OF RAINFALL SERIES WITHIN REGIONS

Homogeneous groups of stations having been formed further analyses will be simplified if each group can be represented by a single time series. Fitting a linear two-way unreplicated model to the station rainfall series within a region results in a regional temporal effect, regional spatial effect, and residual effect for each region. The residuals represent effects on a local or micro-scale that are unexplained by the general meso effect of a region. Regions 8, 10 and 23 have been excluded from the following analysis because of problems arising from too few degrees of freedom. Furthermore, it was considered that these stations constituted anomalies.

A two-way analysis of variance (ANOVA) on the stations' records within each of the seven regions demonstrates that the regional temporal effect, A, is significantly different from zero, and thus provides an explanatory effect on the rainfall series. The level of rejection for the null-hypothesis, H_o, is taken as 0.01; the information relating to the ANOVA is given in Table 5.2. In all cases, the probabilities of the calculated F values having occurred by chance were less than 0.001 under H_o: $all\ \alpha_i = 0$.

TABLE 5.2 - SUMMARY OF THE ANOVA FOR AN UNREPLICATED FIXED TWO-WAY MODEL ON REGIONS 1 TO 7; H_o: $all\ \alpha_i = 0$.

Region	F	ν_1	ν_2
1	39.68	62	2666
2	18.23	62	1178
3	20.42	62	1426
4	41.34	62	2418
5	7.05	62	372
6	9.78	62	558
7	7.27	62	248

Similarly, the ANOVA demonstrates that H_o: $all\ \beta_i = 0$. can be rejected at a probability level of less than 0.001 (Table 5.3). This result justifies the assumption that the rainfall stations have real micro-climates, the behaviour of which can be determined from analyses of the residuals in the model.

TABLE 5.3 - SUMMARY OF THE ANOVA FOR AN UNREPLICATED FIXED TWO-WAY MODEL ON REGIONS 1 TO 7; H_0: *all* $\beta_i = 0$

Region	F	ν_1	ν_2
1	127	43	2666
2	7333	19	1178
3	390	23	1426
4	274	39	2418
5	23911	6	372
6	2846	9	558
7	1621	4	248

5.4 TREND IN THE REGIONAL TEMPORAL SERIES

Testing each of the regional temporal series for trend over the period 1910-1972 shows that no region is undergoing long term significant positive or negative trend, (Table 5.4). A two-tailed test has been used, with a significance level of five per cent. The critical 't' statistic value for rejecting the null-hypothesis is 2.00 (H_0 : $r(s) = 0$).

The probability of obtaining one positive and six negative $r(s)$ values in a total of seven can be estimated by treating the positive and negative values as successes and failures respectively in a binomial situation. This action is justified because of the manner in which the regional temporal effects are defined. The probability that the distribution of positive and negative signs in Table 5.4 arose by chance is 0.11, based on the assumption that either sign was equally likely to arise.

TABLE 5.4 - SPEARMANS $r(s)$ VALUES WHEN USED TO TEST FOR TREND IN THE REGIONAL TEMPORAL SERIES FOR EACH OF THE REGIONS 1 TO 7.

Region	$r(s)$	't' statistic
1	-0.08	-0.66
2	-0.22	-1.80
3	-0.16	-1.29
4	0.05	0.37
5	-0.02	-0.20
6	-0.13	-1.06
7	-0.12	-0.95

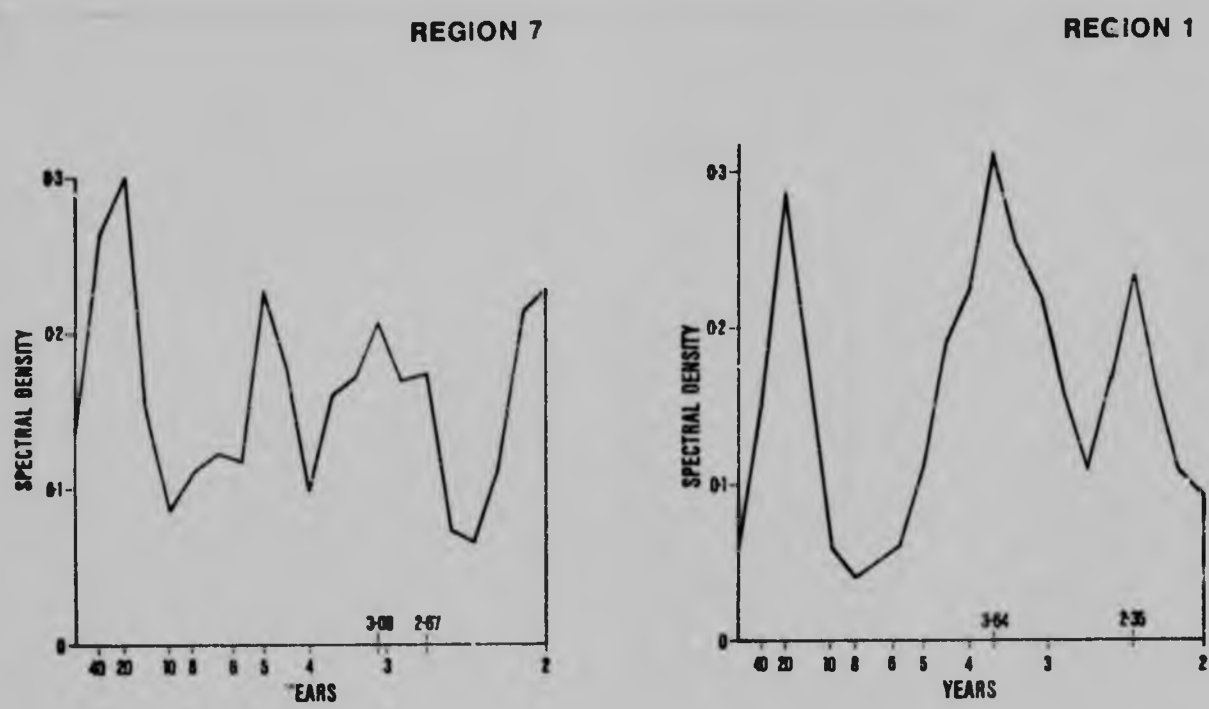
5.5 TREND IN THE MICRO-EFFECT SERIES

Some residuals from the unreplicated two-way linear models show significant trend, (Table 5.5). In a total of 34 stations having significant trend on a micro-scale, 16 show positive trend and 18 negative trend.

The one sample Chi-square test can be used in this instance to test the significance of these results because of the essential independence of the residuals, (Table 5.5). Under the null-hypothesis of no trend, the distribution of 16 positive and 18 negative trend values can be expected to arise by chance with a probability of less than 0.001 (χ^2= 96 with two degrees of freedom). The distribution is random of stations with positive and negative trend over the country (Table 5.5, Fig. 4.1). In some cases adjacent stations have

TABLE 5.5 - STATIONS HAVING SIGNIFICANT MICRO-SCALE (LOCAL EFFECT) TREND IN EACH OF THE REGIONS 1 TO 7

Region	No. of stations in region	Particular stations showing trend	
		Positive	Negative
1	47	151, 143, 114, 102, 75	147, 141, 140, 128, 83
2	20	44, 39	51, 34
3	26	77, 37, 32, 24, 18	23, 19, 17, 12
4	40	97, 93	91, 81
5	7	14	29, 15
6	10	56	62, 55
7	5	-	152
TOTALS	157	16	18



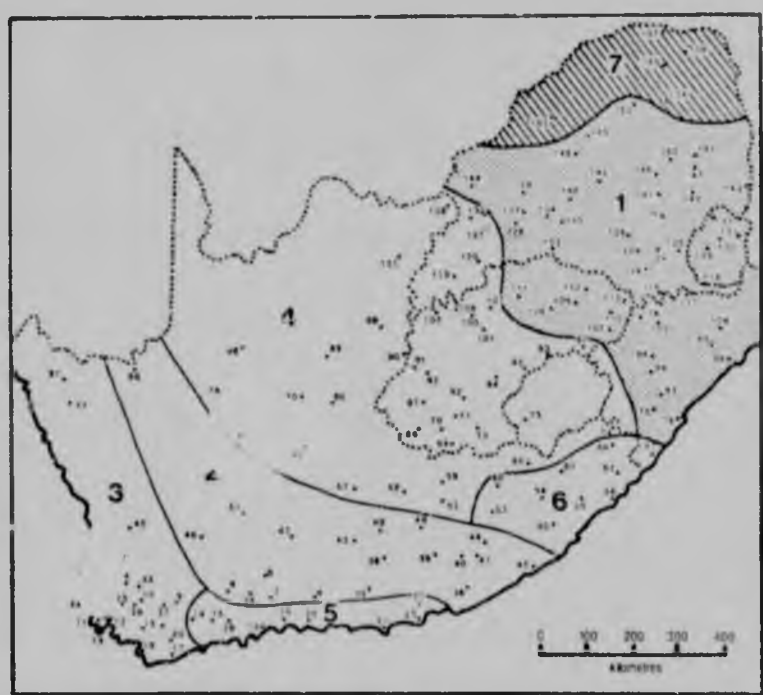


Fig. 5.9: Showing the spectra of the regional temporal series for Regions 7 and 1, together with the regions' geographical locations.

records whose micro-effects show trend of opposite sign; this reflects the true local nature of the micro-effects.

5.6 THE OSCILLATORY BEHAVIOUR OF THE REGIONAL TEMPORAL SERIES

That the regional temporal series are oscillatory in behaviour is demonstrated by the significant peaks in their spectra, (Table 5.6). Clearly defined peaks with wavelengths of 20 years are present in the spectra for the northern and central Transvaal, Regions 7 and 1 respectively. These regions are those having summer seasonal rainfall, (Fig. 5.9).

TABLE 5.6 - WAVELENGTHS AT WHICH PEAKS ARE CENTRED IN THE SPECTRA OF THE TEMPORAL SERIES FOR REGIONS 1 TO 7.

Region	Peaks (wavelengths in years)
1	20**, 3.64**, 2.35*
2	10**, 3.64**, 2.86, 2.10
3	40**, 5, 3.33, 2.67*
4	20, 4.44*, 3.64**, 2.67, 2.35
5	20, 10, 4**, 2.86**, 2.22
6	20, 6.67, 3.64**, 2
7	20**, 5, 3.08, 2.67, 2

(**, * indicates significance at 5 and 10 per cent levels respectively, against an hypothesis of a random generating process)

Areas of the south and south-western Free-State, and north-eastern Cape Province, Regions 4 and 6, have temporal

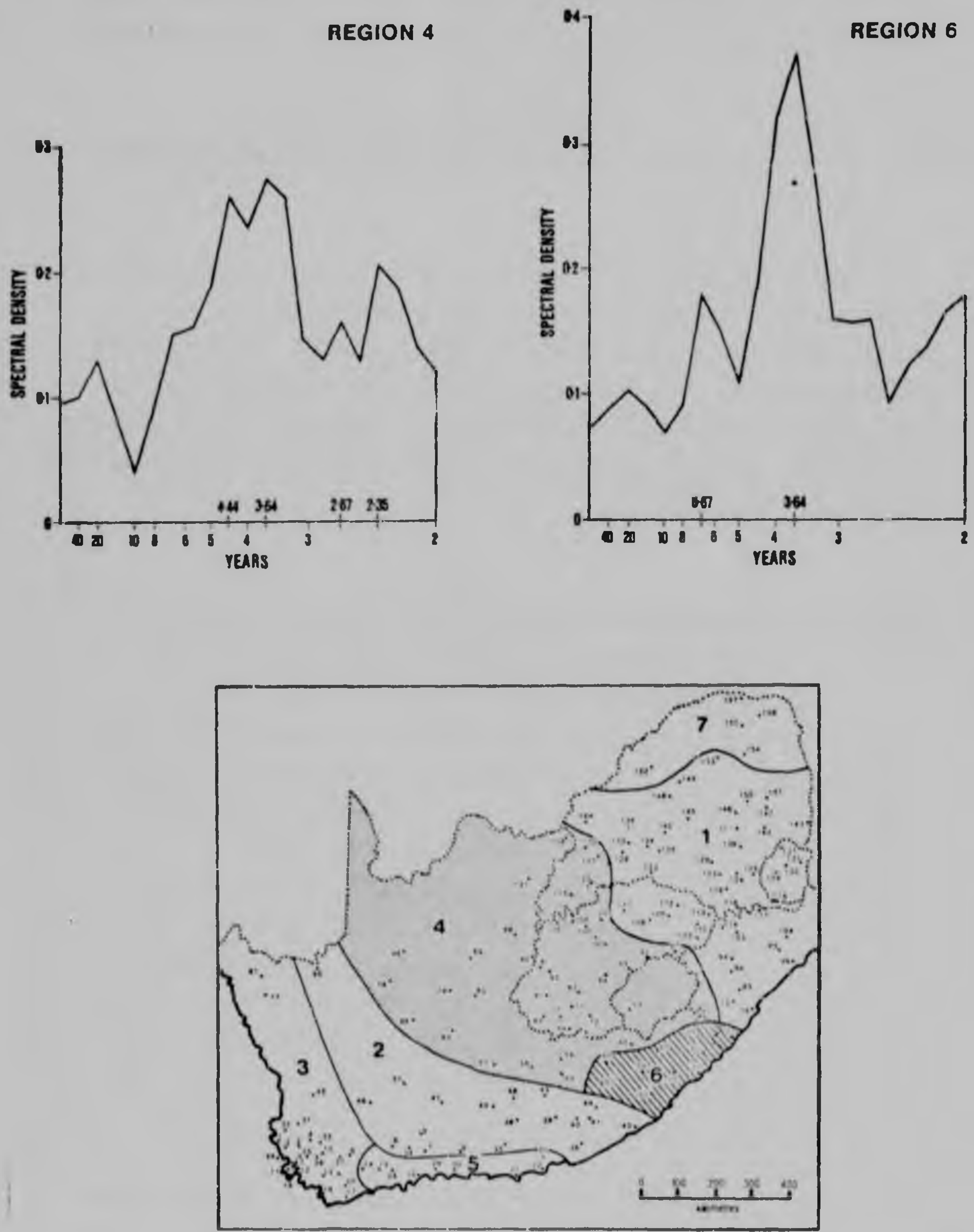


Fig. 5.10: Showing the spectra of the regional temporal series for Regions 4 and 6, together with the regions' geographical locations.

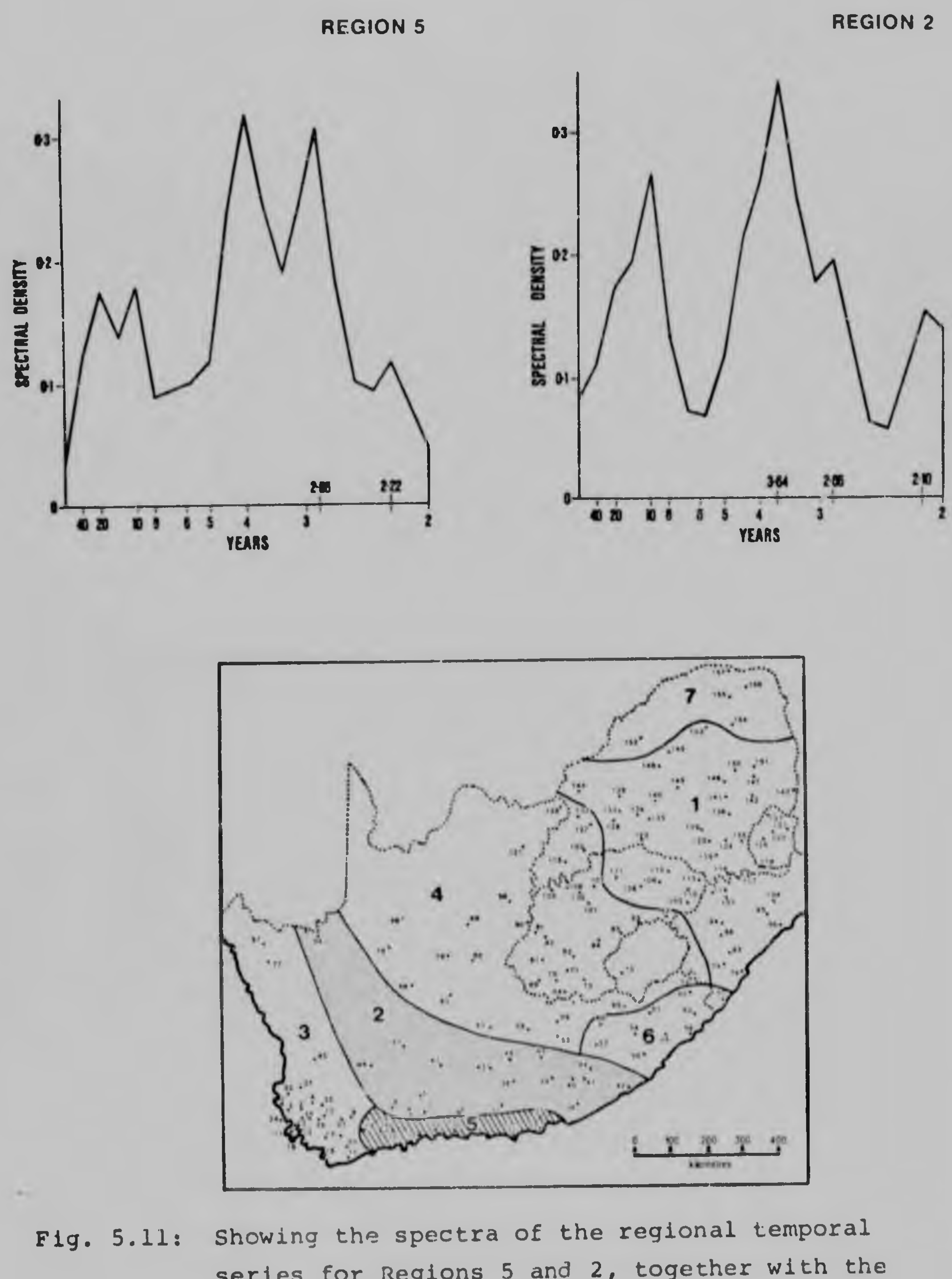
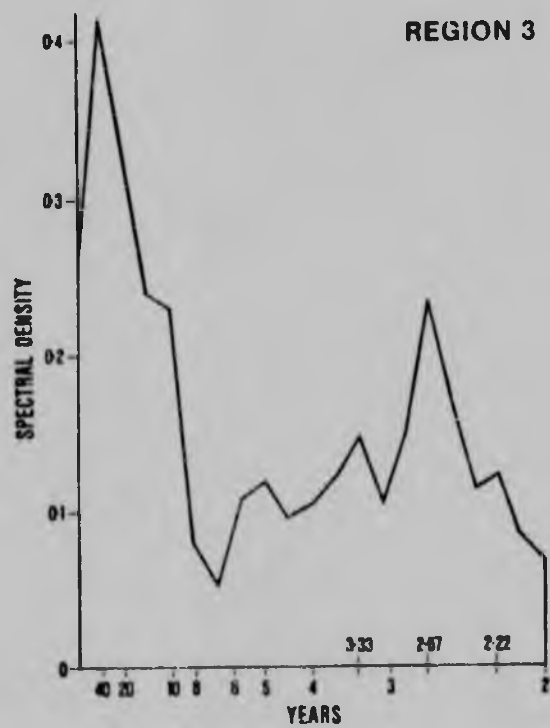


Fig. 5.11: Showing the spectra of the regional temporal series for Regions 5 and 2, together with the regions' geographical locations.



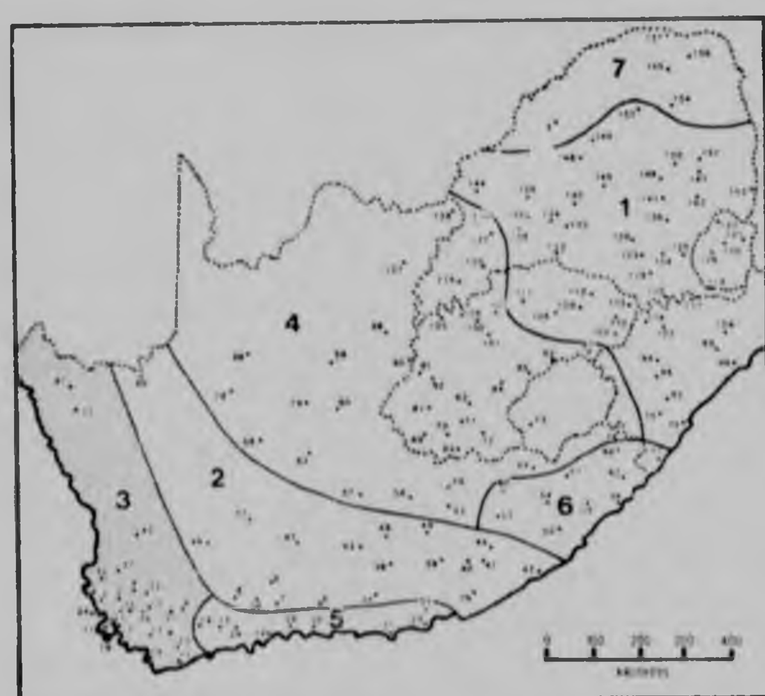


Fig. 5.12: Showing the spectrum of the regional temporal series for Region 3, together with the region's geographical location.

series in which the 20 year oscillation is discernible, but weak, (Fig. 5.10). These Regions, 4 and 6, are in decreasing summer rainfall areas where the predominantly anticyclonic disturbances are weakening. The spectrum of Region 5 is unique, having peaks at both 20 and 10 years, (Fig. 5.11). This southern coastal region borders on Region 2 which forms an area of transition between predominantly anticyclonic activity to the north and cyclonic to the south, (Fig. 5.11). The spectrum for the south-western Cape, Region 3, shows a peak at 40 years, (Fig. 5.12).

At the high frequency end of the spectra an oscillation with periods in the 3-4 year range is ubiquitous. It is weak over the south-western Cape Province, (Region 3) and the far northern area of the country, (Region 7). In the other regions, it is the strongest peak in the spectra, (Figs. 5.9 to 5.12). An oscillation in the 2-3 year waveband is also apparent over most parts of the country. Its wavelength is less stable than that in the 3-4 year interval, and is spatially dependent. Over the south coast, (Region 5) and interior corridor, (Region 2) it appears to peak at 2.86 years. Its wavelength decreases both to the north and south of these regions. It could be argued that peaks at the high frequency end of the spectrum are caused by aliasing. However, apart from the fact that the quasi-biennial oscillation has been noted frequently elsewhere, aliasing presents no serious problem here because sampling is not from instantaneous points on a continuous time scale, (Dzerdzeevski, 1968; Godin, 1972; Trenberth, 1975).

The overall level of significance for the peaks noted in Table 5.6 is less than any of the individual levels quoted. This decrease arises from the use of the multiplicative rule to determine the probability of a number of independent events. In this case the non-adjacent spectral ordinates may be considered independent by virtue of using Parzen's lag window.

Furthermore, the individual spectra are on independent time series. If a uniform distribution of significant ordinates over frequency is taken as a null-hypothesis, the application of the one sample Chi-square test shows that the distribution of noted significant ordinates shown in Table 5.6 has, approximately, a probability of less than 0.02 of having arisen by chance.

The cross-spectral analysis indicates that for the 20 year oscillation significant coherence - greater than 0.34, (Jenkins and Watts, 1968) - exists between Region 1 and Regions 4 and 7. Region 1 leads 4 by about 1.5 years, and is in phase with Region 7. Coherence between Region 1 and Regions 5 and 6 for this wave is insignificant. With regard to the 10 year wave, only Regions 2 and 5 have significant coherence and the two records are in phase. For the 3-4 year wave, Region 1 lags those regions to its south that possess this oscillation, but it leads Region 7 to its north. However, the largest out of phase value is only 0.25 years, with Region 6, and it could be assumed that all regions are virtually in phase for this wave. It is clear from Table 5.6 that the 2-3 year oscillation is widely varying in its value, over the regions. This is borne out by the cross-spectral analysis, which gave rise to no useful information regarding the oscillation in the 2-3 year interval. These results are considered in more detail in a later part of this thesis.

* * * * *

Seven statistically independent rainfall regions have been defined from a principal component analysis. A spectral analysis on the seven regional temporal series shows clearly defined peaks in both the high and low frequency ends of the spectra. An oscillation of 20-years is strongest over the summer rainfall area of the country. A 40-year wave is confined to the south-west Cape coast, Region 3. Region 2 which forms an area of transition between the predominantly anticyclonic circulation to the north and zone of cyclonic disturbances to the south experiences a wave with period 10-years. The southern coastal strip is affected by both the 10- and 20-year waves. A 3-4 year wave is ubiquitous and is strong over the country as a whole, except in Regions 3 and 7. A quasi-biennial oscillation is present over the country and is spatially dependent.

PART III

DISCUSSION OF RESULTS

CHAPTER 6

RAINFALL CHANGES DURING THE PERIOD OF METEOROLOGICAL RECORD

6.1 PREAMBLE

The results of trend analysis on the one hand and evidence for the occurrence of spatially dependent oscillations on the other having been presented in Part II, it is now necessary to submit these findings to critical scrutiny.

Before this is done, however, it will be advantageous to recapitulate briefly. In Part II of the thesis, rainfall series for 157 stations were analysed for trend, both linear and non-linear, using the method of Kendall. Following this a number of approaches were adopted to investigate the oscillatory behaviour of the series. Fourier analysis enabled the variance associated with selected waves to be mapped. Secondly, rainfall deviations smoothed with a 5-termed binomial filter were grouped according to the distribution of cumulative totals for given periods and the distribution of stations showing like temporal patterns of change were mapped. Thirdly, correlogram and spectral analysis were used to determine objectively the behaviour of the raw data series over the time and frequency domains. Relative variance associated with integrated frequency bands was mapped to show the spatial gradients of specific spectral components.

6.2 THE QUESTION OF TREND

The analysis of rainfall records at 157 stations shows that, owing to the lack of independence of the data, no conclusive evidence can be assembled to demonstrate that rainfall has declined generally over the country during the period 1880-1972. Although the number of stations showing negative trend is larger than would be expected by chance (assuming independent data), the impossibility of providing any physical explanation for only a small fraction of stations experiencing this decline militates against accepting the suggestion that South Africa as a whole is undergoing progressive desiccation. In some instances spurious negative trends can be explained by the excessive rainfall in the decade 1880-1890. It would appear that, as the length of record increases so the tendency for the occurrence of trend decreases. In some cases negative trend was introduced only because of the occurrence of the peak in the rainfall series centred on the interval 1880-1890. Only 24 stations out of 157 showed persistent negative trend over the whole period of record. It is inconceivable that the 157 records are or may be treated as independent. Thus statistical hypothesis testing in this case, although mechanically simple to implement, becomes extremely difficult to interpret. Recourse to physical (i.e., climatological) explanation and reasoning is a far more powerful tool in throwing doubt on the hypothesis that South Africa as a whole has undergone progressive desiccation.

6.3 AN ALTERNATIVE TO TREND

Instead of a steady decline, it appears that South African rainfall has oscillated during the period 1880-1972.

Different and independent techniques of analysis suggest the reality of clearly discernible but weak oscillations with periods 16-20, and 10-12 years. The ubiquity of an oscillation of 3-4 years and the distinctive spatial occurrence of a quasi-biennial fluctuation have also been shown.

The quasi-biennial oscillation occurs almost exclusively in that area of the country having a six-monthly seasonal rainfall regime. The 16-20 year fluctuation can be discerned over thousands of square kilometres of the northern, eastern and southern subcontinent having a summer rainfall regime. It is weakest in the south and predominates in the north-east. Though weak, the 10-12 year fluctuation shows a distinctive spatial distribution in the coastal and adjacent inland area of the southern Cape Province experiencing an all-seasons rainfall. That this oscillation is more than simply an harmonic of its 16-20 year counterpart is suggested by the fact that it fails to account for much variance in areas where the 16-20 year fluctuation is most pronounced. Its reality is confirmed by the analysis of periods of rainfall maxima and minima actually observed.

6.4 RECOGNITION OF SOME INHERENT WEAKNESSES IN THE ANALYSIS

Each of the three techniques used to determine the nature of rainfall oscillations is individually open to criticism. The arbitrary choice of a fundamental period for the fitting of cosine curves to the data and the fact that the observed fluctuations are quasi-periodic, rather than truly s[illegible] two factors which make the Fourier analysis a cru[illegible] Furthermore, a basic assumption for this analysis [illegible] the random components must be mutually independe[illegible] [illegible]is condition is open to serious doubt in climatological time series. However, Fourier analysis is unique in that no other type of analysis obtains results in the required form. Bearing in mind the weaknesses, it is still possible to draw meaningful conclusions from such an analysis.

In the case of the sorting of filtered data into periods of maximum and minimum rainfall there is the risk that the smoothing procedure itself may introduce pseudo-periodicities into the data. However, the results of the simulation study show that not only is it difficult to reproduce the correct temporal similarities for many stations from randomly generated data, but it is impossible to reproduce the high degree of observed spatial clustering for the stations that do show appropriate temporal similarities.

The spectral analysis deals with records of short length which makes it impossible to obtain good resolution at the low frequency end of the spectra. Furthermore, the series analysed are not independent. This makes for difficulties when attempting to determine overall statistical significance for the spectral ordinates. However, Granger and Hughes, (1968), do suggest that series of the length used can produce meaningful spectral estimates. In addition, even allowing for lack of independence, the fact that records for many different stations show similar spectra, and that spectral components show a noticeable, cohesive and in-phase spatial organisation, offsets the extent of the doubts one may have concerning the reliability of individual spectra.

Notwithstanding these and other objections, the fact that curve fitting, simple sorting of filtered data into periods of maximum and minimum rainfall, and spectral analysis give independently derived results in substantial agreement, suggests the physical reality rather than the possible statistically spurious nature of the observed fluctuations.

6.5 TEMPERATURE RAINFALL RELATIONSHIP

Yet further support for the physical reality of the 16-20 year rainfall oscillation is to be found in the similar but inverse oscillation in temperature that has been observed over the country, (Table 6.1).

TABLE 6.1 - AMPLITUDES OF, AND PHASE LAGS BETWEEN, 20-YEAR RAINFALL AND TEMPERATURE WAVES AS DETERMINED BY FOURIER AND CROSS-SPECTRAL ANALYSIS RESPECTIVELY, (AFTER KEEN, 1971).

Station No.	Amplitude rainfall	mm, °C) temperature	Phase lag years
157	25	1.7	9
153	35	1.6	9
140	43	2.2	10
136	36	1.2	10
135	44	1.1	11
86	58	1.6	10
67	11	0.4	9
53	11	0.8	10
36	14	1.5	11
29	16	0.9	10
23	21	0.8	9
17	31	1.1	9
49	16	2.7	7
38	17	2.5	15
79	8	2.0	4

At the 80per cent of the stations for which annual temperature data were available the 20-year rainfall and temperature oscillations were out of phase by approximately 160 degrees, (Fig. 6.1).

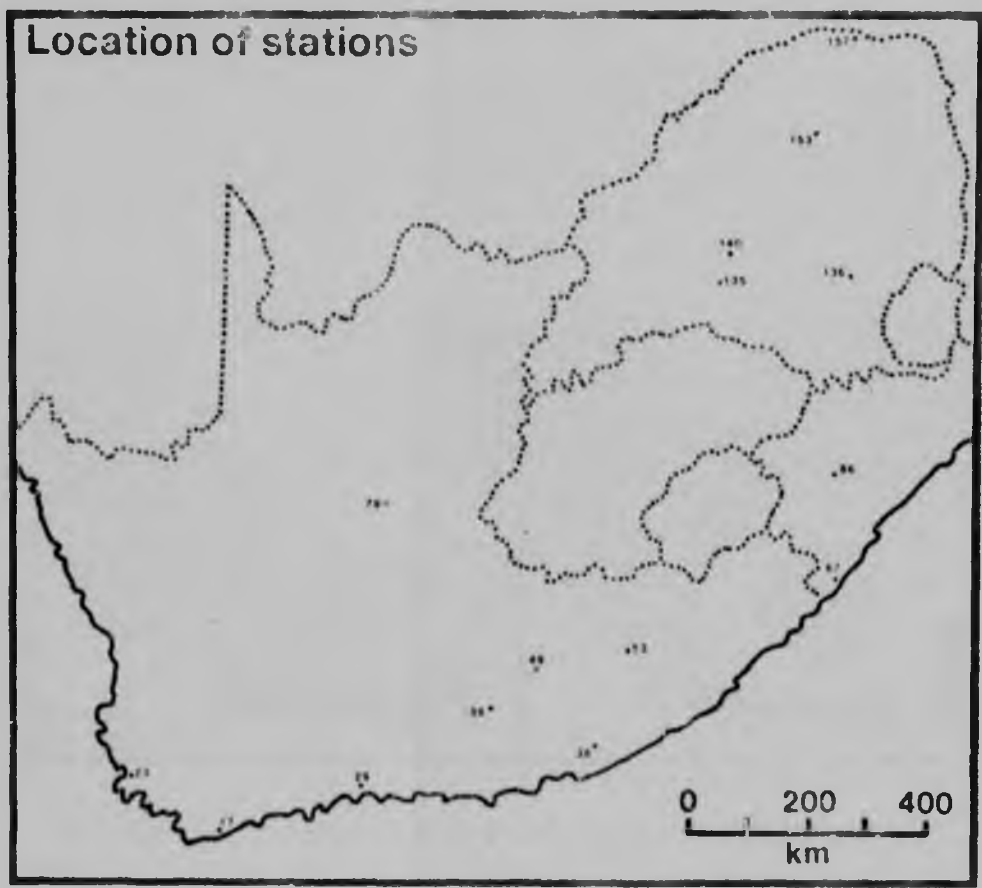


Fig. 6.1: Showing the geographical locations of stations used in the temperature and rainfall analysis.

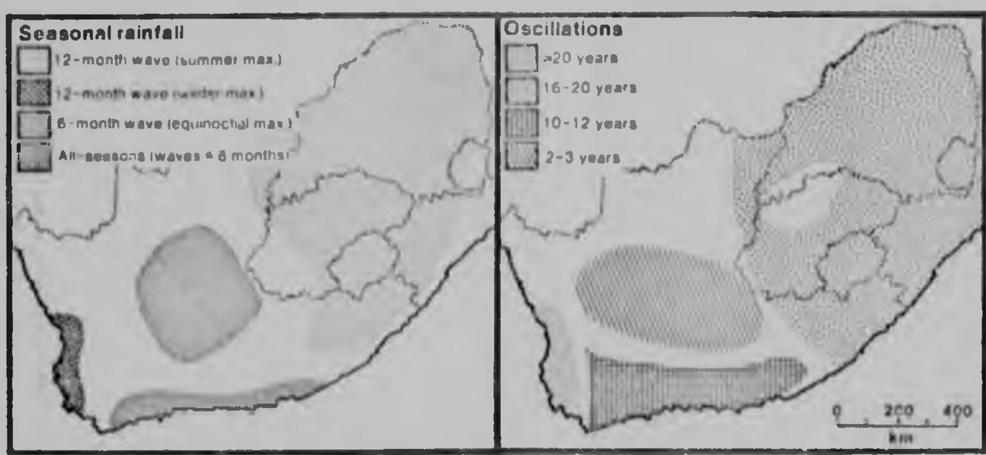


Fig. 6.2: Showing the similarity between the spatial incidence of seasonal rainfall regimes and predominant oscillations in the rainfall records.

The fact that the 16-20 year temperature and rainfall fluctuations vary inversely suggests that the fluctuations in the two parameters reflect similar persistent changes in the general circulation of the atmosphere over the subcontinent, since similar but inverse fluctuations may be expected if the controlling mechanism is the general circulation. Periods of higher temperature and lower rainfall are likely to be brought about by persistent anticyclonic conditions favouring subsidence, clear skies, dry air and optimum insolation.

Similarly, lower temperatures and higher rainfall are likely to be due to persistently greater cyclonic activity, increased convergence and cloud cover, accompanied by decreased insolation, (Schulze, 1965). Such a relationship holds over a time scale of months, when a close correlation exists between seasonal anomalies of precipitation and pressure departures over southern Africa, (Rubin, 1956). An abnormally wet summer over the eastern part of the subcontinent is associated with a general negative departure of pressure, stronger cyclonic activity in the westerlies and a deeper continental trough. By contrast, dry summers coincide with positive anomalies over and to the east of the continent.

The similarity between the distribution of seasonal rainfall regimes, being the manifestation of different rainfall producing systems in different parts of the country, and the distribution of the dominant regional oscillations is striking, (Fig. 6.2). The one-to-one correspondence between seasonal rainfall areas, and regions claimed to have spatially dependent oscillatory rainfall patterns lends further support to the hypothesis that these oscillations have real physical significance and are linked to the general circulation of the atmosphere over Southern Africa.

* * * * *

No conclusive evidence can be found to support the view that South Africa's annual rainfall budget has decreased over the period 1880-1972. Dependence between station records has made statistical inferences unreliable. In some cases spurious trend was introduced as a result of the excessively wet decade between 1880 and 1890.

A number of methods show the presence of spatially dependent fluctuations in the rainfall data. The 16-20 year oscillation is strongest in the summer rainfall area; the 10-12 year one in the coastal and adjacent inland area of the southern Cape Province experiencing an all-season rainfall; a 3-4 year wave is ubiquitous; and a quasi-biennial oscillation has its strongest influence over that area of the country having a six-monthly seasonal rainfall regime. Strong evidence in support of the 16-20 year oscillation is obtained from the fact that temperature and rainfall, over the area where the wave has influence, are 180 degrees out of phase. Evidence supporting the presence of waves in general is given by the similar spatial distribution of seasonal rainfall and these waves.

CHAPTER 7

FURTHER ASPECTS OF RAINFALL CHANGES OVER SOUTH AFRICA

7.1 PREAMBLE

The discussion in the previous chapter reveals that lack of independence between the rainfall series restricts the value of statistical inference. In order to remove this restraint and permit an analysis of rainfall data on a meso-scale, a different approach was adopted. Principal component analysis was used to delimit the country into rainfall regions. In that exercise, rotation of the factors retained from the principal component analysis added greater clarity to the regionalisation of the country in terms of climatic change. The fitting of linear stochastic models to the station records within each group derived from principal component analysis solved the problem of dependence between the rainfall series. Furthermore, it enabled the component series into which rainfall could be decomposed to be analysed individually without being contaminated by the other components.

7.2 THE PARSIMONIOUS REPRESENTATION OF 157 RAINFALL RECORDS

By grouping stations according to the rotated factor with the highest loading common to each station, ten rainfall regions were formed. Of these, seven were retained in the subsequent analyses.

The linear model enabled the regional temporal effect, regional spatial effect, and local effect, to be isolated for each region. Of extreme value is the fact that, by definition, the newly derived series are statistically independent. This permits meaningful conclusions to be made from the results of statistical tests applied to these series.

7.3 A RECONSIDERATION OF TREND IN RAINFALL

By submitting the regional temporal series to a two-way analysis of variance, it was possible to establish that, at the 0.01 level, they were all significantly different from zero; this result also applies to the regional spatial effects.

These regional temporal series were submitted to trend analysis; the results showed that no region was undergoing trend. Of the seven series analysed, the unbalanced distribution of the values of *r(s)* was cause for concern, even though none was significant at the five per cent level. Because the *r(s)* values are independent, the distribution of positive and negative signs follows the binomial distribution under the null-hypothesis. In fact, the probability of obtaining six negative signs from a total of seven is 0.11 by chance. Hence it can be concluded with 95 per cent confidence that South Africa is not undergoing progressive desiccation.

This result makes the one regarding trend, given in Chapter 4, even more puzzling; this is so even when taking account of the dependence problem. However, the residuals in the linear model represent local effects that are unexplained by the general meso-effects of the regions. Examining these residuals for trend showed that some possessed significant trend at the five per cent level. Of the 157 residual series examined, a total of 34 showed significant trend. The distribution was 16 stations with positive and 18 stations with negative trends. Applying the one-sample Chi-square test resulted in this distribution of signs having occurred by chance with a probability of less than 0.001. The spatial distribution of stations with trend on a micro-scale shows no spatial organisation, stations with positively trending residuals often being adjacent to those with negative trending values. It may therefore be concluded with a high degree of confidence that some stations are affected by trend on a local scale, but that there is no bias toward trend in either direction.

7.4 THE OSCILLATORY BEHAVIOUR OF THE REGIONAL TEMPORAL SERIES

In submitting the regional temporal series to spectral analysis the advantages of the linear two-way unreplicated model become apparent. The seven spectra are independent by the model, consequently any peak common to some or all of these spectra gains added support for its reality. One such peak is that in the 16-20 year waveband. It appears in the spectra of Regions 1,4,5,6 and 7. This peak is statistically significant in Regions 1 and 7, and weakens in spectra for regions away from the summer rainfall area of the country. In terms of coherence this oscillation of 20 years is common to Regions 1,4 and 7. Lack of significant coherence between these regions and Regions 5 and 6 suggests that in the southern part of the country the 20 year wave is an harmonic of a wave of longer period implied by the peak at the low frequency end of the spectrum for Region 3.

The 10-12 year oscillation has statistical significance only in the spectrum of Region 2. It does, however, appear in the spectrum of the adjacent coastal area, Region 5. Since this peak is absent in the spectra of regions where the 20-year one is strong, it may be concluded that it is not an harmonic of the 20-year wave present over the summer rainfall area of the country. That the 10-year wave is of common origin to both Regions 2 and 5 is indicated by the highly significant coherence value, (0.79) that is present between these two regions for this wave. Although the southern part of the subcontinent experiences cyclonic disturbances which weaken in a northerly direction, the 10-year wave would not appear to be associated with these disturbances. This follows from the fact that its area of greatest influence is over Region 2, and weakens toward the south coast and the accompanying cyclonic activity. The existence of this wave may be due to a beat effect between the 20-year oscillation to the north and the one of longer wavelength to the south-west. The ubiquity of an oscillation with wavelength in the 3-4 year interval is illustrated by the presence of a relevant peak in the spectra for the regions as a whole. Taking Region 1 as an area of reference, those regions to the south lead this region with respect to the phase of this wave. On the other hand, the one region to the north of Region 1 lags behind it. However, the phase differences are small in all cases. The largest phase difference for this wave is between Region 1 and Region 6, where it is of the order of 0.25 years. This 3-4 year oscillation would therefore appear to exist in its own right, originating from some process that is independent of the mechanisms responsible for the various oscillations of longer wavelength that are in evidence over the country as a whole.

The presence of the quasi-biennial oscillation is widespread. This wave is, however, more variable in its wavelength than is the 3-4 year oscillation. The coherence for this wave between the regions is generally low. This leads to the hypothesis that it is of purely local importance. However, the fact that it has been observed over other wide areas of the world mitigates against this theory.

7.5 PRINCIPAL COMPONENTS AND RAINFALL STATION GROUPINGS

Principal component analysis and rotation of the eigenvectors accounting for significant variance in the original data matrix provided a means of delimiting the country into homogeneous rainfall regions. Fitting a linear two-way unreplicated model to each of the regions so formed enabled the regional temporal effects to be isolated for further analysis. From an inspection of the spectra of these temporal series, it is clear that the principal component analysis has pulled out, or grouped, those stations whose records have common oscillatory patterns.

Factor one has pulled out those stations whose records are most strongly affected by the quasi-20 year oscillation. Those stations experiencing a quasi-10 year oscillation have been grouped under factor 2, whilst factor 3 has formed a rainfall region from the stations whose annual totals fluctuate with a period of the order 40 years.

* * * * *

Using principal component analysis to form rainfall regions and fitting linear models to the sets of data contained in each region has resulted in a clear picture being obtained regarding rainfall behaviour on a regional basis. The linear model enables rainfall amounts to be divided into three components, regional temporal, regional spatial, and local effects. These can be examined separately, free of contamination from each other. Trend is shown to be absent on a regional scale, but present on a micro one. Hence it can be concluded that South Africa's rainfall receipts have not decreased over the period 1910 to 1972. Furthermore, by the linear model, tests for trend have been carried out on independent series, thus making the conclusion reliable. Evidence for the existence of fluctuations in rainfall within definite frequency intervals is provided by the the spectra of the regional temporal series. The summer rainfall area has a 16-20 year wave; all year seasonal area, a 10-year wave; the all year southern coastal strip has both the 10 and 20-year waves, though weakly so; and the area with a mediterranean climate has a 40-year wave. At the high frequency end of the spectrum, the 3-4 year wave is ubiquitous; the quasi-biennial wave is spatially dependent.

PART IV

STOCHASTIC MODELS AND RAINFALL PREDICTION

OVER SOUTH AFRICA

CHAPTER 8

SOME APPROACHES TO SHORT AND LONG TERM FORECASTING OF RAINFALL OVER SOUTH AFRICA

8.1 INTRODUCTION

Both series and spectra have shown that the regional temporal rainfall series are oscillatory. This fact suggests that rainfall over South Africa may lend itself to prediction by models applicable to time series data of this type. The investigation dealing with forecasting models is an exploratory one. Two regions have been chosen to illustrate the results of this section of the work. The summer rainfall region of the north-eastern part of the country (Region 1), experiencing a quasi- 20 year oscillation, and the southern region with a quasi- 10 year oscillation (Region 2).

Despite the considerable attention that has been given to climatic change in the literature, few attempts have been made to model such changes mathematically (Schneider and Dickinson, 1974). The difficulties attendant on such modellings are formidable and for this reason, and since model building constitutes a major topic in itself, no more than preliminary statistical modelling will be explored in this chapter. Four approaches suggest themselves; autoregression, multiple linear regression, exponential smoothing, and trigonometric curve fitting. Between them, these models cover short and long term forecasts. Before considering each of these in turn, it is necessary to consider certain theoretical aspects of forecasting models in general.

8.2 THE FORECASTING MODELS

AUTOREGRESSION AND MULTIPLE LINEAR REGRESSION MODELS

Autoregression models have been fitted to the regional temporal series (Pagano, 1972). The estimation of the parameters in these models can be done by either solving the Yule-Walker equations (Kendall, 1973), or by adopting multiple linear regression techniques as advocated by Kendall and Stuart (1966). In this thesis the multiple linear regression techniques have been used.

The final order of a model is based on the prediction sum of squares as the fitting criterion (Allen, 1971a, b, 1972, 1974). The classical multiple linear regression model may be defined as

$$\underline{Y} = \underline{X}\underline{\beta}+\underline{\varepsilon} \tag{8.1}$$

where $\underline{Y}$ is the $(n \times 1)$ column vector on the dependent variable; $\underline{X}$ is the $(n \times k)$ matrix on the independent variable and of full rank; $\underline{\beta}$ is the unknown $(k \times 1)$ column vector of regression parameters to be estimated; and ε is the $N(0, I\sigma^2)$ column vector of random variables.

The prediction sum of squares relating to this model is defined as

$$P(l) = \sum_{i=1}^{n} (Y_i - \hat{Y}_{il})^2 \tag{8.2}$$

where Y_i is the *i-th* value of the dependent variable; and $\hat{Y}_{il}$ is the least squares estimator of $E(Y_i)$ using the *l-th* subset of predictor variables from the *k* possible, for $o \leq l \leq k$.

Each observation is therefore predicted from an equation whose parameters have been estimated using the remaining $(n - 1)$ observations on Y_i. This method simulates prediction as the analysis progresses, and the criterion can be shown to be a weighted sum of squares of residuals. The weights are functions of the variances of the independent variables. The residual sum of squares, which is the classical criterion in regression analysis, is but a simple sum. Prediction sum of squares will have a relatively low value when associated with subsets that have small squared residuals related to independent variables with large variances. Independent variable selection by the prediction sum of squares usually results in a more parsimonious subset, than that given by conventional means.

A reduction in the computation time of repeatedly calculating the prediction sum of squares is obtainable from a re-arrangement of the expressions for ordinary estimators in the multiple linear regression model. The presence of vectors in the algebra that follows is clear; vector notation is therefore dropped.

Initially

$$Y = X\beta + \varepsilon \tag{8.3}$$

and

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{8.4}$$

Let $\hat{\beta}_{(i)}$ be the column vector of regression parameters estimated on $(n-1)$ observations; that is, the i-th observation removed where $1 \leq i \leq n$ then

$$\hat{\beta}_{(i)} = (X'_{(i)} X_{(i)})^{-1} X'_{(i)} Y_{(i)}$$

$$= (X'X - x'_{(i)} x_{(i)})^{-1} (X'Y - x'_{(i)} y_{(i)}) \qquad (8.5)$$

where $x_{(i)}$, $y_{(i)}$ are the observations deleted from the calculation, which in future will be denoted by x, and y respectively.

Now, using the matrix identity

$$(A - uv')^{-1} = A^{-1} + \frac{A^{-1} uv' A^{-1}}{(1 - v' A^{-1} u)} \qquad (8.6)$$

$$\hat{\beta}_{(i)} = \left[(X'X)^{-1} + \frac{(X'X)^{-1} x'x \, (X'X)^{-1}}{(1 - x(X'X)^{-1} x')} \right] (X'Y - x'y)$$

$$= (X'X)^{-1} X'Y + \frac{(X'X)^{-1} x'x (X'X)^{-1} X'Y}{(1 - x(X'X)^{-1} x')}$$

$$- \frac{(X'X)^{-1} x'x (X'X)^{-1} x'y}{(1 - x(X'X)^{-1} x')} \qquad (8.7)$$

now, let

$$g = (X'X)^{-1} x' \qquad (8.8)$$

then,

$$\hat{\beta}_{(i)} = \hat{\beta} + \frac{gx\hat{\beta}}{(1-x(X'X)^{-1}x')} - gy - \frac{gxgy}{(1-x(X'X)^{-1}x')}$$

$$= \hat{\beta} + \frac{gx\hat{\beta}}{1-xg} - gy - \frac{gy'x'y}{(1-xg)}$$

$$= \left[I + \frac{gx}{(1-xg)} \right] \hat{\beta} - g \left[y + \frac{g'x'y}{(1-xg)} \right]$$

$$\hat{\beta}_{(i)} = \left[I + \frac{gx}{(1-xg)} \right] \hat{\beta} - \frac{gy}{(1-xg)} \qquad (8.9)$$

Equation 8.9 facilitates evaluation of the $\hat{\beta}_{(i)}$'s for each of the l subsets $o \leq l \leq k$ with a minimum of computing time.

For relatively small samples, bias exists in the estimates of the regression coefficients; this bias is due to autocorrelation in the residuals. Quenouille (1956) introduced a statistic, now known as the generalised Jacknife statistic, for the express purpose of reducing bias in *any* estimator. The statistic has been used in a general method of forming confidence intervals where normal, and asymptopic theory did not apply (Tukey, 1958). Gray and Schucany (1972) have investigated Quenouille's earlier work and explored the possibility of its extension and adaptability to wider fields of estimation. The usefulness of this statistic has been explored further by Bissel and Ferguson (1975). Quenouille defined his statistic as,

$$J(\hat{\theta}) = n\hat{\theta}_n - (n-1)\,\bar{\hat{\theta}}_{n-1} \qquad (8.10)$$

where $\hat{\theta}_n$ is the original biased estimator based on n observations in a sample of size n, and $\bar{\hat{\theta}}_{n-1}$ is the mean of the $\hat{\theta}$'s obtained by using the n possible subsets of $(n-1)$ observations.

Equation 8.9 and the regression with the minimum $P(l)$, $o \leq l \leq k$, are used to obtain the Jacknife estimates from

$$J(\hat{\beta})_l = n\hat{\beta}_{n,l} - (n-1)\,\bar{\hat{\beta}}_{(n-1),l} \qquad (8.11)$$

where $J(\hat{\beta})_l$ is the Jacknife estimate of a regression coefficient in a regression of order l; Y_i is the sample of dependent variables $1 \leq i \leq n$; $\hat{\beta}_{n,l}$ is the biased estimate of a regression coefficient based on the n observations in a regression of order l; and $\bar{\hat{\beta}}_{(n-1),l}$ is the mean of the $\hat{\beta}$'s obtained by using the n possible subsets of $(n-1)$ observations in a regression of order l.

An estimate of the standard error of the residuals is given by

$$S.E. = \left[\frac{P(l)}{n-l}\right]^{\frac{1}{2}} \qquad (8.12)$$

where l is the order of the regression as determined from 8.2. Uncertainty regarding the distributions being dealt with when using autoregressive models suggest that tolerance regions, rather than confidence intervals, are more appropriate. These may be estimated from

$$\hat{Y} \pm 1.96\,(S.E.) \qquad (8.13)$$

for 95 per cent regions where $\hat{Y}$ is obtained from

$$\hat{\underline{Y}} = \underline{XJ}(\hat{\beta})_l \tag{8.14}$$

Confidence intervals may be set up for the regression coefficients β_j, $1 \leq j \leq l$. Initially, as each Y_i, for $1 \leq i \leq n$, is deleted from a sample

$$J_i(\hat{\beta}) = n\hat{\beta} - (n-1)\hat{\beta}_i \qquad i = 1, 2, \ldots, n \tag{8.15}$$

where the estimate $\hat{\beta}_i$ of $\hat{\beta}$ is obtained from the subsample when the *i-th* observation Y_i has been deleted. These pseudovalues $J_i(\hat{\beta})$ can be treated as independent identically distributed random variables (Tukey, 1958). Hence the expression

$$\frac{(J(\hat{\beta}) - \beta)(n(n-1))^{\frac{1}{2}}}{\sqrt{\sum_{i=1}^{n} (J_i(\hat{\beta}) - J(\hat{\beta}))^2}} \tag{8.16}$$

which is a vector in a regression situation, can be assumed to be asymptotically distributed as a normal random variable with zero mean, and unit variance (Gray and Schucany, 1972). The Jacknife $J(\hat{\beta})_l$, can be used to form an approximate confidence interval for any β.

THE EXPONENTIAL MODEL

In a model of the type (Kendall and Stuart, 1966)

$$u(t+1) = au(t) + a(1-a)u(t-1) + a(1-a)^2 u(t-2) + \ldots,$$
$$+ a(1-a)^k u(t-k) + \varepsilon(t+1) \tag{8.17}$$

the terms contribute a diminishing influence on $u(t+1)$ as t decreases backwards in time providing $|1-a| < 1$. In 8.17 $u(t+1)$ is the one unit ahead value of the time series in which $u(t)$ is the latest observation or known value; a the smoothing constant; k, $(o \leq k \leq n)$, is the number of historical terms used to forecast $u(t+1)$; and $\varepsilon(t+1)$ is the error variable.

The importance of the most recent observation when forecasting $u(t+1)$ is determined by the value of a. Estimating the optimal a is not straightforward. With computer assistance, a is varied for different values of k and the minimum mean sum of squares of the errors of the forecasts is used as the criterion for fixing both a and k. As the section of work on forecasting progressed, the final orders of the models chosen suggested that the extra computation time involved in using the prediction sum of squares as a criterion was probably not justified here. The exponential model can be more flexible than the autoregressive type, and for ill-behaved time series is often more appropriate, because the parameter is easily adjusted in the event of the generating process undergoing change (Barnard, 1959; Brown, 1959, 1963; Box and Jenkins, 1970; Cox, 1961; Lowe, 1970; Nelson, 1973; Salzman, 1968; Ward, 1963; Winters, 1960, 1969).

The basic exponential model can be extended to double and triple types (Brown, 1963). For the triple model,

$$u(t+1) = b(t) + c(t) + 0.5d(t) + \varepsilon(t+1) \qquad (8.18)$$

Initial estimates, which are henceforth updated, of the coefficients $b(t)$, $c(t)$ and $d(t)$ are obtained from

$$\hat{b}(t) = u(1) - \hat{c}(t) - 0.5\ \hat{d}(t) \tag{8.19}$$

$$\hat{c}(t) = u(2) - u(1) - 1.5\ \hat{d}(t) \tag{8.20}$$

$$\hat{d}(t) = u(1) - 2u(2) + u(3) \tag{8.21}$$

As new observations become available the updating functions are

$$\hat{b}(t) = u(t) + (1-a)^3 (\hat{u}(t) - u(t)) \tag{8.22}$$

$$\hat{c}(t) = \hat{b}(t-1) + \hat{c}(t-1) - 1.5a^2 (2-a) (\hat{u}(t) - u(t)) \tag{8.23}$$

$$\hat{d}(t) = \hat{d}(t-1) + a^3(\hat{u}(t) - u(t)) \tag{8.24}$$

These updated coefficients are used in

$$\hat{u}(t+1) = \hat{b}(t) + \hat{c}(t) + 0.5\ \hat{d}(t) \tag{8.25}$$

to give one step ahead forecasts.

In theory, this model can be used to provide forecasts even when no historical data are available. Usually in estimating the coefficients $b(t)$, $c(t)$, $d(t)$ either of two methods are used. The complete set of available data can be used in the estimation, whilst incrementing a to obtain the minimum mean sum of squares of the forecasting errors. This has the disadvantage that initially the coefficients may be particularly biased, because they are based on the use of few terms. The result of this is reflected in the mean sum of squares of the residuals, and it becomes difficult to assess how well the model could fit the data.

Alternatively, the data set may be divided into three subsets (Salzman, 1968). The first subset is used to provide initial estimates for the coefficients; in this way their values are given the opportunity to stabilise. The second subset is used to finalise the estimation and determine a. Finally, the goodness of fit, as indicated by the mean sum of squares of the residuals can be determined from the third subset. Both methods have been used in the present work.

THE TRIGONOMETRIC MODEL

Regression models of the form

$$Y_i = D_o + \sum_{j=1}^{k} \left[D_i \cos \frac{2\Pi i}{\lambda_j} + E_i \sin \frac{2\Pi i}{\lambda_j} \right] + \varepsilon_i \qquad (8.26)$$

$$\text{for } o \leq i \leq n-1, \ 1 \leq j \leq k$$

are fitted by ordinary least squares to the regional temporal series, where D_i, E_i are the regression coefficients; n is the number of observations on Y_i; k is determined from the appropriate spectrum for each region; and the wavelengths, λ_j, are chosen so as to minimise the sum of the squares of the residuals. Initial estimates of the wavelengths are obtained from the spectrum for a particular region. Thereafter they are incremented positively and negatively about the starting value in a loop within the fitting computer programme. Unlike conventional harmonic analysis, which is inappropriate for non-deterministic data, the wavelengths are not necessarily integer variables (Granger and Hughes, 1971; Kupper, 1972). Furthermore, they are not constrained to be multiples of the sample size used in the computations.

8.3 APPLICATION OF THE FORECASTING MODELS

AUTOREGRESSION ANALYSIS ON THE REGIONAL TEMPORAL SERIES

Fitting autoregressive models to the regional temporal series for Regions 1 and 2 result in constant models (zero order) for both cases (Table 8.1). The model for each region having the best predictive value, i.e., the one with the minimum value of $P(l)$, is underlined in the table.

The Jacknife estimates of the coefficients in these models are -4.83 and -4.74 for Regions 1 and 2 respectively. For comparison, the ordinary least squares estimates are -4.87 and -4.75 for Regions 1 and 2 respectively. Since the difference between the two estimates is obviously not significant, it follows that the number of data points taken in the estimation was sufficient to reduce the problem of bias which can be one of major difficulty when using lagged variables in regression models (Johnston, 1972). In addition, it follows that the use of the Jacknife estimates has not significantly refined the model.

The sample estimate of the variance (i.e., observed variance) for the regional temporal series in Region 1 is 16329 and in Region 2 is 5572. For the best autoregressive model, that of zero order, there is little difference between these observed variances and the prediction sum of squares obtained from the model. Therefore, because the model accounts for so little of the observed variance of the regional temporal series, they must be rejected as being of no climatological significance or usefulness in the present study.

TABLE 8.1 - MEAN PREDICTION SUM OF SQUARES, $P(l)$, FOR REGIONS 1 AND 2 USING 53 DATA POINTS (10 VALUES LOST IN THE ESTIMATION PROCESS) FOR AUTOREGRESSIONS OF INCREASING ORDER

Region	Order of autoregression										
	0	1	2	3	4	5	6	7	8	9	10
1	<u>15173</u>	15590	16133	16146	16094	16492	16214	16417	15346	15414	15956
2	<u>5171</u>	5306	5457	5676	5911	5772	5810	5684	5717	5845	6102

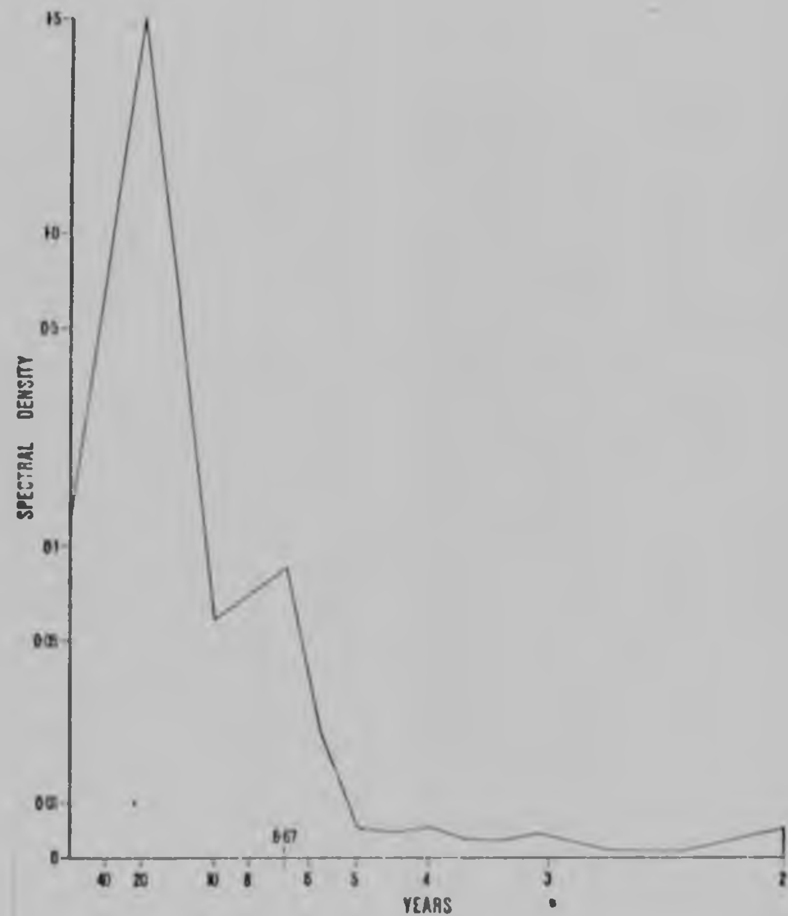


Fig. 8.1: Spectrum of the 20-year (double) sunspot series.

MULTIPLE REGRESSION USING SUNSPOT NUMBERS AS PREDICTOR VARIABLES

The literature reflects a longstanding interest in possible relationships between climatological parameters and solar activity, e.g. Koppen (1873). However, as Lamb (1972) points out, the work of Walker (1915) and Brooks (1923, 1934) did much to discourage further research in this field. In the case of Walker's world-wide survey and Brooks' study of the variation in level of Lake Victoria, insufficient thought was given to the formulation of the conclusions that were drawn. In Lamb's opinion, the general conclusion that no relationship existed between solar activity and weather was probably unfounded.

Further interest shown in this area of study in the 1950s by over enthusiastic amateurs working in isolation brought more discredit to the subject. Of the many reviews of work carried out in this field of research, those of particular value can be found in: Baur (1949, 1956, 1958), Belinski (1957), Julien et al (1957), Lamb (1972), Mironovitch (1960), Mitchell (1965), Mustel (1967), and Roberts (1963).

Renewed interest in this field of research has been shown by King (1973, 1974) and Willett (1964, 1965, 1968, 1974). The Willett (1968) paper is one of the few that attempts to describe a possible physical foundation for solar/weather relationships.

The oscillations present in both the individual rainfall series and regional temporal series of order 20 and 10 years show a striking similarity with solar activity as measured by sunspot numbers. For the time interval 1910-1972 the spectra of the double and single cycles peak at 20, and 10 years respectively (Fig. 8.1,8.2).

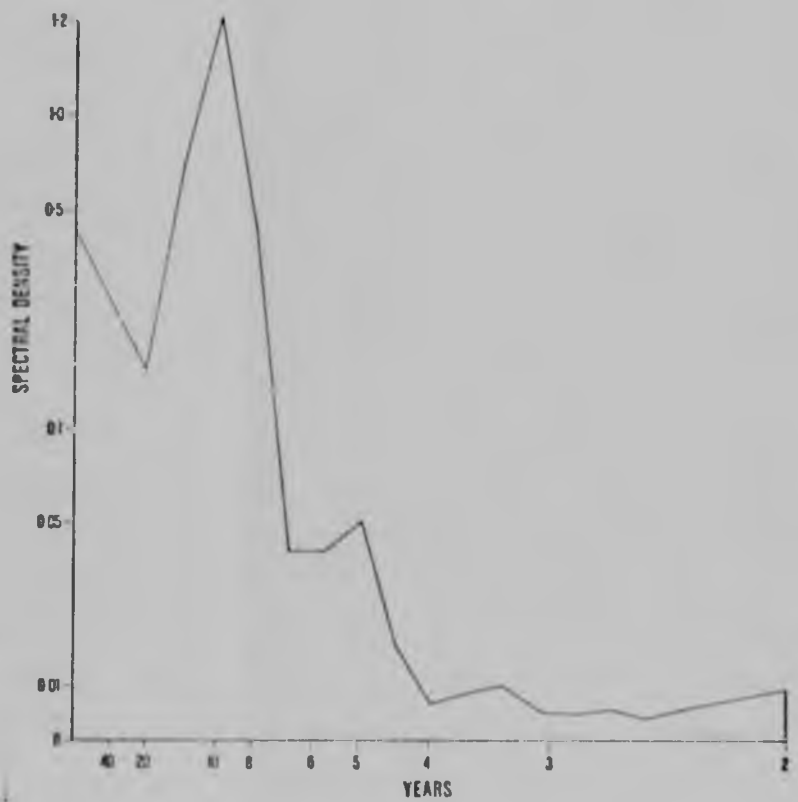


Fig. 8.2: Spectrum of the 10-year (single) sunspot series.

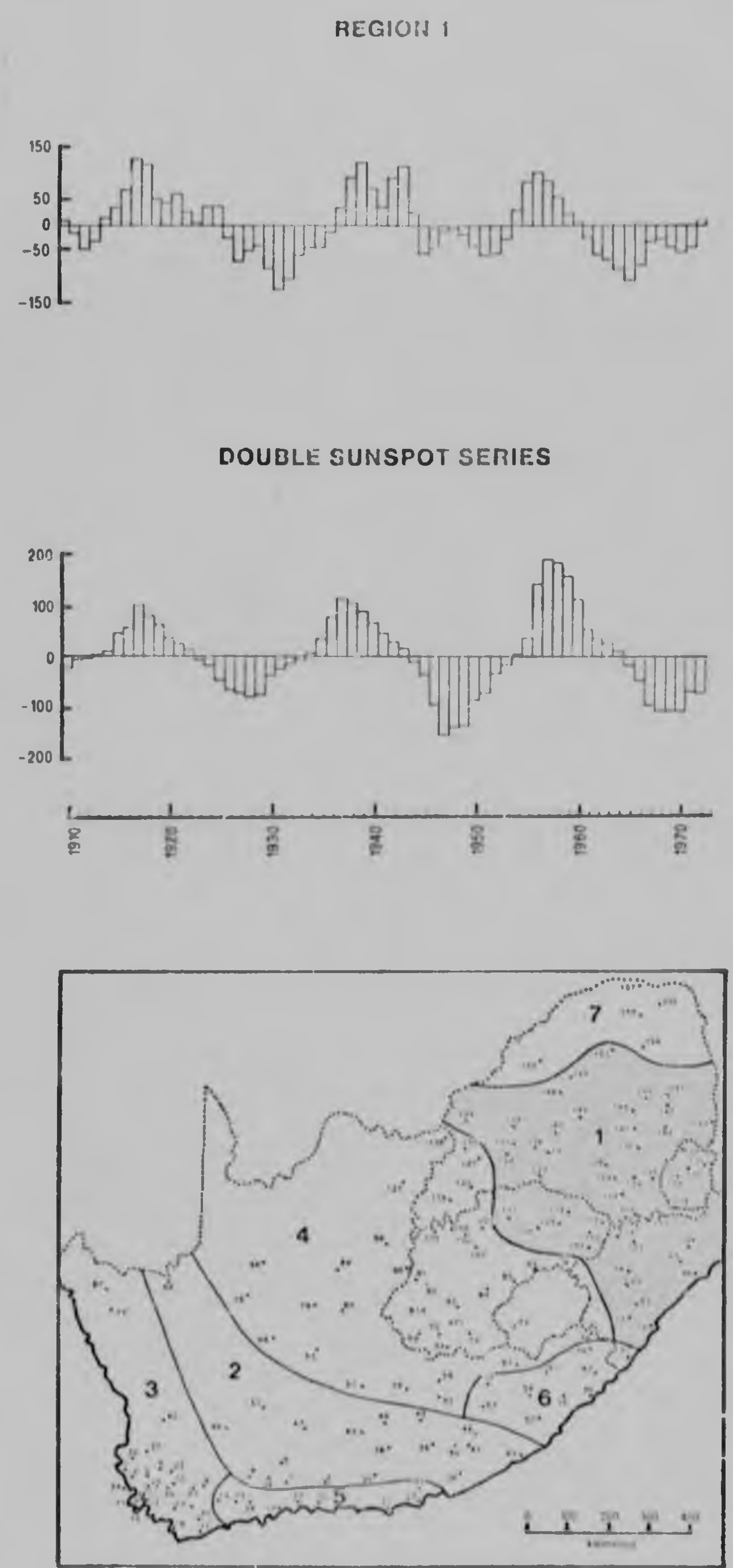


Fig. 8.3: Showing the in-phase similarity between the temporal series for Region 1 and the double sunspot series.

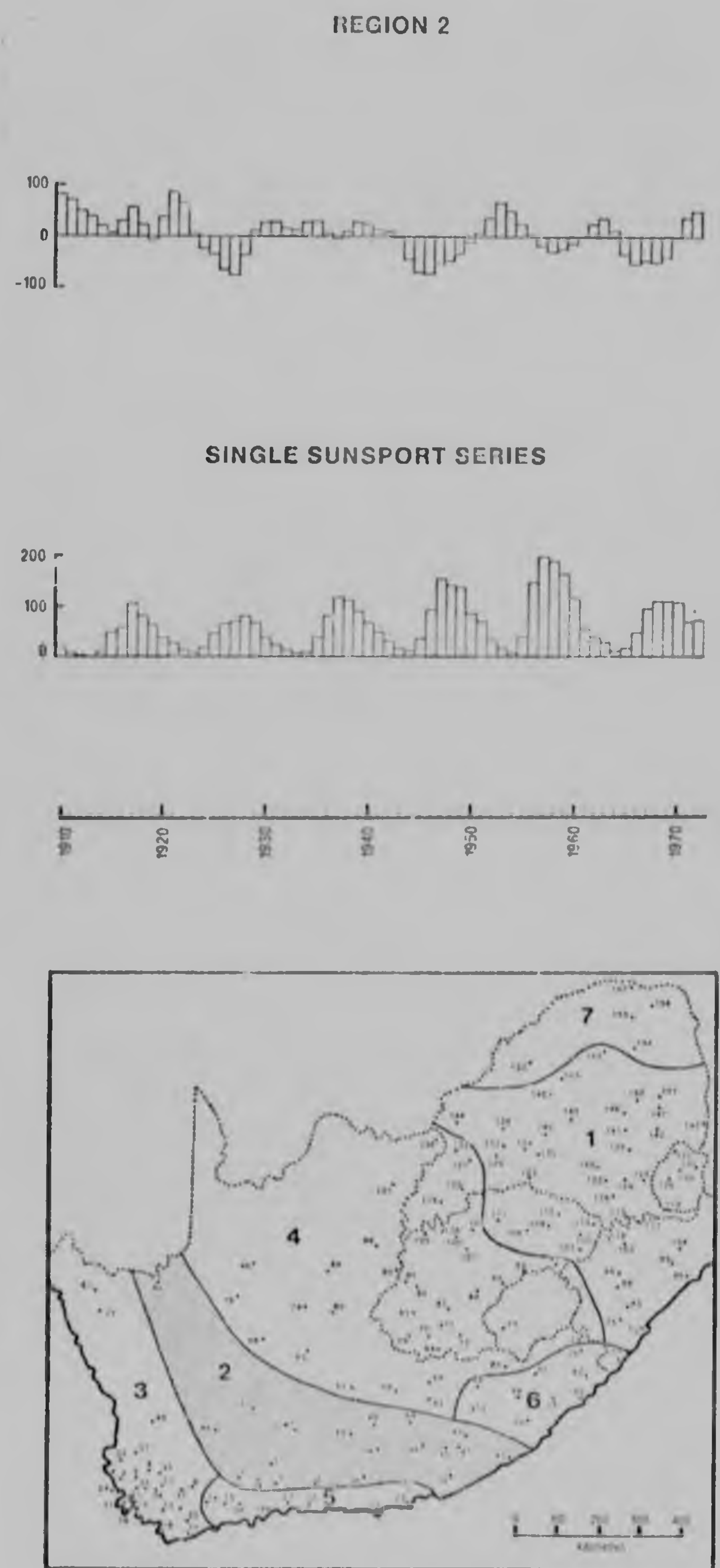


Fig. 8.4: Showing the anti-phase similarity between the temporal series for Region 2 and the single sunspot series.

The summer rainfall area of the country, Region 1, has a regional temporal series whose variation resembles the double sunspot series closely, (Fig.8.3). Cross-spectral analysis reveals a significant coherence of 0.53 between the two series at a wavelength of 20 years. Solar variation leads rainfall by about 0.5 years, and therefore the two series can be assumed to be in phase.

The all seasons rainfall area, Region 2, has a strong similarity between its regional temporal series and the single sunspot series, (Fig. 8.4). The coherence between these two series at the wavelength of the single solar cycle is 0.85, with a phase difference of about 4 years. These two series are therefore not far from being in perfect anti-phase.

Comparison of the coherence of 0.54 between the regional temporal series for Region 1 and the double sunspot cycle, and the coherence of 0.85 between the regional temporal series for Region 2 and the single sunspot cycle is, at first glance, confusing when considered in conjunction with the graphs given in Fig.8.3. The apparent contradiction lies in the fact that the regional temporal series for Region 1 tends to be bi-modal.*

Because the series are relatively short, the coherence values can be taken as no more than a guide to the frequencies that are common between two series (Granger and Hughes, 1968). In these circumstances the coherencies tend to be biased heavily toward 0.5. Confidence limits for the phase differences are useless and therefore not given. This fact arises from the necessity of having to use the sample coherence

---

* *Further research may establish the combined effect of the double solar cycle and the lunar declination cycle of 18.6 years, the latter being suggested by Nevill (1908).*

TABLE 8.2 - REGRESSION MODELS OF DIFFERENT ORDERS AND THEIR MEAN PREDICTION SUM OF SQUARES, $P(l)$, FOR REGIONS 1 AND 2 BASED ON THE SINGLE SUNSPOT CYCLE AS PREDICTOR

Region	Order of regression										
	0	1	2	3	4	5	6	7	8	9	10
1	16581	<u>16254</u>	16555	17006	17357	17705	17912	18310	18379	18688	17295
2	5659	<u>5392</u>	5547	5712	5694	5839	5969	6212	6214	5908	5902

TABLE 8.3 - REGRESSION MODELS OF DIFFERENT ORDERS AND THEIR MEAN PREDICTION SUM OF SQUARES, $P(l)$, FOR REGIONS 1 AND 2 BASED ON THE DOUBLE SUNSPOT CYCLE AS PREDICTOR

Region	Order of regression										
	0	1	2	3	4	5	6	7	8	9	10
1	16581	<u>15543</u>	15924	16292	16491	16704	17086	17339	17575	17511	17970
2	5659	5590	<u>5371</u>	5425	5638	5822	6011	5982	6089	6291	6243

TABLE 8.4 - COMPARISON OF JACKNIFE AND ORDINARY ESTIMATES FOR REGRESSION COEFFICIENTS WHEN TEMPORAL SERIES FOR REGIONS 1 AND 2 ARE REGRESSED ON THE SINGLE AND DOUBLE SUNSPOT SERIES

Region	Jacknife estimates		Ordinary estimates	
	single cycle	double cycle	single cycle	double cycle
1	-31.14	-1.47	-31.16	-1.46
	0.52	0.48	0.52	0.43
2	24.29	-0.19	24.27	-0.20
	-0.40	0.77	-0.40	0.76
		0.62		0.62

estimate as the true coherence when deriving confidence intervals for the phase differences.

Fitting multiple linear regression models using sunspot numbers as a lagged predictor variable on the regional temporal series gives rise to the models and their prediction sum of squares values, Tables 8.2 and 8.3. Table 8.2 refers to the single sunspot cycle as predictor variable whereas Table 8.3 refers to the double cycle as the predictor.

In the case when single sunspot numbers are taken as the predictor variable, the order of regression having the minimum prediction sum of squares is unity for both Region 1 and 2. When the double sunspot numbers are taken as the predictor variable Region 1 is best represented by a model of order one, and Region 2 by a model of order two. Jacknife and ordinary least squares estimates of regression coefficients for regression models relating to the single and double sunspot series are given in Table 8.4.

The Jacknife estimates do not differ significantly from the ordinary least squares estimates and thus in no way improve the model. The climatological value of the best models can be assessed by comparing observed variances of the regional temporal series for Region 1 and 2 and the minimum prediction sum of squares underlined in Tables 8.2 and 8.3. Such a comparison shows that since the respective values are very similar neither the single nor double sunspot series provided useful predictor variables in the multiple linear regression models.

EXPONENTIAL EXTRAPOLATION

In fitting exponential models to the regional temporal series, orders up to seven have been considered.

Therefore, 56 terms have been used to obtain the mean sum of squares of the forecasting errors, seven terms being lost in the initialisation process. The smoothing constant is not estimated from the data in this type of model, but is chosen so that for a given order model, the mean sum of squares of the forecasting errors is a minimum.

Initially, the smoothing constant, a, has been taken from zero to 0.7 in steps of 0.1. However, since a number of minimum sum of squares is obtained for $a = 0.1$, the process has been repeated. In this case a has been taken from 0.02 in increments of 0.02 up to 0.1. The final value of a, and the order k, of the model using the minimum sum of squares, *m.s.e.*, as the fitting criterion are given for the regional temporal series of Region 1 and 2 in Table 8.5.

TABLE 8.5 - ESTIMATED CONSTANTS IN THE BEST FITTING EXPONENTIAL MODELS FOR REGIONS 1 AND 2

Region	k	a	*m.s.e.*
1	4	0.06	16690
2	4	0.02	5576

The smaller the value of a the less influence has the latest term in the series on the forecast. Because a is so small in this case, Table 8.5, the fitting process has been repeated leaving the latest term out of the model, (Table 8.6).

A comparison of the entries in Tables 8.5 and 8.6 shows that for neither region is the fit of the model improved by omitting the latest term in the series. As in the case of the autoregressive and multiple linear regression models the fitting criterion, in this case the mean sum of squares of the residual errors, does not differ significantly from the observed variance and the model must be rejected for predicting rainfall changes. Before rejecting exponential models altogether, however, it is necessary to explore Brown's (1963) assertion that the triple exponential model may be more superior.

TABLE 8.6 - ESTIMATED CONSTANTS IN THE BEST FITTING EXPONENTIAL MODELS FOR REGIONS 1 AND 2 WHEN THE LATEST TERM IN THE SERIES HAS NOT BEEN INCLUDED IN THE MODEL

Region	k	a	$m.s.e.$
1	3	0.12	16406
2	3	0.04	5579

The triple exponential model produces excessively large minimum mean sum of squares of the forecasting errors, namely 51301 and 29124 for Regions 1 and 2 respectively. Some improvement in the fit of this model can be obtained when the estimation of a was obtained using three subsets of the data series. However, the mean sum of squared residuals in this case is still excessive, being 25568 and 11735 for Regions 1 and 2 respectively, showing that the triple exponential model is even less satisfactory than the single.

TRIGONOMETRIC REGRESSION AND THE PREDICTION OF REGIONAL TEMPORAL SERIES

The initial values for the λ_j in equation 8.26 have been estimated from each of the peaks in the spectra of the regional temporal series (Table 5.5). Values of λ_j have been taken on each side of a spectral peak and these have been used to estimate the regression coefficients. The values of the wavelengths giving the minimum mean sum of squares of the residual errors have been used in the final model. In most cases, the spectral peaks occur at wavelengths not far removed from the wavelengths of the best fitting trigonometric function. Values for D_i, E_i and λ_j are given in Table 8.7 for Regions 1 and 2. Regression coefficients D_i and E_i that are significant at the 5 per cent level are indicated thus *.

TABLE 8.7 - VALUES OF D_i, E_i AND λ_j IN EQUATION 8.26 FOR REGIONS 1 AND 2.

Region	α_o	α_i	β_i	λ_j (years)
1		-20.77	-66.03*	18.6
		40.76*	51.62*	3.60
		60.44*	10.58	2.34
	2.92			
Standard error of the residuals = 105.27 mm				
2		16.7	34.49*	10.4
		37.87*	4.88	3.50
		32.16*	2.45	3.00
		24	9.41	2.10
	-0.32			
Standard error of the residuals = 63.09 mm				

TABLE 8.8 - VALUES OF THE FITTING CRITERIA FOR EACH OF THE FORECASTING MODELS FITTED TO THE REGIONAL TEMPORAL SERIES FOR REGIONS 1 AND 2

Region	Auto regr.	Mult. regr. cycle		Exponential		Trig. regr.	Est. variance regional series
		single	double	single	triple		
1	15173	16254	15543	16690	51301	10009	16329
2	5171	5392	5371	5576	29124	3242	5572

Using equation 8.26 and Table 8.7 it is possible to determine the fitting criterion of mean squares of the residual errors. These mean squares of the residual errors for Regions 1 and 2 are 10009 and 3242 respectively. These values are not closely similar to the actual observed variance of the regional temporal series for both summer and all-seasons rainfall areas and suggest that the trigonometric models offer the best potential for forecasting future rainfall changes. For both Regions 1 and 2 the goodness of fit for the trigonometric model is significant at the 1 per cent level.

8.4 COMPARISON OF MODELS

Autoregression, multiple linear regression using sunspot numbers as predictor variables, single and triple exponential, and trigonometric models have been fitted to the regional temporal series of Regions 1 and 2 The results obtained from this exercise are given in Table 8.8. In this table the values of the fitting criteria are given for each type of model.

It is clear that the only worthwhile model is that using trigonometric curve fitting. For both the autoregressive and multiple linear regression models the reduction in the sum of squares provided by the models is negligible. The triple exponential models are a good example of what can happen when a completely inappropriate model is used. The model is incapable of providing forecasts that converge on the true value, and what is commonly termed an explosion takes place, (Box and Jenkins, 1970; Kendall and Stuart, 1966; Tintner, 1952).

It is only trigonometric curve fitting that provides a basis for predicting future rainfall changes but at best this approach is somewhat crude.

8.5 PREDICTING MAXIMA AND MINIMA SOLAR ACTIVITY

As has been pointed out, the general pattern of rainfall over much of South Africa is similar to the variation in solar activity as measured by sunspot numbers. Over the summer rainfall region the double sunspot cycle and regional temporal series (Region 1) are in phase, whereas in the southern Cape area the single sunspot cycle and the regional temporal series (Region 2) are in anti-phase. Thus if it is possible to predict sunspot variation, statements can be made concerning likely rainfall conditions.

Using the data presented by Lamb (1972) concerning the inter peak and trough time periods associated with the single sunspot cycle it is possible to relate the time taken to reach a peak from the previous trough and vice versa. The regression equation for so doing is equation 8.27.

$$\hat{T}_{max} = 1.230 + 0.466\ t_{min} \tag{8.27}$$

where $\hat{T}_{max}$ represents the period of time to reach a maximum in solar activity and t_{min} is the corresponding preceding period of fall. An analysis of variance shows that the goodness of fit of this linear model is significant at the 5 per cent level. In addition the estimated standard error of the residuals and regression coefficient are 1.025 and 0.178 respectively. Equation 8.27 therefore provides adequate forecasts for the next period of rise to a maximum of solar activity given the previous period of fall to a minimum.

In a similar manner the prediction equation for the time of fall given the previous time of rise is given by

$$\hat{T}_{min} = 9.204 - 0.587\ t_{max} \tag{8.28}$$

Again, an analysis of variance for this regression indicates that the fit is significant at the 5 per cent level. The estimated standard errors of the residuals and regression coefficient are 1.221 and 0.235 respectively. Equation 8.28 will predict the time of the next minimum given the previous time of rise adequately.

One may apply the same method to the double sunspot cycle except that instead of predicting the time to the next maximum given the previous period of fall and vice versa, in this case one predicts the time to a maximum from zero from the rise from a minimum to that zero. Thereafter the fall from a maximum to the next zero is predicted from the rise from the previous zero to the next maximum and so on.

8.6 SOME PREDICTIONS FOR THE FUTURE

To date it has proved impossible to forecast long term rainfall totals over South Africa in any way at all. The work presented in this thesis shows that, whereas forecasts of rainfall one or two years ahead is not feasible without further research, the prediction of extended spells of above and below normal rainfall is possible, both qualitatively and semi-quantitatively. Such forecasts for the summer rainfall area (Region 1) and the all-seasons area of South Africa (Region 2) will now be attempted.

The proviso that applies to all predictions for the future is that the general circulation of the atmosphere over Southern Africa and adjacent oceans will obey the dictates of its controls in the next few decades as it has in the past half dozen, so that the rhythm and patterns of rainfall changes of the past 60 years or so continues to repeat itself in the near future.

A QUALITATIVE ASSESSMENT

Subjective extrapolation of mean temporal series

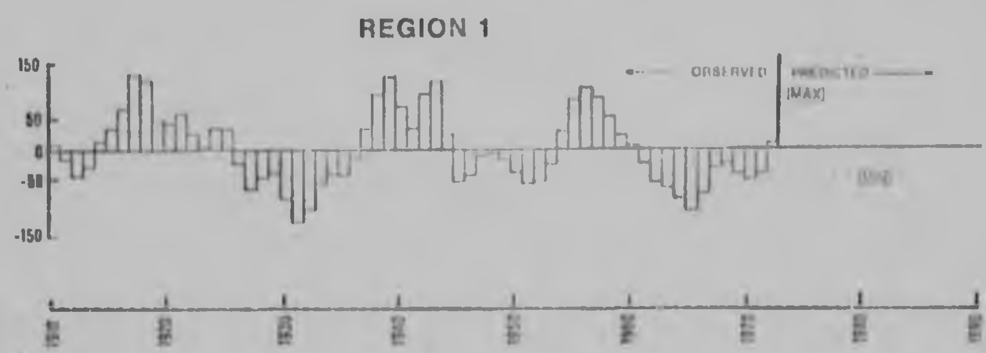


Fig. 8.5: Showing a subjective extrapolation of the temporal series for Region 1. If conditions continue unchanged, the present wet period is suggested to have reached its maximum in an interval centred on 1975/76. Similarly, a drier than normal period is indicated for the early to middle 1980's.

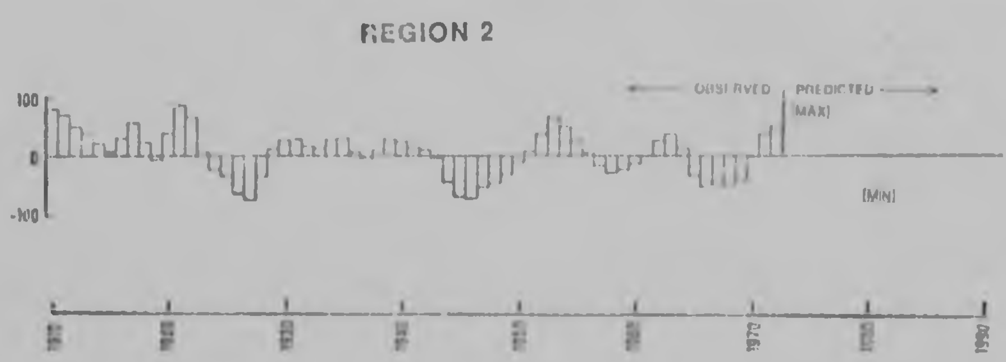


Fig. 8.6: Showing a subjective extrapolation of the temporal series for Region 2. Past and prevailing conditions of the atmospheric circulation suggest that a period of above normal rainfall will be experienced around 1978. Below normal conditions can be expected for a period centred on the early eighties.

into the future suggests that in the case of the summer rainfall region the present wet seasons that the country is experiencing will continue, but the peak has probably been reached (1975/76) and the above average rainfall will decrease toward the end of the decad (Fig. 8.5). Likewise the beginning of the next decade might be drier than normal with a trough centred on 1986 (Fig.8.5). This is not to say that for the rest of the seventies one or two years will not experience droughts or that the middle and late eighties will not experience occasional floods. On the basis of past experience all that can be said is that there is a good chance that the cumulative *total rainfall* in the summer rainfall areas will be above normal for the rest of this decade and below normal for the next. The range of the fluctuation will vary between 20 and 30 per cent of the total annual precipitation at individual stations covering thousands of square kilometres over the north-eastern part of the country.

Likewise in the southern Cape area (Region 2) subjective extrapolation of the regional temporal rainfall series given in Fig. 8.6 suggests that the interval centred on 1975 will be a period of slightly above normal rainfall receipt whilst the late eighties will experience more serious deficits. Below normal rainfall can be expected over an interval centred on 1982. The remarks made earlier, with respect to Region 1, regarding the year to year variation in rainfall behaviour apply equally well to Region 2. The range of the fluctuation will vary between 15 and 25 per cent of the total annual precipitation at individual stations in the southern Cape area of the country.

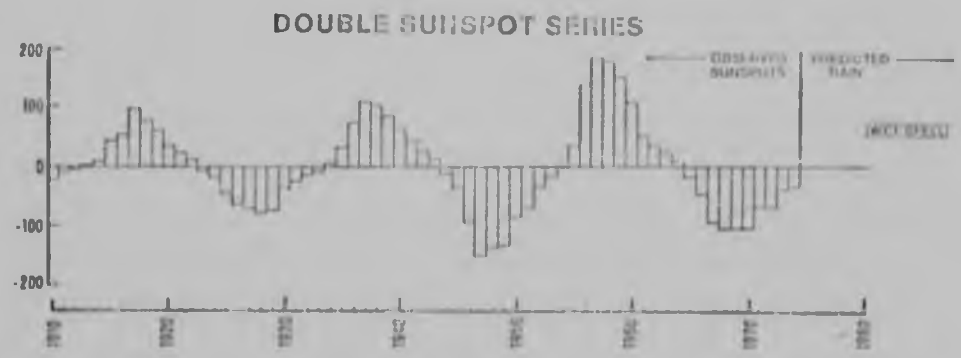


Fig. 8.7: Using solar activity as measured by the double sunspot series as a forecasting aid, indicates that Region 1 will continue to experience above normal rainfall during the remainder of the seventies and early eighties.

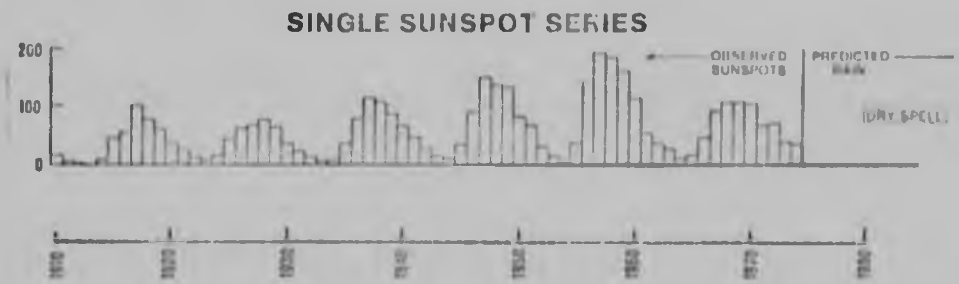


Fig. 6.8: Variation in solar activity as measured by the single sunspot series suggests that the rainfall receipts over Region 2 will be below normal during the late seventies and early eighties.

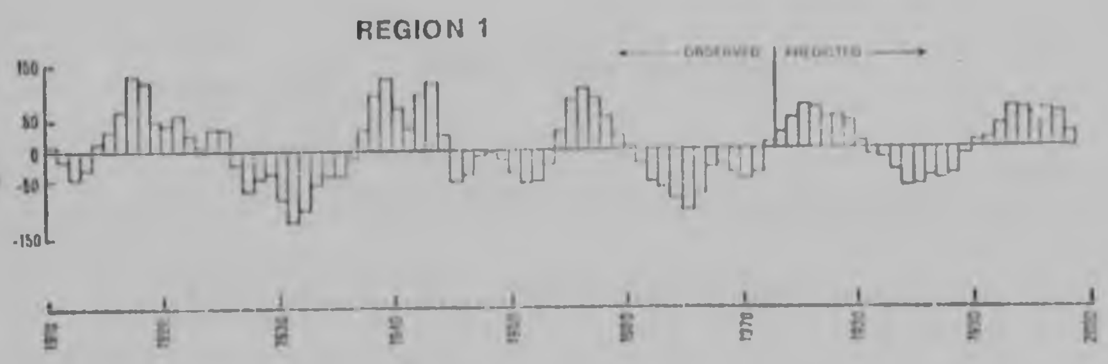


Fig. 8.9: Forecasting rainfall receipts using a trigonometric model suggests that Region 1 will experience above normal rainfall receipts during the seventies and nineties. It is likely to be drier than normal during much of the eighties.

RAINFALL ESTIMATION FROM SOLAR ACTIVITY

Using this approach predictions of future sunspot action may be related to the appropriate rainfall changes bearing in mind the phase relationships referred to earlier. Thus in Fig. 8.7 it is suggested that in the summer rainfall region of the country the present spell of above average rainfall will have reached its peak by about 1978. It is impossible to make a prediction further into the future than this until further sunspot data are available for the years 1975 to about 1980. Once this is at hand a prediction can be made for the ensuing period.

Similarly in the southern Cape region this forecasting approach suggests that the interval centred on 1979 will see rainfall receipts that are below average in value (Fig. 8.8).

EXTRAPOLATION FROM THE TRIGONOMETRIC MODEL

Whereas the approach used in estimating future sunspot changes is dependent only on the pattern of variation in a few preceding years, extrapolation using the trigonometric model uses all the available data and therefore forecasting future changes must again be based on the premise that the pattern and rhythm of change that has prevailed since at least the turn of the century will continue to prevail into the near future. Given that it does, trigonometric extrapolation into the future suggests that in the summer rainfall region the present wet interlude is centred on 1975/76. For the next few years the average rainfall receipts will decrease, and maximum drier than normal conditions can be expected to prevail over an interval centred on 1985 (Fig. 8.9).

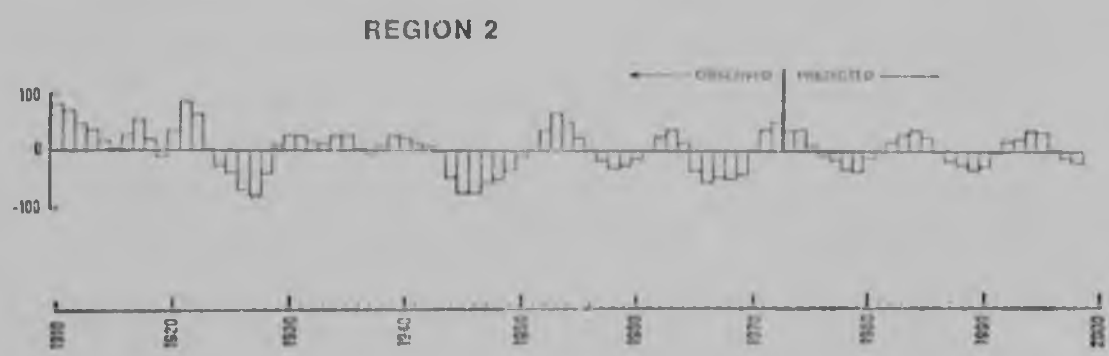


Fig. 8.10: Forecasting rainfall over Region 2, using a trigonometric model indicates generally droughty conditions during the late seventies and late eighties. The early eighties and middle nineties are likely to be periods when above average rainfall receipts will be experienced.

The southern Cape Province can expect a generally droughty period over the interval centred on 1978. This period when a deficit in rainfall is experienced will give way to above average conditions likely to be centred on 1984. A period experiencing a greater rainfall deficit than that of 1978 will probably occur around 1989 (Fig. 8.10).

* * * * *

Prediction of exact year to year changes has not proved feasible. On the other hand it would appear that estimation of future spells of wet and dry years is possible providing the general circulation over South Africa continues only to change rhythmically as it has done previously and not to undergo a radical new mode of behaviour.

PART V

CONCLUSIONS

CHAPTER 9

SUMMARY

South Africa's rainfall has been a popular topic for discussion among both informed scientists and enthusiastic amateurs since the inception of European settlement. In particular, theories of progressive dessication have been propounded over the years and until very recently the debate has continued as to whether South Africa is becoming wetter or drier. The majority of research workers have, in many studies, sought to resolve the controversy by analysing rainfall data for trends. Relatively few workers have been concerned with the oscillatory nature of South African rainfall and when they have they have been concerned only with isolated stations widely distributed over the country. In almost all studies there has been a tacit assumption that the country as a whole has behaved in a similar fashion in response to climatic changes.

In this thesis attempts have been made to resolve the controversy of progressive dessication, to examine the oscillatory nature of South African rainfall and to delimit regional changes of climate. The specific results of the investigation are best itemised as follows.

A. TREND

1. Over the whole of South Africa, and out of 157 cases, only 26 showed trend significant at the five per cent level. Of

the cases showing trend only two exhibited positive trend. The 24 cases of negative trend were distributed fairly evenly over the country.

2. Where it has been possible to extend trend analysis back in time to 1880, it has been clearly demonstrated that the presence or otherwise of trend is highly time dependent. In only three cases out of 30 could persistent trend be found over the total period 1880-1972. In many cases the effect of an excessively wet decade 1880-1890 introduced spurious trend into the data.

3. Statistical testing of regression data shows that it is not possible to prove conclusively, in a statistical sense, that South Africa has not undergone progressive dessication. However, this is hardly surprising since all statistical testing that is based on the assumption of independent data will prove inconclusive. The *a priori* assumption, and the results of this thesis, showed quite conclusively that South African rainfall data are not independent.

4. Although the number of individual stations showing negative trend is larger than might be expected by chance (assuming independent data), the impossibility of providing any physical explanation for why only a small fraction of stations should experience this decline mitigates against accepting the suggestion that South Africa as a whole is undergoing progressive dessication.

5. Delimiting rainfall regions using principal component analysis and subsequent factor rotation provides the data matrix input for using a linear two-way fixed effects model to isolate the independent components of annual rainfall totals. These components are: regional temporal effect, regional spatial effect, and micro or local effect.

6. Trend analysis on regional temporal series shows that neither individual regions nor South Africa as a whole have experienced either decreasing or increasing rainfall receipts over the period of meteorological records.

7. Trend analysis on the micro-effects obtained from the linear two-way fixed effects model, shows significant trend at 34 stations, 16 positive and 18 negative. The probability of such a distribution having arisen by chance is less than 0.01.

8. The stations with records whose local effect series denote trend are randomly distributed over the country. In many cases stations with positive local effect series components of trend are adjacent to others with negative trend.

B. OSCILLATORY SECULAR VARIATION

1. Inspection of smoothed rainfall series, determination of pentade and decade periods of maximum and minimum rainfall, Fourier analysis, spectral analysis and principal component analysis all show that it is possible to isolate, on a regional scale, areas of South Africa experiencing like patterns of secular rainfall variation.

2. These regions may be identified with in-phase oscillations of the order of 20, 10, 3-4 and 2-3 years.

3. The 16-20 year oscillation is strongest over the northern part of the country which experiences a seasonal summer rainfall. To a lesser extent this wave is present in the all-year rainfall belt running along the south coast. The 10-year wave appears over the southern Cape Province and adjacent inland area, experiencing an all-year rainfall regime. At the higher frequency end of the spectra, the 3-4 year oscillation shows no spatial dependence. On the other hand, the 2-3 year wave is strongest over that area of the country having a seasonal six-monthly rainfall regime.

4. Spatial gradients of change between one region and another have been determined by mapping the variance associated with Fourier waves and the variance associated with specific spectral bands determined by spectral analysis.

5. Whereas individual spectral peaks may not be statistically significant, the fact that large numbers of spectra possess peaks in the same frequency bands may offset doubts one may have about the individual peaks.

6. Isolating regional temporal effects using the linear two-way fixed effects model and subjecting the time series of these effects to spectral analysis shows that the spectral peaks for regions as a whole *are* statistically significant.

7. The amplitude of the 20-year wave is between 20 and 30 per cent of the annual mean rainfall throughout the summer rainfall area in the north-eastern part of South Africa. The amplitude of the 10-year wave is between 15 and 25 per cent of the annual mean rainfall throughout the all-seasons rainfall area in the southern Cape Province.

8. Cross-spectral analysis between temperature and rainfall data shows that an inverse correlation exists between the two variables at a period of the order 20 years. It is significant that at this wavelength the rainfall and temperature oscillations are in almost perfect anti-phase.

C. FORECASTING FUTURE CHANGES

1 Subjective extrapolation of composite time series of regional temporal effects suggests that the present wet spell over the summer rainfall region will decrease toward the end of the present decade. In the southern Cape area a drier than average period will be experienced around 1978, to be followed by a wetter than average period centred at about 1982.

2. Autoregressive, multiple linear regression and single and triple exponential smoothing have all been shown to be inadequate for forecasting future changes.

3. Using a simple linear regression model it has proved possible to predict both single and double sunspot cycle variations for limited periods in the future, based on the fact that precipitation in the summer rainfall area is currently in phase with the double sunspot cycle and that precipitation in the all-seasons southern Cape rainfall region is in anti-phase with the single sunspot cycle, it is possible to suggest that

   (i) in the summer rainfall region the present wet spell arises from an above average rainfall decade centred on 1976.

   (ii) in the southern Cape the pentad centred on 1979 will be a period when the rainfall receipts will be below average.

4. The trigonometric regression model fits the observed regional temporal effect data at the one per cent level of significance. Nevertheless the standard error of forecasting estimates are still too large to enable meaningful one year ahead predictions to be made. However, it is possible to predict wet and dry spells with some qualitative assurance. On this basis the model suggests that

   (i) the summer rainfall region will become drier toward the end of the 1970s to be replaced by a dry spell with maximum severity centred on 1985.

   (ii) the southern Cape will experience a deficit over the interval centred on 1978, to be replaced by a wetter than average interval centred on 1984.

5. All the above forecasts are contingent upon the fact that the general circulation of the atmosphere, and therefore the consequent rainfall changes, will oontinue to vary in the future as it has over the period of meteorological record to date.

This thesis has investigated secular variation in rainfall over South Africa. On the basis of the research carried out, the often popular belief that South Africa is undergoing progressive dessication must be rejected. But although there is no significant evidence to suggest that the general circulation over Southern Africa and its surrounds is experiencing long-term *permanent* change, there is considerable evidence to support the theory that the general circulation, which in this case must involve the southern hemisphere's high pressure belt, is behaving in an *oscillatory* manner. These oscillations in the atmosphere manifest themselves over South Africa as a 20-year wave in rainfall over the summer rainfall area, and a 10-year wave over the all-seasons rainfall area of the southern Cape Province. In addition to these secular variations a spatially ubiquitous 3-4 year oscillation is clearly in evidence, as too, is the spatially dependent oscillation in the interval 2 to 3 years occurring over the six-monthly rainfall region of the southern Cape Province.

The oscillatory behaviour of rainfall over South Africa has enabled forecasts for future wet and dry spells to be made. Forecasts using three different methods are in essential agreement regarding the centres of periods experiencing above and below average rainfall. In the summer rainfall area these periods relate to decades, whilst in the southern Cape Province pentades are the appropriate intervals. The fundamental proviso governing such predictions is that the general circulation of the atmosphere continues to behave in the near future as it has in the best part of the last century. If this continues to be so, then the estimation of near future rainfall changes made in this thesis should prove valid. Despite its high degree of organisation the general circulation of the atmosphere is a turbulent system, thus because the circulation over South Africa has experienced regular changes previously does not mean it will continue to do so for an extended period of time and there is no way of predicting when the present pattern of rhythmic change will alter or even disappear altogether.

LIST OF THE MAIN SYMBOLS USED IN CHAPTERS 3 AND 8

A	the regional temporal effect
a	a coefficient in the exponential model
$A(k)$	coefficient in the co-spectrum
A_h	cosine coefficient in Fourier series
a_j	weighting factor in smoothing procedure
am_h	amplitude of Fourier wave
B	the regional spatial effect
b	a coefficient in the exponential model
$B(k)$	coefficient in the quadrature spectrum
B_h	sine coefficient in Fourier series
c	a coefficient in the exponential model
$c(k)$	the serial covariance
C_h	cosine term in Fourier series
$\overline{C}(f)$	smoothed spectral power
$\overline{co}(f)$	smoothed co-spectrum
$\overline{Coh}(f)$	smoothed coherency spectrum
d	a coefficient in the exponential model
D_i	cosine coefficient in trigonometric regression
E_i	sine coefficient in trigonometric regression
$E(\ )$	expected value
F	multiple of maximum lag in spectral analysis
f	the frequency
g	$(X'X)^{-1}x'$
i	summation index
j	summation index
$J(\hat{\theta})$	..ife

k	a summation index
l	a summation index
m	maximum lag in spectral analysis
n	number of observations, or terms in a time series
P	counting variable in Tau
p	order of binomial smoothing
$P(l)$	prediction sum of squares
$\overline{ph}(f)$	smoothed phase-spectrum
$\bar{q}(f)$	smoothed quadrature spectrum
$R(f)$	frequency response function
$r(s)$	correlation coefficient
$r'(s)$	alternative form of serial correlation
$\bar{R}(f)$	smoothed spectral density
$S(t)$	elements in binomially smoothed series
S_n	sine term in Fourier series
t	time
$u(t)$	an observation at time t
$Var(\ )$	variance
$\underline{X}$	matrix on independent variable
X_i	independent variable in regression model
$x_{(i)}$	independent variable removed in $P(l)$
X_{ij}	variable in analysis of variance model
$\underline{Y}$	vector on dependent variable in regression model
$Y(i)$	dependent variable in regression
$y_{(i)}$	dependent variable removed in $P(l)$
$\hat{Y}_{(il)}$	least squares estimate of $E(Y_i)$ based on l-th subset of k possible predictor subsets

α_i	regional temporal series
$\underline{\beta}$	column vector of regression coefficients
β_i	regional spatial series
$\hat{\beta}_i$	estimated β on $(n-1)$ terms in a regression
Δt	sampling interval
$\underline{\varepsilon}$	column vector of residuals
ε_{ij}	residuals in analysis of variance model
ϕ_h	phase angle
$\lambda(k)$	Parzen's lag window
λ_j	wavelength
μ	overall mean in analysis of variance model
τ	Kendall's correlation coefficient
$\hat{}$	estimate

REFERENCES

Afifi, A.A. and Azen, S.P., 1972: *Statistical Analysis; a computer oriented approach*, Academic Press, London.

Agric. Jour. Union South Africa, 1913-1914: Is South Africa drying up? Vol.6, 1913, 693-694, 1027-1029; Vol.7, 1914, 113-114, 150-154, 259-261, 423-424, 523-538, 732-739, 894-898; Vol.8, 1915, 107-113, 262-265.

Allen, D.M., 1971a: Mean square error of prediction as a criterion for selecting variables. *Technometrics*, 13, 469-75.

Allen, D.M., 1971b: The prediction sum of squares as a criterion for selecting predictor variables. Univ. of Kentucky, Dept. of Statistics, *Tech. Report 23*.

Allen, D.M., 1972: Biased prediction using multiple linear regression. Univ. of Kentucky, Dept of Statistics, *Tech. Report 36*.

Allen, D.M., 1974: The relationship between variable selection and data augmentation and a method of prediction. *Jour. Am. Statist. Soc.*, 16, 125-127.

Anderson, T.W., 1971: *The Statistical Analysis of Time Series*, John Wiley, London.

Barber, F.H., 1970: Is South Africa drying up? *Agric. Jour.*, Cape of Good Hope, 86, 167-170.

Bargman, D.J., 1960: *Tropical Meteorology in Africa*, Editor, Munitalp Foundation, Nairobi.

Barry, R.G. and Perry, A.H., 1973: *Synoptic Climatology: methods and applications*, Methuen, London.

Baur, F., 1949: Zurückführung des Grosswetters auf solare Erschemungen. *Archiv. f. Met., Geophys. und Biokl., A*, 1, 358-74, Vienna.

Baur, F., 1956, 1958: *Physikalisch-statistiche Regeln als Grundlagen für Wetter-und Witterungsvorhersagen*, I, 1956; II, 1958. Frankfurt (Akad. Verlag).

Belinski, N.A., 1957: *Use of certain features of atmospheric processes for long-range forecasts*, Leningrad (Gidromet. Izdat.) (English translation Ohio U.S. Air Force base, 1959).

Bendat, J.S. and Piersal, A.G., 1968: *Measurement and analysis of random data*, John Wiley, New York.

Bissell, A.F. and Ferguson, R.A., 1975: The Jacknife: toy, tool or two-edged weapon? *The Statistician*, 24, 79-100. *

Blackman, R.B. and Tukey, J.W., 1958: *The measurement of Power Spectra*, Dover Publications, New York.

Box, G.E.P. and Jenkins, G.M., 1970: *Time Series Analysis, forecasting and control*, Holden-Day, London.

Brook, G.A. and Mametse, M.N., 1969: Rainfall trend patterns in South Africa. *S.A. Geog. Jour.*, 52, 134-138.

Brooks, C.E.P., 1923: Variations in level of the central African lakes. *Geophys. Mem.*, 20, London, (Met. Office).

Brooks, C.E.P., 1934: Variation of the annual frequency of thunderstorms in relation to sunspots. *Quart. Jour. Roy. Met. Soc.*, 60, 153-65.

Brown, R.G., 1959: *Statistical forecasting for inventory control*, McGraw-Hill, New York.

Brown, R.G., 1963: *Smoothing, Forecasting and Prediction of discrete time series*, Prentice-Hall International Inc., London.

Cattell, R.B., 1952: *Factor analysis*, Harper, New York.

Cattell, R.B., 1965: Factor Analysis: An introduction to essentials I and II. *Biometrics*, 21, 190-215; 405-435.

Chatfield, C. and Pepper, M.P.G., 1971: Time Series Analysis: example from Geophysical Data. *Jour. Roy. Statist. Soc.*, C, 20, 217-238.

Child, D., 1970: *The essentials of Factor Analysis*, Holt, Rinehart and Winston, London.

Colgate, M.G., 1975: An attempt to predict the anomalies in the monthly mean sea level pressure field a month ahead. *Quart. Jour. Roy. Met. Soc.*, 101, 267-280.

Cox, G.W., 1925: Periodicity in rainfall. *Trans. Roy. Soc. South Africa*, 12, 295-299.

Craddock, J.M., 1965: A meteorological application of principal component analysis. *The Statistician*, 15, 143-156.

Craddock, J.M. and Flintoff, S., 1970: Eigenvector representations of northern hemispheric fields, *Quart. Jour. Roy. Met. Soc.*, 96, 124-129.

Craddock, J.M. and Flood, C.R., 1969: Eigenvectors for representing the 500 mb geopotential surface over the northern hemisphere. *Quart. Jour. Roy. Met. Soc.*, 95, 576-593.

Currie, L., 1972: Special Spectra? *Ann. Assoc. Amer. Geogr.*, 62, No. 3.

de Blij, H.J., 1964: *A geography of subsaharan Africa*. Rand McNally, Chicago.

Dyer, T.G.J., 1974: Distribution of Rainfall over South Africa. *S. Afri. Journ. Sci.*, Vol. 70, 6, 180-182.

Dzerdzeevskii, B.L., 1963: *Sukhovies and Drought Control*. Editor, Israel Program for Scientific Translations.

Dzerdzeevskii, B.L., 1968: Climatic fluctuations and the problem of super long range forecasting. *Izv. Akad. Nauk.*, Ser.Geog, (Moscow), 43-55 (in Russian).

Erat, S.J., 1973a: Time series analysis of annual temperatures. *Mon. Wea. Rev.*, 101, 501-504.

Erat, S.J., 1973b: Spectral analysis of daily maximum and minimum temperature series on the East Slope of the Colorado Front Range. *Mon. Wea. Rev.*, 101, 505-509.

Fabricius, A.F., 1964: Drought Investigation: Northwestern Transvaal. *S. Afri. Jour. Sci.*, 60, 267-275.

Glasspole, J., 1925: Relation between annual rainfall over Europe, Oxford and Glenqnoich. *Brit. Rainfall*, 65, 254-69.

Glover, J., Robinson, R., Henderson, J.P., 1954: Provisional Maps of the reliability of annual rainfall in East Africa. *Quart. Jour. Roy. Met. Soc.*, 80, 602-609.

Godin, G., 1972: *The Analysis of Tides*, Liverpool University Press, Liverpool, U.K.

Gower, J.C., 1966: Some distance properties of latent roots and vector methods used in multivariate analysis. *Biometrika*, 53, 325-338.

Granger, C.W.T. and Hatanaka, M., 1966: *Spectral Analysis of Economic Time Series*, Princeton Univ. Press, New Jersey.

Granger, C.W.T. and Hughes, A.O., 1968: Spectral Analysis of short time series: a simulation study. *Jour. Roy. Statist. Soc.*, A, 130, 83-99.

Granger, C.W.T. and Hughes, A.O., 1971: A new look at some old data: the Beveridge wheat price series. *Jour. Roy. Statist. Soc.*, A, 134, 413-28.

Gray, H.L. and Schucany, W.R., 1972: *The Generalised Jackknife Statistic*, Marcel Dekker, New York. *

Griffiths, J.F., 1972: *World Survey of Climatology, Vol.10; Climates of Africa*, Elsevier Pub. Comp. London.

Grimmer, M., 1963: The space filtering of monthly surface temperature anomaly data in terms of pattern, using empirical orthogonal functions. *Quart. Jour. Roy. Met. Soc.*, 89, 395-408.

Gudmendsson, G., 1971: Time series analysis of imports, exports, and other economic variables. *Jour. Roy. Statist. Soc.*, A, 134, 383-412.

Hannan, E.J., 1967: *Time Series Analysis*, Methuen, London.

Hildebrand, F.B., 1956: *Introduction to Numerical Analysis*, McGraw-Hill, New York.

Hofmeyr, W.L. and Schulze, B.R., 1963: Temperature and rainfall trends in S. Afr. during the period of meteorological records. Proc.Rome Conf. on changes of climate with special reference to arid zones. *Arid Zone Research Series, Unesco*, 20, 81-85.

Hutchins, D.E., 1888: Cycles of drought and good seasons in South Africa. *Times Office*, Wynberg, 137.

Hutchins, D.E., 1897: Rainfall of South Africa. *Agric. Jour.*, Cape of Good Hope, 11, 701-10.

Hutchins, D.E., 1898: Agricultural weather forecasts in England and in South Africa. *Agric. Jour.*, Cape of Good Hope, 12, 138-143; 211-19; 267-76.

Hutchins, D.E., 1905: Weather forecasts. *Agric. Jour.*, Cape of Good Hope, 26, 105-15.

Hutchins, D.E., 1907: The cycle year 1905 and the coming season. *Trans. S. Afric. Phil. Soc.*, 16, 237-50.

Jackson, S.P., 1952: Atmospheric circulation over South Africa. *S.A. Geog. Jour.*, 34, 1-15.

Jackson, S.P., 1961: *Climatological Atlas of Africa*, CCTA/CSA, Nairobi.

Jaganathan, P. and Parthasarathy, B., 1973: Trends and periodicities of rainfall over India. *Mon. Weath. Rev.*, 101, 371-374.

Julian, P., Ruff, I., Nord, J., 1957: Possible responses of terrestrial atmospheric circulation to changes in solar activity. *Tech. Rep. No. 1*, Boulder (Univ. Colorado High Altitude Observatory, Institute of Solar-terrestrial Research).

Jeffers, J.M.R., 1967: Two case studies in the application of principal component analysis. *App. Stat.*, 16, 225-36.

Jenkins, G.M., 1961: General considerations in the analysis of spectra. *Technometrics*, 3, 133-66.

Jenkins, G.M. and Watts, D.G., 1968: *Spectral Analysis and its applications*, Holden-Day, San Francisco.

Johnston, J., 1972: *Econometric Methods*, McGraw-Hill, New York.

Jolliffe, I.T., 1972: Discarding variables in a principal component analysis. I: artificial data. *Jour. Roy. Statist. Soc.*, C, 21, 161-173.

Kaiser, H.F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.

Kaiser, H.F., 1961: A note on Guttman's lower bound for the number of common factors. *Brit. Jour. Statist. Psychology*, 14, 1-2.

Kanthack, F.E., 1930: The alleged desiccation of S. Africa. *Geogr. Jour.*, Vol. 76, No. 6, Dec. 1930.

Keen, C.S., 1971: *Rainfall spectra and the delimitation of rainfall regimes in S. Africa*, Unpublished M.Sc. thesis, Univer. of Witwatersrand, Johannesburg.

Keen, C.S. and Tyson, P.D., 1973: Seasonality of South African rainfall: a note on its regional delimitation using spectral analysis. *Arch. Met. Geoph. Biokl.*, Ser. B., 21, 207-214.

Kendall, M.G., 1945: An analysis of oscillatory time series. *Jour. Roy. Statist. Soc.*, 108, pt I-II, 93-129.

Kendall, M.G., 1962: *Rank Correlation Methods*, Charles Griffin, London.

Kendall, M.G., 1973: *Time Series*, Charles Griffin, London.

Kendall, M.G. and Stuart, A., 1961: *The Advanced Theory of Statistics*, Charles Griffin, London, Vol. 2.

Kendall, M.G. and Stuart, A., 1966: *The Advanced Theory of Statistics*, Charles Griffin, London, Vol. 3.

King, J.W., 1973: Solar radiation changes and weather. *Nature*, Vol. 245, No. 5426, 443-446.

King, J.W., 1974: Weather and the Earth's magnetic field. *Nature*, Vol. 247, No. 5437, 131-134.

Kokot, D.F., 1948: *An Investigation into the evidence bearing on recent climatic changes over Southern Africa*, Irrigation Dept., Pretoria.

Köppen, W., 1873: Über mehrjährige Perioden der Witterung, nisbesondere über die 11-jährige Periode der Temperatur. *Z. f. Met.*, 8, 241-8; 257-67, Vienna.

Kupper, L.L., 1972: Fourier analysis and spherical harmonic regression *Jour. Roy. Statist.*, C, 21, 121-130.

Kutzbach, J.E., 1967: Empirical eigenvectors of sea-level pressure, surface temperature, and precipitation complexes over North America. *Jour. Appl. Met.*, 6, 791-802.

Landsberg, H., 1960: *Physical Climatology*, Gray Printing, Du Bois, Pennsylvania.

Lamb, H.H., 1972: *Climate: present, past, and future*, Vol. 1, Methuen, London.

Levinkind, L., 1941: Droughts and South Africa: a preliminary study of their extent, severity, and frequency of occurrence. *Farming in S. Afri.*, March, 84-89.

Livingstone, D.,1857: *Missionary Travels and Research in South Africa,* London.

Meldrum, C., 1873: On a supposed periodicity of the rainfall. *Quart. Jour. Roy Met. Soc.,* 1, 130-33.

Mironovitch, V., 1960: Sur l'évolution séculaire de l'activité solaire et ses liaisons avec la circulation atmosphérique générale. *Met. Abhandl.*, 9, 3. Berlin (Inst. f. Met. Geophys. der Freien Universität).

Mitchell, J.M., 1965: *The solar inconstant, Proc. Seminar on possible responses of weather phenomena to variable extraterrestrial influences,* NCAR Tech., Note, TN-8, Bouler, Colorado.

Mitchell, J.M., et al., 1966: Climatic Change. *W.M.O., Techn. Note 79.*

Moffat, R., 1842: *Missionary Labours and Scenes in Southern Africa,* London.

Moran, P.A.P., 1949: The spectral theory of discrete stochastic processes. *Biometrika,* 36, 63.

Morrison, J.T., 1900: On some periodical changes in the rainfall at the Roy. Obs., Cape of Good Hope, since 1841. *Report Met. Comm., Cape Town,* 195.

Mustel, E.R., 1967: *Manifestation of Solar Activity in the Troposphere and the Stratosphere,* Moscow, (Astron. Council of the U.S.S.R. Acad. Sci. and U.S.S.R. Hydromet. Centre). In English.

Nelson, C.R., 1973: *Applied Time Series analysis for managerial forecasting,* Holden-Day, London.

Nevill, E., 1908: The rainfall in Natal. *Agric. Jour., Natal,* 11, 1531-1533.

Otnes, R.K. and Enochson, 1972: *Digital Time Series Analysis*, Wiley.

Pagano, M., 1972: An algorithm for fitting autoregressive schemes. *Jour. Roy. Statist. Soc.*, C, 21, 274-281.

Panofsky, H.A. and Brier, G.W., 1965: *Some applications of statistics to meteorology*, Penn. State Univ., University Park.

Parthasarathy, B. and Dhar, O.N., 1974: Secular variations of regional rainfall over India. *Quart. Jour. Roy. Met. Soc.*, 100, 245-257.

Parzen, E., 1957: On choosing a stimate of the spectral density function of a stationary time series. *Annals Math. Stat.*, 28, 921-32.

Parzen, E., 1958: On assymptotically efficient consistent estimates of the spectral density function of a stationary time series. *Jour. Roy. Statist. Soc.*, B, 20, 303-22.

Parzen, E., 1961: Mathematical considerations in the estimation of spectra. *Technometrics*, 3, No. 2, 167-90.

Paulsmeier, H.J.C., 1946: *Is South Africa drying up?* John Meinert, Windhoek.

Peres, M.A., 1930: Preliminary investigation of the rainfall of Lourenco Marques. *A. Afr. Jour. Sci.*, 27, 132-35.

Quenouille, M., 1956: Notes on bias in estimation. *Biometrika*, 43, 353-360.

Roberts, W.O., 1963: Does variable solar activity affect stratospheric circulation? *Met. Abhandl.*, 36, 341-52, Berlin (Inst. f. Met. Geophys.der Freien Universiät).

Rubin, M.J., 1956: The associated precipitation and circulation patterns over Southern Africa. *Notos*, 5, 53-63.

Rummel, R.J., 1970: *Applied Factor Analysis*, Evanston: Northwestern Univ. Press.

Salzman, L., 1968: *Computerised Economic Analysis*, McGraw-Hill, New York.

Schneider, S.H. and Dickinson, R.E., 1974: Climate Modelling. *Rev. Geophy. and Space Res.*, Vol. 12, No. 3, 447-493.

Schulze, B.R., 1965: *General Survey, Climate of South Africa, Part 8*, Government Printer, Pretoria, 81-85.

Schumann, T.E.W. and Thompson, W.R., 1934: A study of South African rainfall, secular variations and agricultural aspects. *Univ. Pretoria Series*, No. 1, 28.

Schwartz, E.H.L., 1918: The progressive desiccation of Africa: the cause and the remedy. *S. Afr. Jour. Sci.*, 15, No. 3, 139-190.

Schwartz, E.H.L., 1919: The Kalahari Lake scheme. *S. Afr. Min. (Eng.) Jour.*, 28, Pt. II, Nos. 1433-1437.

Schwartz, E.H.L., 1920a: *The Kalahari or Thirstland Redemption*, Maskew Muller, Cape Town.

Schwartz, E.H.L., 1920b: The Kalahari and its possibilities. *Jour. Afr. Soc.*, 20, No. 77, October.

Schwartz, E.H.L., 1923: The Kalahari scheme as the solution of the South African drought problem. *S. Afr. Jour. Sci.*, 20, 208-22.

Schwartz, E.H.L., 1925: The Kalahari project. *Met. Mag.*, 60, 258, 262-65.

Schwartz, E.H.L., 1926: The Northern Kalahari. *S. Afr. Geog. Jour.*, 9, 27-36.

Schwartz, E.H.L., 1927: The Kalahari project. *Met. Mag.*, 62, p. 47.

Sellers, W.D., 1965: *Physical Climatology*, Chicago Univ. Press.

Slutsky, E., 1927: The summation of random causes as the source of cyclic processes. *Econometrica*, 5, 105-46.

South African Weather Bureau, 1965: *Climate of South Africa; Part 8, General Survey*, Pretoria, 19-28.

South African Weather Bureau, 1972: *Climate of South Africa; Part 10, District Rainfall*, Pretoria, 5-22.

Stidd, C.K., 1967: The use of eigenvectors for climatic estimates. *Jour. Appl. Met.*, 6, 255.

Stringer, E.T., 1972: *Techniques of Climatology*, W.H. Freeman, San Franciso.

Taljaard, J.J., 1955: Stable stratification in the atmosphere over South Africa. *Notos*, 4, 217.

Taljaard, J.J., 1964: Cyclogenesis, cyclones and anticyclones in the Southern Hemisphere during autumn, 1958. *Notos*, 13, 31-36.

Taljaard, J.J., 1965: Cyclogenesis, cyclones and anticyclones in the Southern Hemisphere during the period June-December, 1958. *Notos*, 14, 73-84.

Taljaard, J.J., Schmidt, W. and van Loon, H., 1961: Frontal analysis with application to the Southern Hemisphere. *Notos*, 10, 25-58.

Taljaard, J.J. and Schumann, T.E.W., 1940: Upper air temperatures and humidities at Walvis Bay, South Africa. *Bull. Meteorol. Soc.*, 21.

Taljaard, J.J. and van Loon, H., 1962: Cyclogenesis, cyclones and anticyclones in the Southern Hemisphere during the winter and spring of 1957. *Notos*, 11, 3-20.

Taljaard, J.J. and van Loon, H., 1963: Cyclogenesis, cyclones and anticyclones during summer, 1957-1958. *Notos*, 12, 37-50.

Thatcher, A.R., 1971: Year-to-year variations in the earnings of individuals. *Jour. Roy. Statist. Soc.*, A, 134, 374-82.

Thompson, W.R., 1936: *Moisture and Farming in South Africa*, Central News Agency, S. Africa.

Tintner, G., 1952: *Econometrics*, John Wiley, New York.

Tocher, K.D., 1963: *The Art of Simulation*, English Univ. Press, London.

Trenberth, K.E., 1975: A quasi-biennial standing wave in the Southern Hemisphere and interrelations with sea surface temperature. *Quart. Jour. Roy. Met. Soc.*, 101, 55-74.

Tripp, W.B., 1888: Rainfall of South Africa, 1842-1886. *Quart. Jour. Roy. Met. Soc.*, 14, 108-23.

Tripp, W.B., 1890: Rainfall of the globe: comparative chronological account of some of the principal records. *Quart. Jour. Roy. Met. Soc.*, 16, 193-205.

Tukey, J.W., 1958: Bias and confidence in not quite large samples. *Ann. Math. Statist.*, 29, 614.

Tukey, J.W., 1961: Discussion, emphasising the connection between analysis of variance and spectral analysis. *Technometrics*, 3, No. 2, 191-220.

Tylden, G., 1946: South Africa's droughts in perspective; an historical survey covering a period of 284 years. *Farmers Weekly*, 70, 1383.

Tyson, P.D., 1970: Rainfall fluctuations over South Africa during the period of meteorological records. Convention: *Water for the future*, 1-14, Pretoria.

Tyson, P.D., 1971: Spatial variation of rainfall spectra in South Africa. *Annals Assoc. Am. Geogr.*, 61, 711-20.

Tyson, P.D., 1972: Rainfall spectra and recent climatic variation in Southern Africa. *International Geography, 1972*, Toronto Univ. Press, 202-04.

Union Government, 1951: *Report of the Desert Encroachment Committee*, No. 59, 11.

Union of South Africa, 1914: *Report from the select-Committee on droughts, rainfall, and soil erosion*, The Senate, Union Parliament, S.C.2.

Van Reenen, R.J., 1923: A resume of the drought problem in the Union of S. Afr. *S. Afr. Jour. Sci.*, 21, 178-92.

Van Reenen, R.J., 1925: Note on the apparent regularity of the occurrence of wet and dry years in S.W. Africa. *S. Afr. Jour. Sci.*, 22, 94-95.

Van Rooy, M.P., 1966: Truth about the drought. *Farmers Weekly*, 110, 22-23.

Vorster, J.H., 1957: Trends in long range rainfall records in South Africa. *S. Afr. Geog. Jour.*, 39, 61-66.

Vowinckel, E., 1956: Ein Beitrag zur Witterungsklimatologie des suedlichen Mozambiquekanals. *Miscelânea Geofísica Publicada Pelo Serviço Meteorológico de Angola em Comemoração do X Aniversário do Serviço Meteorlogico Nacional*, 63-86.

Walker, G.T., 1915: Correlations in seasonal variations of weather: IV - sunspots and rainfall; V - sunspots and temperature; VI - sunspots and pressure. *Mem. of the Indian Met. Dept.*, 21, Nos. 10-12, Poona.

Wallis, A.H., 1935: Is our rainfall getting less? *The 1820 Magazine*, 6, 35-37.

Whittlesey, J.R.B., 1968: The correlational properties of random number generation for IBM 360. *Comm. Assoc. Comp. Machine 11*, 641-644.

Willett, H.C., 1964: Evidence of solar-climatic relationships. *Climate and our food supply*, 123-51, Ames, Iowa.

Willett, H.C., 1965: Solar-climatic relationship in the light of standardised climatic data. *Jour. Atmos. Sci.*, Vol. 22, No. 2, 120-136.

Willett, H.C., 1968: Remarks on the seasonal changes of temperature and of ozone in the Arctic and the Antarctic stratospheres. *Jour. Atmos. Sci.*, Vol. 25, No. 3, 341-360.

Willett, H.C., 1974: Recent statistical evidence in support of the predictive significance of solar-climatic cycles. *Mon. Wea. Rev.*

Wilson, J.F., 1865: Water supply in the basin of the River Orange or Gariep South Africa. *Jour. Roy. Geog. Soc.*, 35, 106-29.

Yule, G.U., 1926: A study in sampling and the nature of time series. *Jour. Roy. Statist. Soc.*, 89, 1-12.

Yule, G.V. and Kendall, M.G., 1968: *An Introduction to the Theory of Statistics*, Charles Griffin, London.

Footnote

* *Jackknife has alternative spellings in the literature; that due to Bissell, et al, has been used in the text.*